

Understanding the impacts of viruses on microbial methanol utilisation in seawater

Kevin Jamie Purves

Doctor of Philosophy

University of East Anglia,
Norwich, UK

School of Environmental Sciences

In collaboration with

Plymouth Marine Laboratory

September 2019

This copy of the thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that use of any information derived there from must be in accordance with current UK Copyright Law.

In addition, any quotation or extract must include full attribution.

Abstract

Methylotrophs are bacteria which utilise methanol and other one-carbon compounds for assimilative (growth) and dissimilative (energy) metabolism. Methylotrophy has been demonstrated by a large proportion of marine bacteria, influencing carbon cycling and is a significant sink of methanol in the marine environment. Viruses are understood to considerably influence biogeochemical cycles in the oceans, however the extent to which they influence methylotrophs and methanol cycling has remained unclear. Virus-like particles associated with a methanol-utilising methylotroph were isolated to this end and characterised using electron microscopy, indicating morphologies resembling enveloped viruses. A comprehensive seasonal survey in the Western Channel Observatory (WCO) combined virus abundance data with a taxonomic investigation of the microbial community; environmental variables; plankton abundance; bacterial production data and methanol uptake rates. The latter ranged between 0.1 – 10.6 nmol L⁻¹ h⁻¹ throughout the water column in the WCO with little depth variation. Seasonal trends were also consistent throughout the water column, with the highest uptake rates occurring during winter months. For the first time, seasonal virus abundance data was also determined with methanol uptake rates and bacterial production and showed a significant negative correlation with rates of methanol dissimilation. This could suggest that under high viral load bacteria are unable to utilise methanol as readily to meet their energy requirements. However, this reflects total virus abundance and not those that specifically infect methylotrophic bacteria. Furthermore, the diversity and distribution of a recently discovered alternate methanol dehydrogenase utilised by a portion of the methylotrophic community was explored. Sequencing of the *xoxF5* gene was revealed to be highly conserved within the *Rhodobacteraceae* throughout the year in the WCO and was also the dominant associated bacterial group along a transect of the Atlantic. Future studies should focus on viruses specific to known methanol utilisers to understand their role in controlling marine methanol concentrations.

Access Condition and Agreement

Each deposit in UEA Digital Repository is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the Data Collections is not permitted, except that material may be duplicated by you for your research use or for educational purposes in electronic or print form. You must obtain permission from the copyright holder, usually the author, for any other use. Exceptions only apply where a deposit may be explicitly provided under a stated licence, such as a Creative Commons licence or Open Government licence.

Electronic or print copies may not be offered, whether for sale or otherwise to anyone, unless explicitly stated under a Creative Commons or Open Government license. Unauthorised reproduction, editing or reformatting for resale purposes is explicitly prohibited (except where approved by the copyright holder themselves) and UEA reserves the right to take immediate 'take down' action on behalf of the copyright and/or rights holder if this Access condition of the UEA Digital Repository is breached. Any material in this database has been supplied on the understanding that it is copyright material and that no quotation from the material may be published without proper acknowledgement.

Contents

Abstract	2
List of figures	8
List of tables.....	12
Acknowledgements	13
1. Introduction	15
1.1. Methanol in the environment.....	15
1.1.1. The global methanol budget.....	15
1.1.2. Methanol in the oceans	16
1.1.3. Methanol production in the marine environment	19
1.2. Methylo trophy in the environment.....	19
1.2.1. Methylo trophic bacteria	19
1.2.2. Metabolism pathways of aerobic methylo trophs.....	21
1.2.3. Methanol dehydrogenase.....	22
1.2.4. A rare earth element-dependent methanol dehydrogenase	24
1.2.5. Marine methylo trophy.....	27
1.3. Viruses in the environment	31
1.3.1. An introduction to viruses	31
1.3.2. Phage in the marine environment.....	34
1.3.3. Marine biogeochemical influences	38
1.3.4. Phage of methylo trophic bacteria	41
1.4. Project aims and objectives	43
2. Materials and methods.....	46
2.1. Bacterial strains	46
2.2. Cultivation and maintenance of strains	48
2.2.1. Methylo phaga spp.....	48
2.2.2. Marinibacterium anthonyi La6	48
2.3. Timeseries sampling and process	49

2.3.1.	Supplemental timeseries and cruise data.....	50
2.4.	Isolation of methylotrophic bacteria	52
2.5.	Analytical Flow Cytometry (AFC)	53
2.5.1.	Sample preparation and storage	53
2.5.2.	Bacterial numbers analysis by flow cytometry	53
2.5.3.	Viral-like particle analysis by flow cytometry	54
2.6.	Virus Techniques	56
2.6.1.	Filtrate	56
2.6.2.	Incubation variables	57
2.6.3.	Growth of methanol-utilising methylotrophic bacteria in agar	57
2.6.4.	Plaque assay.....	57
2.6.5.	Spot testing	58
2.6.6.	Phage enrichment in liquid culture	59
2.6.7.	Prophage induction by ultraviolet, antibiotic and thermal stress	61
2.6.8.	Tangential Flow Filtration (TFF)	62
2.6.9.	PEG Precipitation and CsCl gradient purification	62
2.7.	Virus dilution and reduction experiments	63
2.7.1.	Virus reduction	63
2.7.2.	Virus dilution.....	63
2.8.	Transmission Electron Microscopy (TEM)	64
2.9.	Extraction of nucleic acids	64
2.9.1.	Environmental DNA extraction	64
2.9.2.	Genomic DNA extraction.....	65
2.9.3.	Pulsed Field Gel Electrophoresis (PFGE)	66
2.10.	Nucleic acid manipulation techniques.....	66
2.10.1.	Quantification of DNA	66
2.10.2.	Standard polymerase chain reaction (PCR).....	67
2.10.3.	Agarose gel electrophoresis.....	69
2.10.4.	DNA purification	69

2.10.5.	Sanger sequencing of PCR amplicons and analysis.....	69
2.10.6.	Primers used for Illumina sequencing	70
2.10.7.	Next Generation Sequencing (NGS) of PCR amplicons	71
2.10.8.	Bioinformatic analysis of NGS amplicons.....	71
2.11.	Microbial uptake rates	73
2.11.1.	Microbial methanol uptake rates	73
2.11.2.	Microbial methanol assimilation into biomass	73
2.11.3.	Microbial methanol dissimilation to carbon dioxide (CO ₂)	74
2.11.4.	Determination of methanol uptake rates	75
2.12.	Leucine incorporation rates	76
2.12.1.	Determination of bacterial leucine incorporation rates	77
2.12.2.	Conversion factors and bacterial production rates	77
2.12.3.	Enzyme kinetics	78
2.13.	Statistical methods	80
2.13.1.	Data pre-treatment.....	80
2.13.2.	BEST: BIO-ENV and BV-STEP.....	81
2.13.3.	Non-metric multidimensional scaling (NMDS).....	81
2.13.4.	RELATE	81
3.	Virus-like particles associated with an obligate methylotroph	83
3.1.	Introduction	83
3.2.	Isolation of methanol-utilising methylotrophic bacteria.....	85
3.3.	Virus-like particle (VLP) abundance.....	89
3.3.1.	Abundance of virus-like particles from environmental samples.....	91
3.4.	Growth Rates.....	93
3.5.	Phage Inoculation	94
3.5.1.	Growth of methanol-utilising methylotrophic bacteria in agar	94
3.5.2.	Plaque assay.....	95
3.5.3.	Phage enrichment and spot testing	96
3.6.	Transmission Electron Microscopy (TEM)	103

3.7.	Purification and nucleic acid extraction of particles.....	105
3.8.	Prophage Induction.....	106
3.8.1.	UV Irradiation	110
3.8.2.	Mitomycin C	110
3.8.3.	Heat Shock.....	111
3.9.	Discussion	113
3.9.1.	Isolation of methanol-utilising methylotrophs	113
3.9.2.	VLPs and isolation of methylotrophic phage.....	114
3.9.3.	A potential chronic infection system	116
3.10.	Conclusion.....	118
4.	Seasonal variation in microbial utilisation of methanol in the Western English Channel.....	120
4.1.	Introduction	120
4.2.	Results.....	121
4.2.1.	Microbial methanol uptake	121
4.2.2.	Virioplankton abundances	125
4.3.	Biogeochemical and abundance dynamics.....	129
4.3.1.	Biogeochemical properties	129
4.3.2.	Nutrients.....	131
4.3.3.	Nano- and picoplankton populations	136
4.3.4.	Bacterial abundance and production	140
4.3.5.	Virus-to-bacteria ratios	145
4.4.	Microbial diversity	149
4.4.1.	16S rRNA bacterial diversity	149
4.4.2.	xoxF5 bacterial diversity.....	156
4.5.	Discussion	161
4.5.1.	Environmental drivers of methanol utilisation in the Western Channel Observatory.....	161
4.5.2.	Microbial community in the Western Channel Observatory.....	173

4.6. Conclusion	179
5. Basin scale variability of microbial and <i>xoxF5</i> diversity	182
5.1. Introduction	182
5.2. Results.....	185
5.2.1. Arctic biogeochemical properties	185
5.2.2. Microbial diversity in Arctic and Atlantic regions.....	189
5.2.3. Variability in <i>xoxF5</i> gene diversity	194
5.3. Discussion	196
5.4. Conclusion	201
6. Synthesis and conclusions	203
6.1. Virus-like particles associated with an obligate methyloph	203
6.2. Viruses and methanol uptake rates in seawater	205
6.3. <i>XoxF5</i> diversity and distribution in the marine environment.....	207
List of abbreviations	210
References.....	212

List of figures

Figure 1.1 Simplified overview of different sources and sinks of methanol.	16
Figure 1.2 Monthly averaged surface concentrations between January 2011 – March 2012 at Station L4 for methanol.	17
Figure 1.3 Major methylotrophic modules involved in the degradation of different C1 compounds.	22
Figure 1.4 Simple structure of methanol dehydrogenase, <i>MxaFI</i>	23
Figure 1.5 Phylogenetic tree of PQQ-dependent dehydrogenases	26
Figure 1.6 Figure of the cycling of volatile organic compounds (VOCs) between phytoplankton and bacteria	29
Figure 1.7 Types of modes of infection utilised by viruses, indicating the various major steps involved or that can occur.....	33
Figure 1.8 Schematic of the stages of a lytic infection upon a bacterial cell.....	33
Figure 1.9 Bacterial resistance observed at the different phage infection stages.	36
Figure 1.10 Flow of cellular material released into the environment by microbes upon lysis.....	40
Figure 1.11 Different virus morphologies isolated from methanotrophic bacteria	42
Figure 2.1 Map of the coast of Plymouth highlighting the sampling stations of the Western Channel Observatory.....	49
Figure 2.2 Flow diagram of the processing of collected seawater used in the timeseries analysis.....	50
Figure 2.3 Flow cytometric distribution of virus subgroups of a natural seawater sample	55
Figure 2.4 Flow visualisation of the preparation of seawater for use as filtrate in experiments.	56
Figure 2.5 Flow diagram of the various phage enrichment methods utilised in experiments.	60
Figure 2.6 A model Lineweaver-Burk transformation plot and a Lineweaver-Burk transformation plot of methanol dissimilation rates at station L4.....	80
Figure 3.1 Agar plates spread with methanol seawater enrichments and incubated for one week at 25°C	86
Figure 3.2 Phylogenetic analysis of <i>16S rRNA</i> gene sequences of L4 isolates and other sequences of members of the genus <i>Methylophaga</i>	88
Figure 3.3 Comparison of virus-like particle groupings in unfiltered surface seawater collected from station L4 and filtrate from three different filters.....	90

Figure 3.4 Flow cytometric plots of samples passed through SFCA and GFCA filter, with gates indicating VLP 1, 2 and 3	90
Figure 3.5 Abundance of total virus-like particles and the three sub fractions in samples collected from the different depths and sediment pore water	92
Figure 3.6 Percentages of individual VLP subgroups which make up the total VLP abundance for each of the samples analysed.....	92
Figure 3.7 Optical density comparison of the growth rates of <i>M. marina</i> and the enrichment isolates from station L4	94
Figure 3.8 Agar plates where cells have been added into the soft agar and where no cells have been added to indicate the differences in turbidity.....	96
Figure 3.9 Sectioned plate from a spot test, indicating regions of no growth within different sections, where filtrate has been spotted on to the top agar.....	98
Figure 3.10 Abundance of VLP 1 and 2, and the optical density (OD) in a phage enrichment and a positive control enrichment using <i>M. marina</i>	99
Figure 3.11 Abundance of VLP 1, 2 and 3, and bacteria in a <i>M. marina</i> phage enrichment	100
Figure 3.12 Results of phage enrichment experiment 3, over 36 hours, utilising a sacrificial approach for each sampling time point.....	102
Figure 3.13 Electron micrograph of virus-like particles.....	104
Figure 3.14 PFGE agarose gel after 20 hours, visualised with UV after an ethidium bromide bath.....	106
Figure 3.15 Annotation of the intact prophage sequence, identified within the genome of <i>Marinibacterium anthonyi</i> La6 by PHASTER	108
Figure 3.16 Phylogenetic analysis of the phage-like protein sequences and integrase identified within the <i>M. anthonyi</i> La6 genome	109
Figure 3.17 Abundance of total VLP, VLP 1, 2 and 3, and the bacterial cell abundance in a culture treated with mitomycin c	111
Figure 3.18 Initial bacterial cell abundance and VLP1, 2 and 3 abundance results from a heat shock experiment of <i>M. anthonyi</i> La6 and <i>M. marina</i>	112
Figure 3.19 Total VLP and bacterial cell abundance in an a) <i>M. marina</i> culture shocked thermally at 42°C and b) a negative control culture.	113
Figure 3.20 Percentage plots of virus groups VLP 1, 2 and 3 derived from samples taken from a) the <i>M. marina</i> culture thermal shock experiment	113
Figure 4.1 Temporal variation of methanol uptake rates throughout the water column and the bacterial growth efficiency for methanol at station L4.....	123

Figure 4.2 Temporal variation of methanol uptake rates and bacterial growth efficiency for methanol throughout the water column at station E1	123
Figure 4.3 Abundances of total virus-like particles, and virus subgroups through the water column at station L4	126
Figure 4.4 Abundances of total virus-like particles, and virus subgroups at the surface and 60 m at station E1	128
Figure 4.5 Depth profile of seawater temperature, salinity and fluorescence values through the water column at station L4	130
Figure 4.6 Depth profile of seawater temperature, salinity and fluorescence values through the water column at station E1	130
Figure 4.7 Nutrient concentrations and the N:P ratio through the water column at station L4	133
Figure 4.8 Nutrient concentrations and the N:P ratio through the water column at station E1	135
Figure 4.9 Abundances of <i>Synechococcus</i> , picoeukaryotes, nanoeukaryotes, coccolithophores, <i>Phaeocystis</i> sp. and cryptophytes at station L4.	137
Figure 4.10 Abundances of <i>Synechococcus</i> , picoeukaryotes, <i>Phaeocystis</i> sp., nanoeukaryotes, cryptophytes and coccolithophores at station E1	139
Figure 4.11 Abundances of total heterotrophic bacteria, high nucleic acid- and low nucleic acid-containing bacteria, bacterial production and bacterial production per cell through the water column at station L4	142
Figure 4.12 Abundances of total heterotrophic bacteria, high nucleic acid- and low nucleic acid-containing bacteria, bacterial production and bacterial production per cell through the water column at station E1	144
Figure 4.13 Total virus-to-bacteria ratio, VLP 1-to-bacteria ratio, VLP 2-to-bacteria ratio, VLP 3-to-bacteria ratio throughout the water column at station L4	146
Figure 4.14 Total virus-to-bacteria ratio, VLP 1-to-bacteria ratio, VLP 2-to-bacteria ratio, VLP 3-to-bacteria ratio at the surface and 60 m at station E1	148
Figure 4.15 Relative abundance (>2%) of the bacterial 16S rRNA community at the surface and 50m depths of station L4 at different taxonomic levels	152
Figure 4.16 Relative abundance (> 2%) of the bacterial 16S rRNA community at the surface and 60m depths of station E1 at different taxonomic levels.	155
Figure 4.17 Relative abundance (>2%) of the bacterial <i>xoxF5</i> community at the surface and 50m depths of station L4 at different taxonomic levels	159

Figure 4.18 Relative abundance (>2%) of the bacterial <i>soxB5</i> community at the surface and 60m depths of station E1 at different taxonomic levels.	160
Figure 4.19 Non-metric multidimensional scaling ordination of (a) transformed surface environmental variables from station L4 and E1, and (b) surface L4 abundance data for the total VLP abundances with methanol dissimilation, assimilation and bacterial production rates for each sample.....	167
Figure 4.20 Schematic indicating the environmental and abundance factors which best describe seasonal variability in methanol uptake rate at station L4.....	172
Figure 5.1 A map of sampling station locations in the Arctic Ocean (JR16006) overlaid with three transect regions and the sampling points from Atlantic cruise (JC039) overlaid with different regions.....	184
Figure 5.2 Measurements of temperature, chlorophyll-a, salinity and total bacterial abundance at the surface of the sampled stations from cruise JR16006.....	186
Figure 5.3 Concentrations of ammonium, phosphate, silicate, nitrate and nitrite at the surface of the sampled stations from cruise JR16006.	188
Figure 5.4 Relative abundance (> 2%) of the bacterial community (<i>16S V4 rRNA</i>) in surface samples of (a) the Arctic and (b) Atlantic Ocean at the Class level	190
Figure 5.5 Relative abundance of the <i>soxB5</i> community in surface samples of the Atlantic Ocean at the (a) Class and at the (b) Family level	195
Figure 6.1 Schematic indicating how the phage enrichment approach could be better optimised to increase the probability of infecting a lab strain with a lytically infecting virus	204
Figure 6.2 Updated simplified overview of different sources and sinks of methanol in the environment.	205
Figure 6.3 Simple schematic overview of the seasonal variation in methanol uptake rates, virus abundance and the bacterial production.....	206

List of tables

Table 2.1 List of organisms used in this study	46
Table 2.2 All media components used when making the individual mediums.	46
Table 2.3 All media reagents used when making the individual mediums.....	47
Table 2.4 PCR primers used in this study, and the positive control species used for each of the genes.	68
Table 2.5 Indicates the sequence that was attached to the 5' end of all primer sets used.	70
Table 2.6 Spacer sequences added in between Illumina adapter sequences and the specific gene primer sets.	71
Table 2.7 V_{max} and K_m values of methanol dissimilation rate kinetic experiments conducted at station L4 in the Western Channel Observatory.	79
Table 3.1 The enrichment conditions that each strain was cultivated from, phylogeny, and some basic characterisation of the strains isolated from station L4 in the Western Channel Observatory	87
Table 3.2 Different conditions and locations of filtrate that were obtained, which resulted in regions of no growth	98
Table 4.1 Pearson correlation coefficient values between microbial methanol dissimilation and assimilation at station L4 and E1 in the WCO with environmental variables	168

Acknowledgements

I would like to give a massive thanks to Dr Jo Dixon, Prof Willie Wilson, Prof Colin Murrell and Dr Ben Temperton, with an additional special appreciation for Dr Susan Kimmance – thank you for the fantastic supervision and support whilst at PML and UEA, you all taught me so much and I'm proud to have learned from you all! I want to also thank the NERC and the EnvEast doctoral training partnership for funding my PhD studies and for the many opportunities and friendships gained through training and social events, hearing your problems and listening to mine always made me feel better and reassured.

A special thanks must be reserved for the crew of the RV *Plymouth Quest* for sampling constantly throughout the year and to the biogeochemistry group for coffee and cakes. To the Allen group – I know I wasn't one of you, but I always felt included and you constantly helped me out by talking through my PhD (specifically Mike), getting help in the lab (especially Tracey – you are the best and I wish I had your brain!) and for involving me in your group meals. I also must thank my 2015 cohort (particularly Rachel, Becca and Jo), as I was the youngest and only boy you took me under your wings like a little brother and kept me sane throughout. This has also to be said for Kieran, Rich and Kristian (thank you for including me in everything in those first years) and also to the two Chris' for making me laugh throughout the day at work and whilst bouldering! To everyone in the PML PhD office, some of you are wacky (I miss your punishing puns, Paul), some of you are serious (sorry Chris...) and some of you are brutally hysterical (that's you Dan), you made that office feel like a home and should never change.

Finally the big ones, thank you to my friends, amazing mum and dad and incredible siblings for always being there to chat about anything and everything, for giving me the strength to finish my thesis, for keeping me grounded and for always, always making me cry with laughter whenever I came home to Scotland! To my partner Jenny, you came into my life halfway through my PhD and I honestly do not know how I was ever planning to finish it without you around. You always listen to me, laugh with me and love me. I can only hope that I can support you through your time in Dublin half as well as you supported me through mine in Plymouth.

Chapter 1

Introduction

1. Introduction

1.1. Methanol in the environment

1.1.1. *The global methanol budget*

Methanol is an oxygenated volatile organic compound (OVOC), which is pervasive throughout the atmosphere and is the second most abundant organic gas after methane (Lewis *et al.*, 2005). As a biogeochemically active compound, it controls tropospheric ozone in the presence of nitrogen oxides via methanol oxidation (Singh *et al.*, 2000; Tie, 2003), subsequently influencing the oxidising capacity of the atmosphere. It has also been associated with the formation of acid rain, due to photochemical reactions converting methanol into formic acid (Heikes, 2002).

Early estimates placed the global source inventory of methanol to be approximately 122 Tg yr^{-1} (Singh *et al.*, 2000). Where biogenic sources (75 Tg yr^{-1}), methane oxidation (18 Tg yr^{-1}) and decaying plant matter (20 Tg yr^{-1}) were determined to be the greatest contributors to the global methanol budget (Figure 1.1). These biogenic sources are primarily derived from plant growth, where plants will produce methanol as a by-product of restructuring pectin to increase cell wall stability through the use of pectin methyl esterase enzymes, which liberate methoxyl groups that then convert to methanol (Oikawa *et al.*, 2011; Oikawa and Lerdau, 2013). However, although biogenic sources were identified as the primary methanol source and thereby dictate the total budget, this was based on poorly quantified estimates (ranging from between 50 and 125 Tg yr^{-1}) (Singh *et al.*, 2000). Additional estimates have placed the global methanol emissions at between 104 Tg yr^{-1} and 312 Tg yr^{-1} (Galbally and Kirstine, 2002; Guenther, 2002). This large variability was likely driven by a lack of studies and the difficulties involved in accurately determining methanol concentrations (Wohl *et al.*, 2019), which in turn has prompted variability in methanol source and sink estimations (Figure 1.1) (Dixon, Beale and Nightingale, 2013). Regardless, Singh *et al.* (2000) noted that the total estimated methanol sink strength ($40 - 50 \text{ Tg yr}^{-1}$) was far exceeded by the global known sources and that additional removal mechanisms, other than hydroxyl oxidation and slow deposition, must also exist and play an important role. Furthermore, Singh *et al.* (2000) indicated that, based on unpublished data from the tropical pacific in 1996, the ocean could be a

source or a sink of methanol and additional research was vital for understanding the role of the ocean on the global methanol budget.

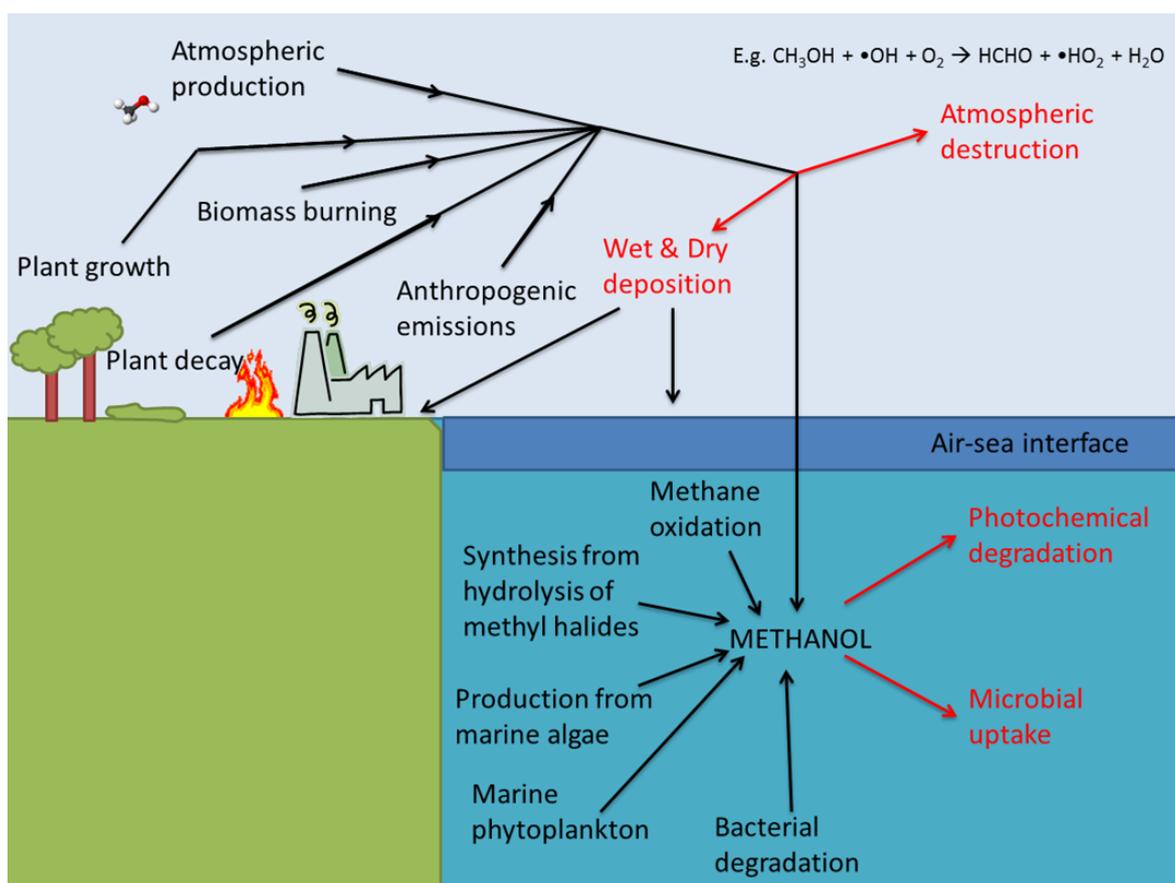


Figure 1.1 Simplified overview of different sources and sinks in the environment. Potential sources are indicated by black lines, whereas potential sinks are represented by red lines. Figure taken from Sargeant (2013).

1.1.2. Methanol in the oceans

A revised methanol inventory of sources and sinks (sources totalling 345 Tg yr⁻¹ and sinks totalling 270 Tg yr⁻¹), again indicated biogenic sources as the principal source (> 80%) (Heikes, 2002). This was one of the first attempts at trying to answer the implication of the marine environment on the global methanol budget, and included an argued for gross emission of 30 Tg methanol yr⁻¹ (Heikes, 2002). This emission rate was chosen to satisfy the model and had a range of 0 – 80 Tg yr⁻¹, where the lower estimate stemmed from the assumption that methanol is readily dissolved in seawater and consumed by organisms, and the upper limit from the maximum flux needed to maintain observed marine boundary layer methanol mixing ratios (Heikes, 2002). Heikes (2002) was intrigued by an implied flux of methanol from the atmosphere to the ocean, primarily by dry deposition (80 Tg yr⁻¹), stating that this

extensive source could be sustaining bacteria that utilise methanol as a carbon source in the marine environment implying an additional sink. Further stoking the uncertainty as to the oceans status as a net source or sink of methanol. Especially as support existed for the biological formation of methanol in the ocean originating from unpublished data, in which methanol was observed in the headspace of phytoplankton cultures (Heikes, 2002), which has been reinforced by a more recent study, which will be discussed in more detail (Mincer and Aicher, 2016).

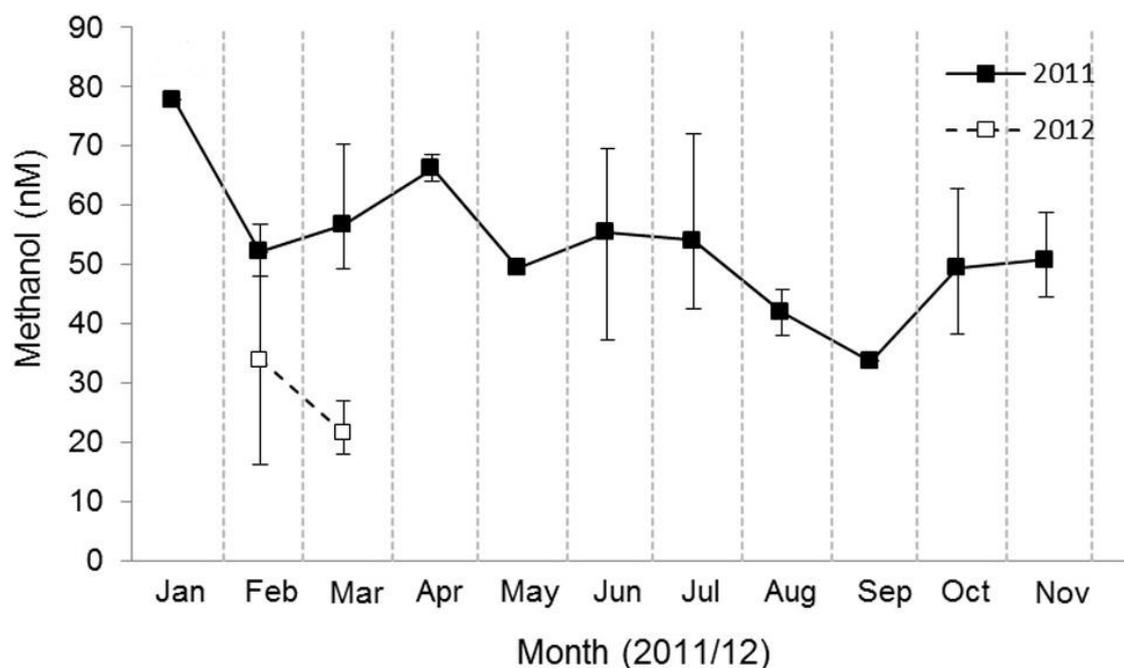


Figure 1.2 Monthly averaged surface concentrations for January 2011 – March 2012 at Station L4 for methanol, which includes error bars representing range of measurements recorded. Figure adapted from Beale *et al.* (2015).

The capability to study compounds in the marine environment at the nano-molar range was a technological advancement which would yield further insights in to the global methanol budget, such as the finding that methanol and other OVOCs are ubiquitous throughout the oceans (Beale *et al.*, 2011). The first such methanol measurements indicated concentrations of 118 nM in the Atlantic (Williams, 2004), and more recent studies indicated inter-annual and seasonal variation and methanol concentrations as high as 78 nM in coastal waters Figure 1.2) (Beale *et al.*, 2015) and a range of 40 – 360 nM across a transatlantic transect (Beale *et al.*, 2013). Flux measurements of the organic species between the atmosphere and the seawater surface were also investigated, with methanol determined to have a flux of -33.7

$\mu\text{mol}^{-1} \text{ m}^{-2} \text{ day}^{-1}$ to the ocean (Williams, 2004). This was indicative of a methanol sink, however positive fluxes were also reported in other ocean regions at different times of the year (Duncan, 2003; Williams, 2004). However, this was significant, as it reinforced theories regarding methanol being metabolised by microbial communities.

In 2011, the first methanol loss rates to the ocean were reported with biological methanol oxidation rates of between $2.1 - 8.4 \text{ nmol L}^{-1} \text{ day}^{-1}$ in northeast Atlantic surface water - with predictions that these rates could reach $29 \text{ nmol}^{-1} \text{ L}^{-1} \text{ day}^{-1}$ in more productive coastal waters (Dixon, Beale and Nightingale, 2011). Dixon et al. (2011a) was the first to show that methanol uptake rates can be inferred by exploiting ^{14}C -labelled methanol to investigate the oxidation of methanol to carbon dioxide (dissimilation) and the uptake into biomass (assimilation) by measuring the ^{14}C radioactive decays with a liquid scintillator. Further methanol uptake rates were also determined in coastal regions, where they ranged between $0.7 - 11.2 \text{ nmol L}^{-1} \text{ h}^{-1}$, reflecting the more productive nature of shelf seas (Sargeant *et al.*, 2016). Dixon et al. (2011a) also demonstrated that methanol was predominantly used as an energy source by microbes (>99%). This was further demonstrated by Sargeant (2013) in coastal and Atlantic samples, where assimilation rates were significantly smaller than dissimilation rates. This is significant for understanding methylotrophic bacteria, as this indicates that methylotrophs instead use further reduced C_1 compounds for assimilation into the cell biomass (Dixon et al. 2011a; Murrell & McDonald 2000). The main loss process present in the marine environment is believed to be by microbial uptake (Heikes, 2002), with methanol turnover times ranging between 1 and 5 days (Millet *et al.*, 2008; Dixon, Beale and Nightingale, 2011). However, Beale *et al.* (2015) determined that the sea surface layer of the Western Channel Observatory (WCO; Plymouth, UK) was under-saturated when compared with the overlaying tropospheric methanol concentration. Combined with microbial methanol uptake rates, it was estimated that the air-sea flux of methanol only contributed 2 – 20% of the total microbial oxidation, implying that the atmosphere is unlikely to be a dominant source of methanol (at the WCO), and that there must be an *in situ* source of methanol to sustain the observed uptake rates (Beale *et al.*, 2015).

1.1.3. Methanol production in the marine environment

As mentioned already, there was indirect evidence to suggest that methanol was being produced within the marine environment, although not measured directly. The most direct results included methanol concentrations increasing around intact macro-algae, when compared to the surrounding seawater (Nightingale, 1991), and concentrations of methanol detected in the headspace of phytoplankton cultures (Riemer, 2000). The hypothesised production of methanol by phytoplankton was eventually determined by Mincer and Aicher (2016), who found that cultures of *Rhodomonas salina*, *Nannochloropsis oculata*, *Phaeodactylum tricornutum*, *Emiliania huxleyi*, *Trichodesmium erythraeum*, *Synechococcus* sp. and *Prochlorococcus marinus* all produced methanol (0.8 – 13.7 μM) after 10 to 24 days. This study was notable as *Synechococcus* and *Prochlorococcus* sp. are known for their vast distribution and abundance in the marine environment, and using conservative estimates of their abundance in the oceans, indicated that phytoplankton sources could be comparable with terrestrial sources of methanol (Flombaum *et al.*, 2013; Mincer and Aicher, 2016). Another proposed source of methanol is chemical synthesis by the hydrolysis of methyl halides in seawater (Elliott and Rowland, 1995), although this process would occur at a relatively slower rate and therefore global production from this source would be comparatively smaller.

1.2. Methylytrophs in the environment

1.2.1. Methylytrophic bacteria

Methylytrophs are organisms which utilise reduced carbon compounds with no carbon-carbon bonds as a primary source of carbon and energy (such as methanol or methane), and must make all carbon-carbon bonds *de novo* (Anthony, 1982). These organisms have been studied to determine their impacts on various important environmental compounds and their cycles such as nitrogen, carbon and sulphur, and also compounds relevant to the climate, including methane and methanol (Boden *et al.*, 2010; Chistoserdova, 2019). Methane can be utilised by methanotrophic methylytrophs as a sole source of carbon, which is oxidised to

methanol (Figure 1.3), and are also capable of efficient growth solely on methanol as a carbon source (Anthony, 1982; Leak and Dalton, 1986; Chistoserdova, 2011). Interest in methylotrophs has driven biotechnological applications for industry where carbon compounds can be reduced and yield a range of industrial products (Kelly, Ardley and Wood, 2015). *Methylorubrum extorquens* alone has successfully been used to produce amino acids, single cell protein, insecticides, bacteriocin and fluorescent protein, with other potential products also determined (Ochsner *et al.*, 2015).

Methylotrophs were initially isolated from the terrestrial environment, such as soil and plants, however have since been indicated as a diverse group of organisms, detectable in a range of environment, from freshwater and wetlands to sediment and marine settings, and have even been isolated from volcanic mudpots and hot springs, indicating an extremophilic nature (Radajewski *et al.*, 2002; Neufeld *et al.*, 2008; Semrau, DiSpirito and Murrell, 2008; Moussard *et al.*, 2009; Pol *et al.*, 2014; Iguchi, Yurimoto and Sakai, 2015). There are many methylotrophic bacteria which have been identified and characterised, with well over 200 described species, belonging mostly to the *Alpha-*, *Beta-* and *Gammaproteobacteria*, however some bacteria also belong to *Verrucomicrobia*, *Actinobacteria*, *Firmicutes* and *Bacteroidetes* (Madhaiyan *et al.*, 2010; Kolb and Stacheter, 2013; Taubert *et al.*, 2015 and references therein). Additionally, knowledge of methylotrophy pathways has increased and led to proposals that the extent of methylotrophy in the environment may be greater than previously speculated (Kalyuzhnaya, Hristova, *et al.*, 2008; Beck *et al.*, 2015; Taubert *et al.*, 2015).

Methylotrophs can be broadly split into two groups: obligate and facultative methylotrophs. Where obligate methylotrophs grow specifically on reduced one-carbon (C1) compounds, and facultative methylotrophs are capable of growth on additional organic multi-carbon compounds (Chistoserdova, 2011). However, although methylotrophs which utilise methane (methanotrophs) are generally obligate, most known methanol-utilising methylotrophs are mainly facultative and aerobic (Trotsenko and Murrell, 2008; Kolb, 2009).

1.2.2. *Metabolism pathways of aerobic methylotrophs*

The last century has involved a host of research in to attempting to understand the diverse metabolic pathways which are employed by aerobic methylotrophs, with almost half a century taken to close just one pathway (Anthony, 2011). To simplify the processes involved in most methylotrophic pathways (Figure 1.3), they can be broadly broken down into three main phases:

1. Oxidation of one-carbon substrates to formaldehyde (HCHO)
2. Oxidation of formaldehyde
3. Incorporation of one-carbon units into cell biomass

The incorporation of one-carbon units into cell biomass can occur at the formaldehyde step for the ribulose monophosphate (RuMP) cycle and at the level of methylene-H₄F and the CO₂ via the serine cycle (Heptinstall and Quayle, 1970). It can also happen at CO₂ step via the Calvin-Benson-Bassham/ribulose biphosphate (RuBP) cycle (Bassham, Benson and Calvin, 1950). The RuBP Cycle is utilised by relatively few methylotrophs and was initially identified in photosynthetic autotrophs, which explains its utilisation to oxidise CO₂ (Anthony, 1982). The RuMP cycle oxidises formaldehyde in an effort to assimilate carbon and following oxidation, it converts glucose 6-phosphate into pyruvate for cell biosynthesis (Anthony, 1982). Four variants of RuMP do exist and are each used to varying extents by different methylotroph groups. However, the 2-keto 3-deoxy 6-phosphogluconate (KDPG) aldolase variant and the fructose 1,6-bisphosphate fructose bisphosphate aldolase bisphosphate variant are the two most common RuMP variants (Strøm, Ferenci and Quayle, 1974). The KDPG variant is used by obligate methylotrophs and the second variant is used by facultative methylotrophs, obligate methylotrophs unable to use methane, obligate methanotrophs and facultative methylotrophs unable to use methane (Anthony, 1982). An additional pathway is the serine pathway (Heptinstall and Quayle, 1970), which is specific to methylotrophs and synthesises phosphoglycerate from formaldehyde and CO₂, but does utilise a number of enzymes from other metabolic pathways (Chistoserdova, 2011).

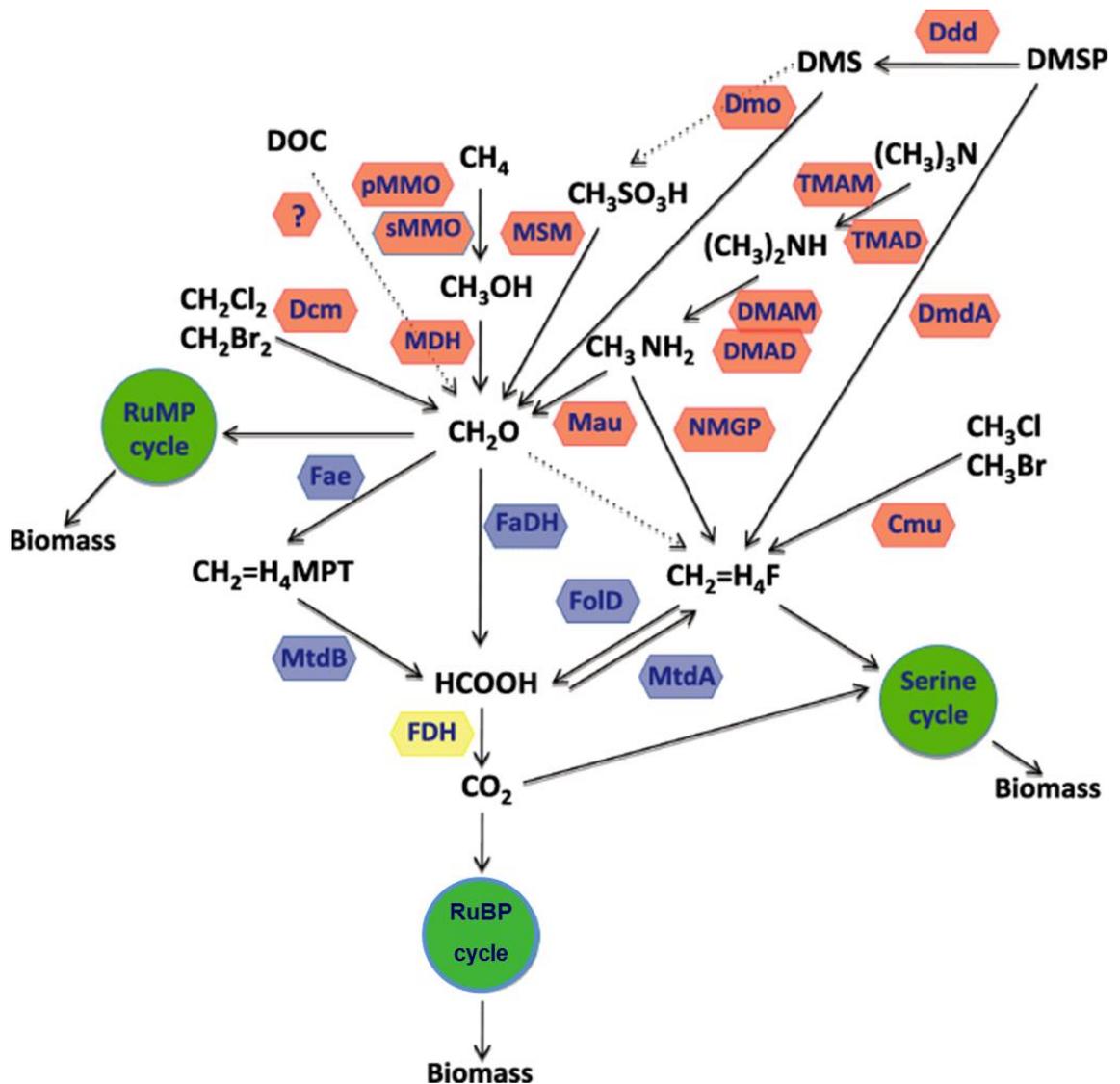


Figure 1.3 A simplified diagram depicting major substrates and intermediates and major methylytrophic metabolic modules. Red boxes indicate primary oxidation, blue boxes indicate formaldehyde handling modules, formate dehydrogenase module is shown in yellow and the assimilation modules are shown in green circles. Dashed lines are used to depict a lack of biochemical knowledge or non-enzymatic reactions (or both). Adapted from Chistoserdova (2011).

1.2.3. Methanol dehydrogenase

To facilitate the oxidation of methanol to formaldehyde, methylytrophic microorganisms utilise a range of different methanol dehydrogenase (MDH) enzymes. Three of the main MDH found in methylytrophs are the: FAD-containing alcohol oxidase; NADP⁺ dependant MDH and the pyrroloquinoline quinone (PQQ)-dependent MDH, which are generally specific to yeast, Gram-positive bacteria and

Gram-negative bacteria, respectively (Anthony and Zatman, 1964; van Dijken, Otto and Harder, 1976; Arfman *et al.*, 1989).

The originally described PQQ-dependent MDH enzyme was first discovered in 1964 by Anthony and Zatman (1964a), and was found to be prevalent in all methylotrophs which had been identified up to the early 1980's (Anthony, 1982). Eventually it was characterised as a heterotrimer (a protein containing 4 non-covalently bound, dissimilar subunits) made up of two large (66 kDa) α -subunits and two smaller (8.5 kDa) β -subunits, a PQQ prosthetic group and a calcium ion (Ca^{2+}) (Anthony and Williams, 2003). This enzyme had an intricate radial symmetrical stereostructure similar to that of an 8-winged propeller, with the 'prop-shaft' consisting of the PQQ molecule and a calcium ion (Figure 1.4) (Anthony and Zatman, 1967; Ghosh *et al.*, 1995; Trotsenko and Murrell, 2008).

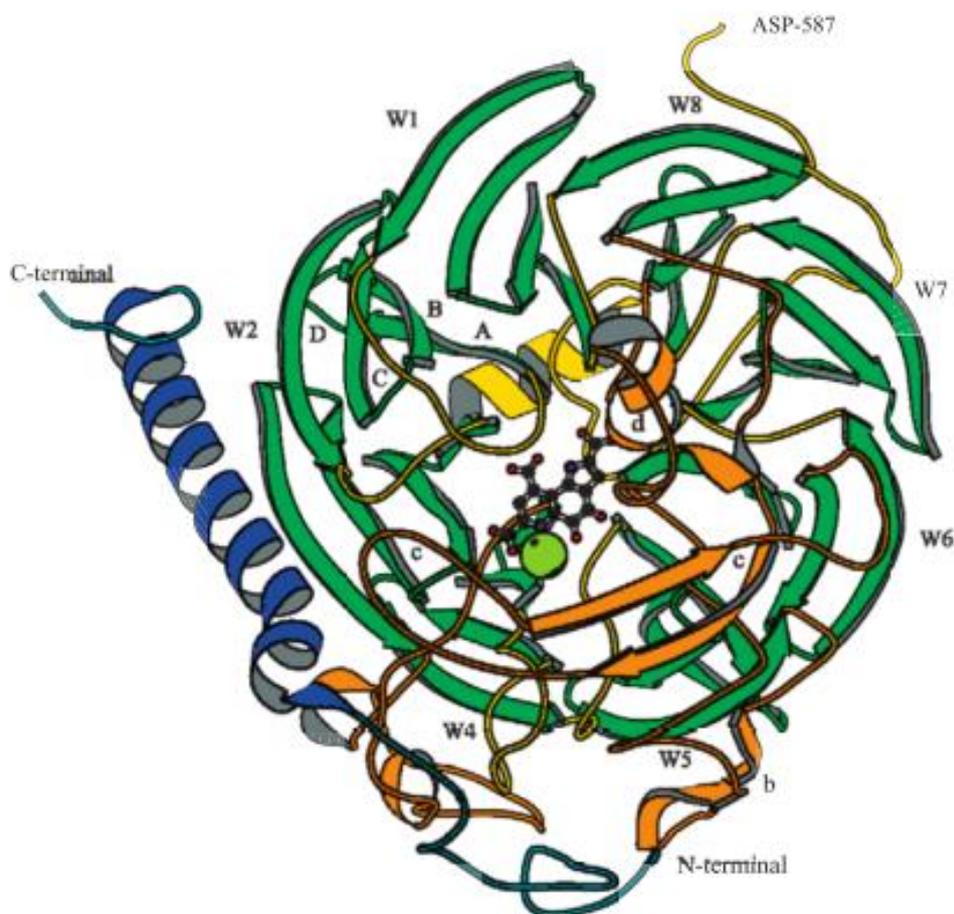


Figure 1.4 One of two $\alpha\beta$ units of methanol dehydrogenase simplified to show the sub unit structures and positioning within the protein. The PQQ prosthetic group is represented in skeletal form in the centre of the structure, with the calcium ion represented as a green sphere. Adapted from (Anthony and Williams, 2003).

PQQ was originally named methoxatin by Salisbury et al. (1979), but was renamed to its current more informative name to emphasise the importance of the proteins fused quinoline and pyrrole ring structure (Anthony and Williams, 2003). As other dehydrogenases had also been characterised with a prosthetic group with similar properties and concluded to be PQQ, the name 'quinoproteins' was therefore coined for this distinct class of PQQ-dependent dehydrogenases (Duine, Frank and van Zeeland, 1979). The prosthetic group in PQQ-dependent MDH is reduced to PQQH₂ during the oxidation of methanol, resulting in the release of formaldehyde, and is then re-oxidised and further reduced (Anthony and Ghosh, 1998). According to Kist & Tate (2013), it is this exchange of electrons that is the catalyst for the oxidation of methanol and the purpose of the calcium ions is to act as a stabilisation mechanic for PQQ by forming a bridge between the quinoprotein and the enzyme, allowing the reaction to continue by supporting the reduction (White *et al.*, 1993). The two different subunits making up the MDH protein are coded for by the *mxoF* gene (large α -subunit) and the *mxoI* gene (smaller β -subunit), additionally is the *mxoG* gene coding for cytochrome c_L (which acts as the primary electron acceptor), however the *mxoF* genes have been termed as the 'essential genes' for methanol oxidation to occur, out of the total 17 *mxo* genes involved in methanol oxidation to formaldehyde (Nunn and Lidstrom, 1986a, 1986b; Lidstrom *et al.*, 1994).

1.2.4. A rare earth element-dependent methanol dehydrogenase

Over 20 years ago, the XoxF protein was identified in *Methylobacterium extorquens* AM1, which was determined to have a high similarity to the MxoF subunit, however the function of XoxF was unclear (Chistoserdova and Lidstrom, 1997). After the genome of *M. extorquens* AM1 was sequenced, two additional prospective MDHs were identified (*xoxF1* and *xoxF2*), with high sequence identity to each other, but only 50% amino acid similarity to MxoF (Vuilleumier *et al.*, 2009). Many conflicting studies attempted to further elucidate the role of *xoxF* (mainly by gene deletion studies), until the first evidence of XoxF as a methanol dehydrogenase was determined using a deletion mutant of *Rhodobacter sphaeroides* and resulted in an inability to grow on methanol (Wilson, Gleisten and Donohue, 2008). Further studies corroborated these findings, and eventually the proposed function of *xoxF* was as an environmental signal for regulation of methanol oxidation genes, and additionally

that only the XoxF large subunit protein is required for growth on methanol in *M. extorquens* AM1 (Skovran *et al.*, 2011).

In 2011, a research group from Gifu University, Japan, carried out research investigating the importance of rare earth elements (REE) on MDH activity and found that lanthanum (La^{3+}) is an important REE for methanol oxidation (Fitriyanto *et al.*, 2011; Hibi *et al.*, 2011). Following on from this work, Nakagawa *et al.* (2012) determined that the post-translational activation of *xoxF* and *mxoF* require La^{3+} and Ca^{2+} , respectively. Furthermore, due to the presence of La^{3+} in plants (0.178 - 3.1 $\mu\text{g/g}$; Nakagawa *et al.* 2012), *M. extorquens* AM1 was capable of generating MDH using the *xoxF* or *mxoF* route, based on which REE was present (Nakagawa *et al.*, 2012). Importantly was the implied connection to the marine environment, as ~35 $\mu\text{g/g}$ of La^{3+} is present in the Earth's crust (Nakagawa *et al.*, 2012) and could therefore also be present in the marine environment. The emerging story of the function of *xoxF* was further enhanced by the isolation of the acidophilic methanotroph *Methylacidiphilum fumariolicum* SolV from a volcanic mudpot, which was unable to grow on methane without the metal-rich mudpot water from which it was isolated (Pol *et al.*, 2014). The limiting factors within the mudpot water was determined to be the lanthanides present, which were cofactors in the only MDH, XoxF, within the genome of *M. fumariolicum* SolV.

Further indications of the importance of *xoxF* genes to methylotrophic bacteria came again from work involving *M. extorquens* AM1, and *Methylomicrobium buryatense* which showed that the transcription of *mxoF* was downregulated and that *xoxF* could be upregulated at higher lanthanum concentrations (>100 nM; Chu, Beck and Lidstrom, 2016; Vu *et al.*, 2016). Various additional *xoxF* sequences have been identified from sequenced genomes, and phylogenetic analysis revealed that *xoxF* sequences cluster into 5 different clades (Figure 1.5; Chistoserdova, 2011). The *mxoF* sequences clustered in between two of the *xoxF* clades, suggesting that *mxoF* could possibly be derived from *xoxF* (Keltjens *et al.*, 2014).

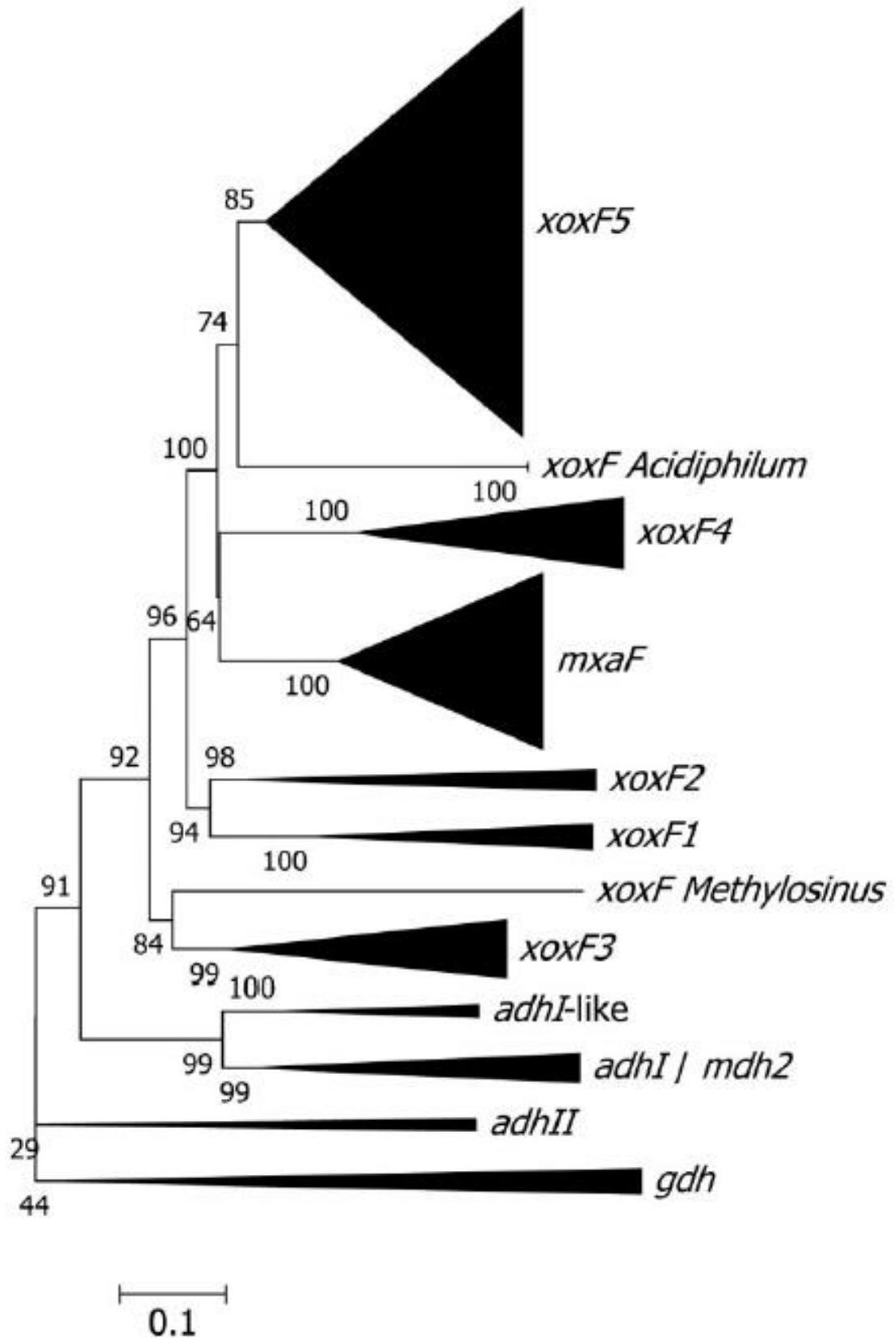


Figure 1.5 Phylogenetic tree indicating the relationship between the different *xoxF* and *mxoF* genes, as well as other PQQ-dependent dehydrogenases. Tree construction carried out using the neighbour-joining method for clustering, and maximum likelihood method for evolutionary distances. Numbers at branches represent bootstrap values over 500 replicates. Scale bar: 1 nucleotide substitution per 10 nucleotides. Taken from Taubert *et al.* (2015).

The five distinct *xoxF* clades vastly vary in their phylogenetic distribution, such that one clade is restricted to the bacterial family *Methylophilaceae* (*xoxF4*) and other clades, such as *xoxF1* and *xoxF2*, have mostly been detected within the *Beijerinckaceae* genus or the *Verrucomicrobia* phylum, respectively (Macey, 2017). On the other hand, the *xoxF3* clade is broadly distributed in members of the *Rhizobiales* (belonging to the *Alphaproteobacteria*), *Methylococcales* (belonging to the *Gammaproteobacteria*) and some *Betaproteobacteria* genera (Macey, 2017). However, *xoxF5* is the most widely distributed of the five clades, detectable within members of the *Alpha*-, *Beta*- and *Gammaproteobacteria* and the *Firmicutes* (Taubert *et al.*, 2015; Watanabe, Kojima and Fukui, 2015). Although much of the research on *xoxF* has been lab-based, less is known about its diversity and true role in the environment. Especially as *xoxF* genes have also been detected in non-methylotrophic bacteria and also play a proposed role in stress response, further muddying the exact function of this gene (Mühlencoert and Müller, 2002; Firsova, Torgonskaya and Trotsenko, 2015).

Attempts to further investigate the environmental relevance of the different *xoxF* clades involved investigating metagenomes (Ramachandran and Walsh, 2015), and creating PCR primer sets to target the different *xoxF* clades (Taubert *et al.*, 2015). These studies indicated the prevalence and diversity of the *xoxF4* and *xoxF5* clades across different marine and freshwater environments, and a large amount of unclassified *xoxF5* sequences at the genus level. *XoxF5* displayed high diversity between the investigated environments, and indicated that the *Alphaproteobacteria* were the dominant OTUs, compared to the two much more shallow sampled locations which indicated more diversity amongst *Proteobacteria* classes (Taubert *et al.*, 2015).

1.2.5. *Marine methylotrophy*

Up until the early 1970's most methylotroph isolations were derived from the terrestrial environment, however the marine environment soon became a target for isolating novel methylotrophic bacteria (Kouno *et al.*, 1973). After initial success, a stream of identification and characterisation studies of marine methylotrophs soon followed (Yamamoto *et al.*, 1980; Lidstrom, Fulton and Ann E. Wopat, 1983; Janvier *et al.*, 1985; Weaver and Lidstrom, 1985; Urakami and Komagata, 1987).

Additionally, the development of a primer set to target the functional *mxoF* gene in 1997 by McDonald and Murrell (1997), allowed for the further identification of methylotrophs and targeting of marine methylotrophs. This primer set has also been successfully used in combination with DNA stable isotope probing (DNA-SIP) experiments to identify marine methylotrophs, using ^{13}C -labelled methanol (Neufeld *et al.*, 2007). DNA-SIP has also been used in combination with metagenomics to identify active methylotrophs, such as the *Methylophaga* genus (Neufeld *et al.*, 2008). It has also been used to characterise an uncultivated *Methylophaga* that was actively incorporating the ^{13}C -labelled methanol, and allowed the almost complete construction of a genome, and was then combined with metaproteomics to reveal the pathways of the strain when utilising methanol (Grob *et al.*, 2015). Complementation studies have shown that the *mxoF* genes can be effectively used for this purpose, due to their broad functional conservation amongst methylotrophs, which was shown by restoring MDH activity in a *M. extorquens* AM1 *mxoF* mutant with cloned *mxoF* genes from a methanotroph (McDonald and Murrell, 1997). This conservation of the *mxo* gene cluster amongst methylotrophs of different subclasses likely arose through horizontal gene transfer (HGT) amongst methylotrophs of different sub-classes (Vuilleumier *et al.*, 2009). This is represented by a high degree and diversity of insertion elements in the genomes of *Methylobacterium*, along with clustered organisations of genes for C1 metabolism (Vuilleumier *et al.*, 2009).

The *Methylophaga* genus was determined as a prevalent group of obligate methanol-utilising marine methylotrophic bacteria by using the DNA-SIP technique in coastal waters (Neufeld *et al.*, 2007, 2008). This genus is readily isolated from the marine environment, so much so that has even been referred to as a 'methylotrophic weed', however its importance in the microbial community is questionable due to its low detectable abundance within the marine microbial community (Janvier, Regnault and Grimont, 2003; Sargeant *et al.*, 2016). However, in 1990, the Sargasso Sea was investigated for genetic diversity amongst bacterioplankton, which unveiled the presence of the SAR11 cluster of *Alphaproteobacteria*, which made up a significant portion of the oligotrophic bacterial community (Giovannoni *et al.*, 1990), as well as coastal waters (Morris *et al.*, 2002). This clade has since been shown to comprise up to 50% of heterotrophic bacteria in surface seawater (Morris *et al.*, 2002) and - more importantly to this project - is also capable of oxidising a variety of single carbon compounds (including methanol; Sun *et al.*, 2011). The study by Sun *et al.* (2011) used ^{14}C -labelled methanol with SAR11 strain

HTCC1062, and appeared to only incorporate 2 - 6% of the methanol as biomass, using the rest as an energy source. This high preference for incorporating methanol as an energy carbon source instead of into biomass in the open ocean, was also demonstrated for methanol uptake rates in coastal and Atlantic seawater (Dixon, Beale and Nightingale, 2011; Sargeant *et al.*, 2016). However, although Dixon, Beale and Nightingale (2011) also showed that 20 – 50% of methanol is used for cell growth in coastal and shelf seas, these combined findings do indicate that DNA-SIP may be inadequate for marine methylotroph investigations if only a small percentage of the methanol is being incorporated within the cell before DNA extraction (depending on the research aims) (Sargeant, 2013). *Rhodobacterales* are another highly abundant bacterial group capable of utilising methanol (for assimilation), and have been found to correlate with changes in assimilation rates in seawater, along with SAR11 which correlated with dissimilation rates in a seasonal study carried out by Sargeant (2013) in the English channel. However, as SAR11 genomes do not contain an *mxoF*-dependent pathway, this was more evidence for other additional major pathways involved in methanol utilisation (Giovannoni *et al.*, 2008; Kalyuzhnaya, Hristova, *et al.*, 2008).

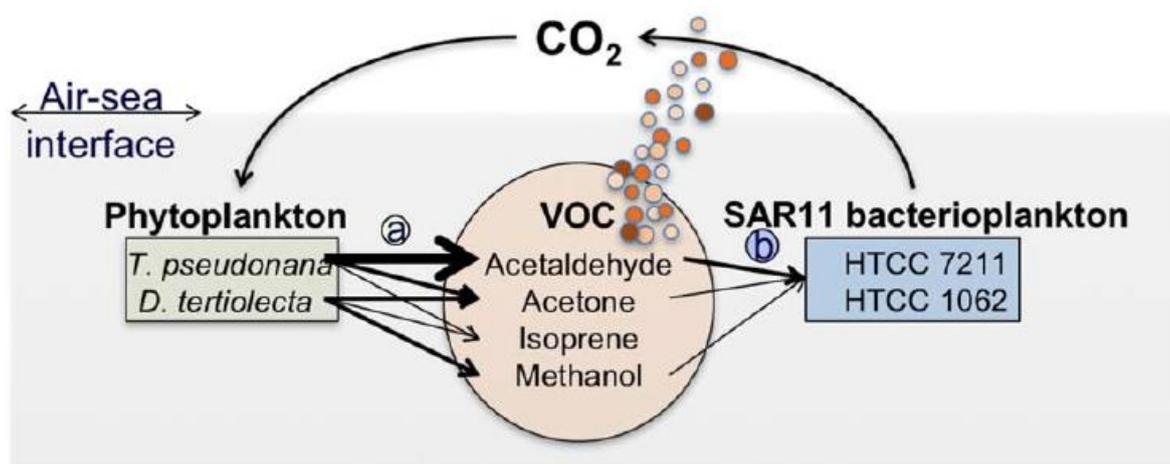


Figure 1.6 Figure of the cycling of volatile organic compounds (VOCs) between a) phytoplankton and b) bacteria. Orange circles are indicative of unutilised VOCs free to escape to atmosphere. Arrow thickness indicates relative rates of VOC production and consumption (for energy or biomass). Taken from Halsey *et al.* (2017).

In addition to the SAR11 clade, the OM43 clade is an abundant group of methylotrophic bacteria also associated with productive coastal waters and

oligotrophic gyres (Rappé, Vergin and Giovannoni, 2000), and contains a XoxF-like methanol dehydrogenase (Giovannoni *et al.*, 2008). Although unable to grow solely on methanol, growth was enhanced when methanol was added to the media used, suggesting methanol is used as an energy source in a similar fashion to SAR11 (Halsey, Carter and Giovannoni, 2012; Halsey *et al.*, 2017). Additionally this clade has been shown to be dominant during a phytoplankton bloom, which is all the more interesting when the proposed production of methanol by phytoplankton is considered (Figure 1.6) (Morris, Longnecker and Giovannoni, 2006; Mincer and Aicher, 2016). An interesting similarity between the two very successful SAR11 and OM43 clades, is their small genomes (1.3 Mbp), which is a result of genome streamlining (Giovannoni *et al.*, 2005, 2008; Giovannoni, Thrash and Temperton, 2014; Salcher *et al.*, 2019). This is where nutrient resources being utilised efficiently has led to a reduction in genome size and metabolic specialisation in select bacterioplankton lineages (Giovannoni *et al.*, 2005).

1.3. Viruses in the environment

1.3.1. *An introduction to viruses*

One major ecological challenge, will be to understand the connection of methylophony with the many biogeochemical processes and to uncover important carbon cycling components which utilise methylophic pathways (Chistoserdova, Kalyuzhnaya and Lidstrom, 2009). One such component, which has been vastly understudied regarding methylophic bacteria, is the role that viruses play. Viruses are, in most cases, entities that exist as genetic material within a protein shell (capsid) and lacking any form of propulsion are reliant on passive diffusion to encounter a host cell.

Having no inherent metabolic capabilities, virus particles must 'hijack' the internal cellular machinery of a host organism, which they usually gain access to by attaching to external transporter proteins. Typically, virus particles range in size from 20 - 200 nm, with a protein capsid shell, which encloses a nucleic acid genome. The largest discovered virus was 1.5 μm in length, although the largest viruses of bacteria (referred to as bacteriophage - or simply, phage) are smaller than this, with 160 nm capsids and 453 nm tail lengths (Legendre *et al.*, 2014; Yuan and Gao, 2017). Some viruses also have lipid membrane envelopes surrounding the capsid, which has been shown to aid in attachment and adsorption into the host (Bamford, Palva and Lounatmaa, 1976). When a virus particle makes contact with a susceptible host cell, there are various modes of infection that viruses utilise (Figure 1.7), which include lytic, lysogenic, pseudolysogenic and chronic infection cycles (Fuhrman, 1999). Lytic infection is a scenario where virus particles (virions) attach to a host cell and, assuming successful adsorption to the cell surface, its genetic material is released into the cell (Figure 1.8).

Adsorption to the cell surface and the subsequent expression of specific genes and approaches that a lytic virus can utilise varies widely, but can generally be broken down in to three main phases (early, middle and late) (Doron *et al.*, 2016; Howard-varona *et al.*, 2017). The genes that are expressed early on in an infection cycle are believed to overcome the host defences, repurpose the cellular machinery and can significantly reduce the hosts genomic DNA (Doron *et al.*, 2016). One of the many typical bacterial responses to this approach can be to overexpress degradation

enzymes to reduce viral DNA, however this can be countered by an anti-restriction response by the virus (Howard-varona *et al.*, 2017). During this initial host-resistance, the virus particle can also target the hosts RNA polymerases to preferentially select for viral promoter sequences over host sequences which will then express viral genes - priming the cell for the next phase of the cycle (Howard-varona *et al.*, 2017). The middle phase of infection has been associated more with a focus on DNA metabolism, phage DNA replication and energy metabolism (Doron *et al.*, 2016; Howard-varona *et al.*, 2017). Host energy metabolism can be augmented and facilitated by the expression of virally encoded auxiliary metabolic genes (AMGs) (Hurwitz, Brum and Sullivan, 2015; Hurwitz and Ren, 2016). Some of the most explored AMGs belong to the phage of the photosynthetic marine cyanobacteria *Prochlorococcus* and *Synechococcus*, which are increasingly photosynthetically inhibited during the course of a viral infection (Hurwitz and Ren, 2016). However, the expression of virally encoded AMGs in infected cells can support the photosynthetic capacity of the cell and thereby ensure a continued source of energy for virus replication during the final stages of the infection cycle (Lindell *et al.*, 2005). The final infection stage (i.e. late) primarily constitutes virion formation and morphogenesis, as well as additional energy metabolism genes before eventual cell lysis (Doron *et al.*, 2016). The virion particle typically comprises of a capsid shell, made up of individual structural proteins which are assembled via intracellular processes and a terminase subunit is utilised to package the capsid with a replicate of the virus genome (Sun *et al.*, 2012). Eventually the fully formed virus particles are released from within the bacterial host, generally resulting from a build-up of “lysins” within the cell wall and the consequent cleaving of covalent bonds responsible for wall stability resulting in eventual cell lysis (Figure 1.8; Clokie and Kropinski, 2009).

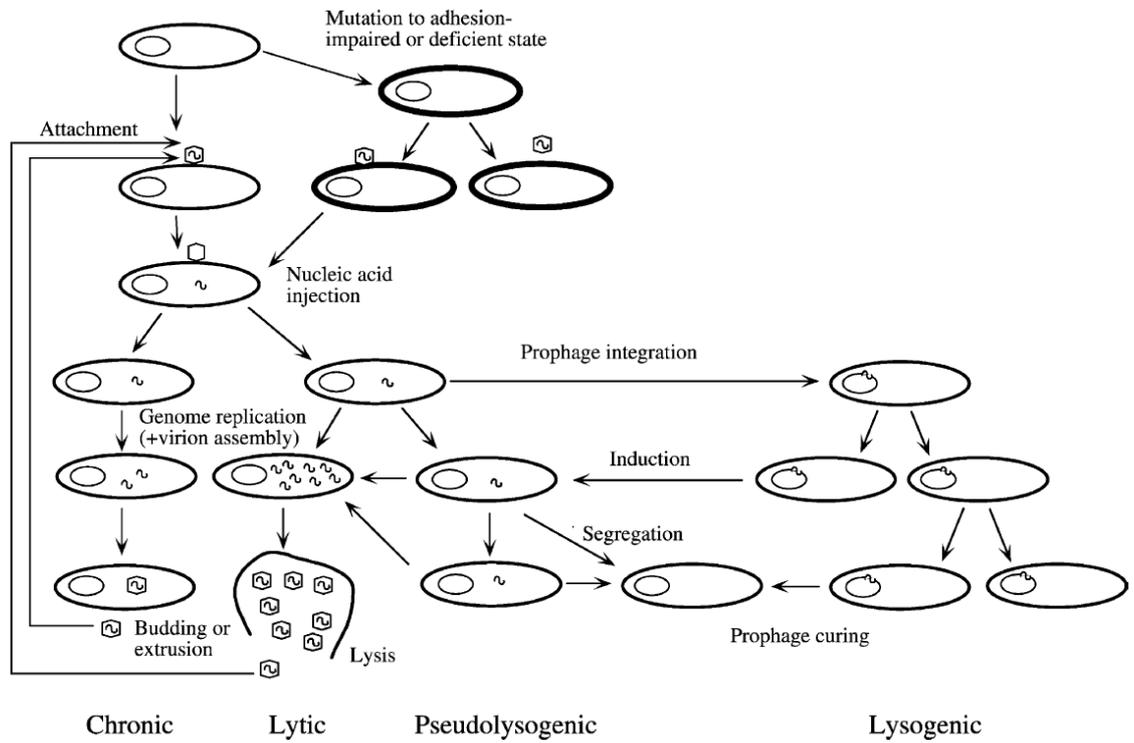


Figure 1.7 Types of modes of infection utilised by viruses, indicating the various major steps involved or that can occur. Taken from Weinbauer (2004).

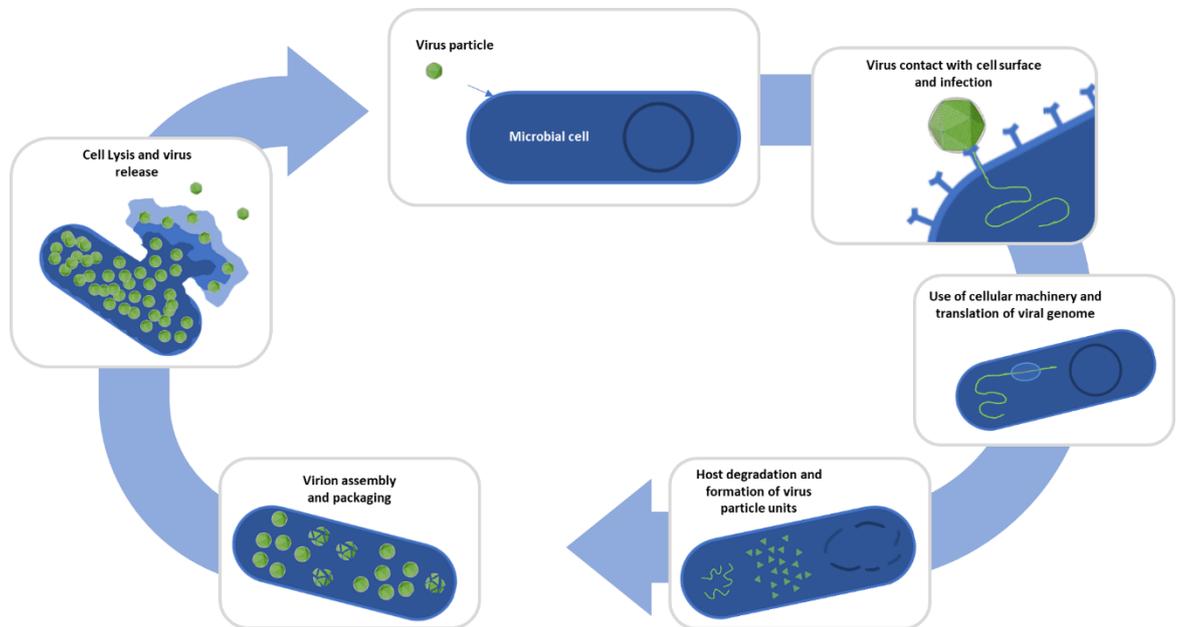


Figure 1.8 Simplistic schematic of the stages of a lytic infection upon a bacterial cell.

A lysogenic life cycle is when the genetic material of a temperate phage inserts itself into the host genome (existing as a 'prophage') and the genetic material of the phage is replicated into the daughter cells of the host. A trigger event can then occur which induces the lytic cycle, once again renewing the sequence. Pseudolysogeny is defined as a more unstable relationship between the host and virus, where the phage genome does not get fully incorporated into the host cell, and yields infected and sensitive virus progeny (Ackermann and DuBow, 1987). A chronic infection is when lysis of the host cell does not occur following infection, and instead progeny viruses are released by alternate routes, such as extrusion or budding from the host (Fuhrman, 1999).

The foundations for the importance and ecological relevance of viruses to the environment was set out by Proctor & Fuhrman (1990), with special significance within the marine environment, stating that viruses could prove to be significant for carbon and nitrogen cycling, biogeochemical cycling, genetic exchange and potentially climate change, and also believed that they should be included as major components in future models of marine systems.

1.3.2. *Phage in the marine environment*

The first marine bacteriophage was reported in 1955 by Spencer, who isolated a phage from seawater collected off the coast of Aberdeen, UK, which lysed the marine bacterium, *Vibrio phosphoreum*. It was also indicated that phage are present in offshore seawater samples but appeared to be sparsely present. However, virus particles have since been found to exist in high abundance in the ocean, ranging up to tens of millions per millilitre and often exceeding the abundance of prokaryotes by as much as 5 – 25 times (Marie *et al.*, 1999; Suttle, 2005; Mojica, Evans and Brussaard, 2014). As prokaryotes are the most abundant group of organisms in the oceans, it would make sense that most marine viruses are also bacteriophages (Sandaa, 2008). As understanding of microbial populations in the environment grew, so too has the potential significance of bacteriophage and other viruses (Proctor and Fuhrman, 1990). Virus infection occurs by diffusion, and subsequent contact with bacterial cells is random, with contact being a function of host abundance and size, thereby making the likelihood of a phage making contact with a bacterium more probable than viruses of higher, less abundant taxonomic groups (Murray and Jackson, 1992b). Other factors are involved in random contact, such as the decay

rates of virus particles, which can become inactivated over time (Fuhrman, 1999). Regardless, cell lysis - induced by a viral infection - is recognised as a major cause of mortality in the marine environment, and has been shown to rival mortality rates caused by grazing from higher trophic levels (Wommack and Colwell, 2000; Sandaa, 2008).

Although the importance of viruses has been largely established, many aspects of virus-host interactions are still poorly understood. Typically this arises from the complications involved in actually culturing viruses for study, which has made the field of virology a very difficult and, in some areas, still elusive research subject (Rohwer and Thurber, 2009). The original approach to isolating specific phage is to target their hosts, as viruses tend to infect species and cells of a single taxonomic group and in many cases are restricted to a specific host species, which consequently results in a massive amount of diversity within the viral community (Breitbart *et al.*, 2007). Regardless, even if the host is in culture, which is not guaranteed due to the observed discrepancy between bacterial counts and cultivatable bacteria known as the 'great plate count' anomaly (Staley and Konopka, 1985), there may still be a low likelihood of a full lytic cycle occurring. A lack of cell lysis could result from many reasons, including imperfection/inactivation of the virus particle itself before infection begins, a phage being less virulent (less able to lyse bacterial cultures) or even resistance mechanisms exhibited by the bacterial host at different stages of infection (Figure 1.9; Wommack and Colwell, 2000; Hyman and Abedon, 2010 and references therein). Regarding the study of viruses, the lytic infection cycle could be interpreted as preferable to other infection types, as it allows for a more robust replication of experiments/results and it can be easier to determine when a successful lytic infection has occurred (due to cell lysis). However, with culture dependent approaches, it can be much more difficult to determine why an infection has not occurred, compared to why one has, or indeed what type of infection has occurred.

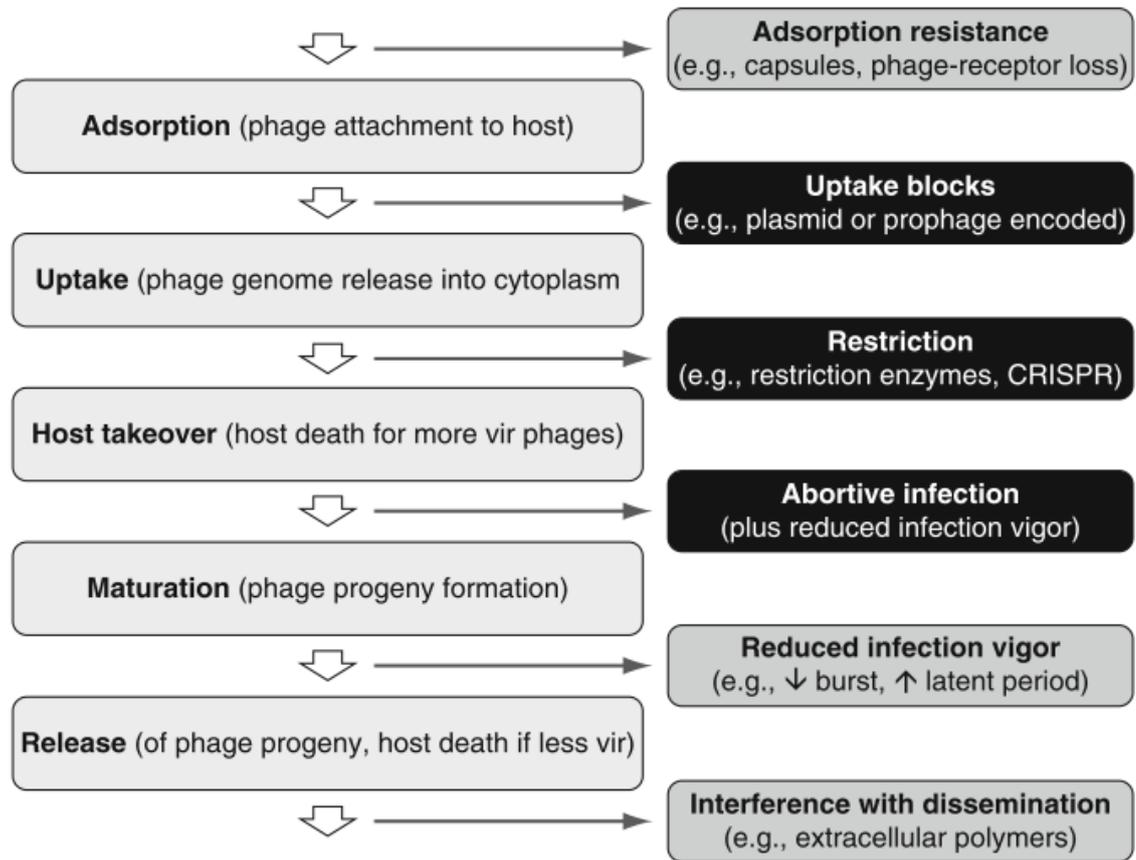


Figure 1.9 Bacterial resistance as a function of different phage infection stages. Darker shaded stages indicate greater levels of reduction of phage fitness. Abbreviation and acronyms: “vir” for “virulent”, CRISPR for clustered regularly interspaced short palindromic repeats and \wedge and \vee for increased and decreased, respectively. Taken from (Hyman and Abedon, 2010).

Alternative culture independent techniques have also been used to investigate viruses, such as the development of primer sets for analysing virus diversity and abundance, however there is no conserved gene available for viruses (e.g. *16S rRNA* gene used in microbial analysis). Therefore more targeted primer sets have been developed to investigate subsets of the virus community, such as the *phoH* gene, which has been used to investigate marine cyanophage (viruses infecting cyanobacteria), however this approach only yielded limited results for approximately 40% of marine cyanophage genomes (Goldsmith *et al.*, 2011), indicating that even genes associated with certain viruses are not conserved within those virus communities. Other culture independent approaches have attempted to link viruses and hosts, such as viral tagging, however this is a very restrictive approach and technically limiting (Deng *et al.*, 2014) and has not been successfully utilised outside the original research group. Other independent approaches include metagenomic sequencing of marine viral communities, however typically 60 – 80% of the genomes

of previously cultured marine isolates show no similarity to GenBank sequences (Paul and Sullivan, 2005). However, metagenomic analysis of viral assemblages has shown that phage genotypes differ significantly between different marine environments, but also showed that 80% of phage genotypes are conserved between the different marine environments (Angly *et al.*, 2006). Implying that viruses in one sample are enormously diverse, but if the viruses were to be pooled from different regions, then they are no more diverse than one individual sample. Additionally, utilising the metagenomic approach has allowed identification and subsequent analysis of the potentially most abundant viruses in the ocean, the pelagiphages, which infect the SAR11 clade of bacteria, previously mentioned in Section 1.2.5 (Zhao *et al.*, 2013). Another recent, more targeted approach involves the use of flow sorting, where viruses can be separated from their natural assemblage, allowing the circumvention of virus cultivation and direct investigation of the genomic information of naturally occurring viruses (Martínez, Swan and Wilson, 2014; Wilson *et al.*, 2017).

Characterising virus communities has also been attempted by investigating the genome sizes within the environment, by using pulsed field gel electrophoresis, which has shown that the main dsDNA viral assemblage in the marine environment ranges between ~15 – 630 kb, and the most dominant of these are the small (20 – 80 kb) and medium (80 - 280 kb) virus genomes (Øvreås *et al.*, 2003). As bacteria are the most dominant host group, then it would follow that the marine phage genome size range sits between 39 – 243 kb, aligning with the most abundant virus genome sizes (Sandaa, 2008). Alternative approaches for investigating viruses in the environment includes determining virus abundance, which has previously been attempted by a plaque forming unit assay, transmission electron microscopy (TEM) and/or epifluorescence microscopy (Weinbauer, 2004; Ortmann and Suttle, 2009). However, flow cytometry is arguably the best approach for doing environmental abundance counts as it can quickly yield high precision, total counts (Weinbauer, 2004; Brussaard, 2006).

The flow cytometric approach also has the added benefit of discriminating virus particles into various groupings, based on the fluorescence and side scatter properties of the stained virus-like particles (Brussaard, 2009). Three main groupings (VLP1, VLP2 and VLP3) have been identified in environmental water samples from various sampling locations in marine and freshwater systems, which

appear to be coupled with the temporal abundance dynamics of different primary and secondary producers (Mojica, Evans and Brussaard, 2014; Zhong et al., 2014; Evans et al., 2017). The VLP1 group has been consistently linked to heterotrophic bacteria based on correlations in shifts in abundance, whereas the VLP3 grouping indicates a closer association with phytoplankton and chlorophyll-a concentrations (Brussaard et al., 2008; Payet and Suttle, 2008a; Evans et al., 2011, 2017; Mojica and Brussaard, 2014; Zhong et al., 2014). This is presumed to be a direct consequence of virus–host relationships, and subsequently indicates that the more abundant VLP1 group consists largely of bacteriophage, whilst the typically less abundant VLP3 group is more associated with algal viruses. However, VLP2 is also considered to represent bacteriophage populations, but shares more of an association with the dynamics of the VLP3 group than VLP1 and is therefore believed to comprise more of cyanobacterial viruses (Evans et al., 2011; Zhong et al., 2014). The abundance of virus particles has been used in combination with bacterial abundance to determine virus-to-bacteria ratios (VBR), which can give insights to the relationship between viral and bacterial populations in different environments. Generally indicating that VBR values are often higher in nutrient-rich, productive environments and are driven by infection rates and cell burst sizes, and decrease with microbial cell density (Wommack and Colwell, 2000; Wigington *et al.*, 2016).

1.3.3. *Marine biogeochemical influences*

Phage are estimated to be responsible for approximately 20 - 50% of the total bacterial mortality in surface waters (Suttle, 1994; Fuhrman and Noble, 1995; Steward, Smith and Azam, 1996). However, some groups of organisms display very little bacterial mortality due to viral predation, and the effect has been hypothesised to be seasonal, local or sporadic (Fuhrman, 1999). Whether this is the case or not, lytic infection releases cellular debris. This debris mainly consists of dissolved molecules (such as monomers, oligomers and polymers), colloids and cell fragments, and when combined together are typically defined as dissolved organic matter (DOM; Shibata *et al.*, 1997). Most, if not all, of the DOM is immediately or eventually available to lower trophic levels (Middelboe, Jørgensen and Kroer, 1996; Gobler *et al.*, 1997). One mixed population study demonstrated that the lysate

derived from a phage-host system was used very efficiently by an unrelated bacterial group within the population, which was previously incapable of significant growth (Middelboe *et al.*, 2003), highlighting that one bacterial population's loss can be another's gain. Typically, it is bacteria that utilise lysis products first, which can produce an inefficient semi-closed trophic loop if the products were from a bacterium in the first place. The effects of a viral presence in a microbial community can be shown in a model by Fuhrman (1992), which states a comparison between two bacterial populations. The population with a viral presence demonstrated a 50% viral mortality rate, which yielded a 27% boost to productivity and to the respiratory rate (Gobler *et al.*, 1997). This would have greater implications in denser populations, and on the microbial community. This view can be reinforced by the same study, which also found a ten-fold increase in the relevant bacterial abundances (Gobler *et al.*, 1997). Further studies came to the conclusion that heterotrophic uptake of lytic products increased by 72%, which was seen to be a response to facilitate the assimilation of these products, as more energy was necessary, instead of potentially being converted into biomass (Higgins, 2005).

A review by Suttle (2007) stated that viral lysis accounts for approximately 20 - 40% of the removal of prokaryotic biomass standing stock within the surface waters of the ocean. This organic matter, released via viral lysis, is capable of being converted into dissolved organic matter (DOM). Viral lysis of bacteria has been found to release larger quantities of DOM than it does particulate organic matter (POM; Poorvin *et al.*, 2004). This process removes this organic matter from the trophic ladder, termed the "viral shunt", and essentially disrupts the microbial loop (Figure 1.10). This loop is a model which describes the flow of carbon through the trophic levels (Azam *et al.* 1983; Fuhrman 1992), where approximately 20 – 25% of the photosynthetically fixed carbon is 'shunted' by viral pathways into pools of dissolved organic matter (Fuhrman, 1999; Wilhelm and Suttle, 1999). The viral shunt is a very large contributing factor to carbon fluxes, supplying dissolved organic matter to the already substantial organic matter budget, and has been seen as a stimulant of growth for heterotrophic bacteria (Figure 1.10) (Fuhrman, 1999). Marine viruses therefore increase the primary productivity occurring in the water column by making more inorganic nutrients available to the autotrophic carbon fixers, thereby boosting secondary productivity by increasing the recycling of organic material available to higher trophic levels (Weitz *et al.*, 2015).

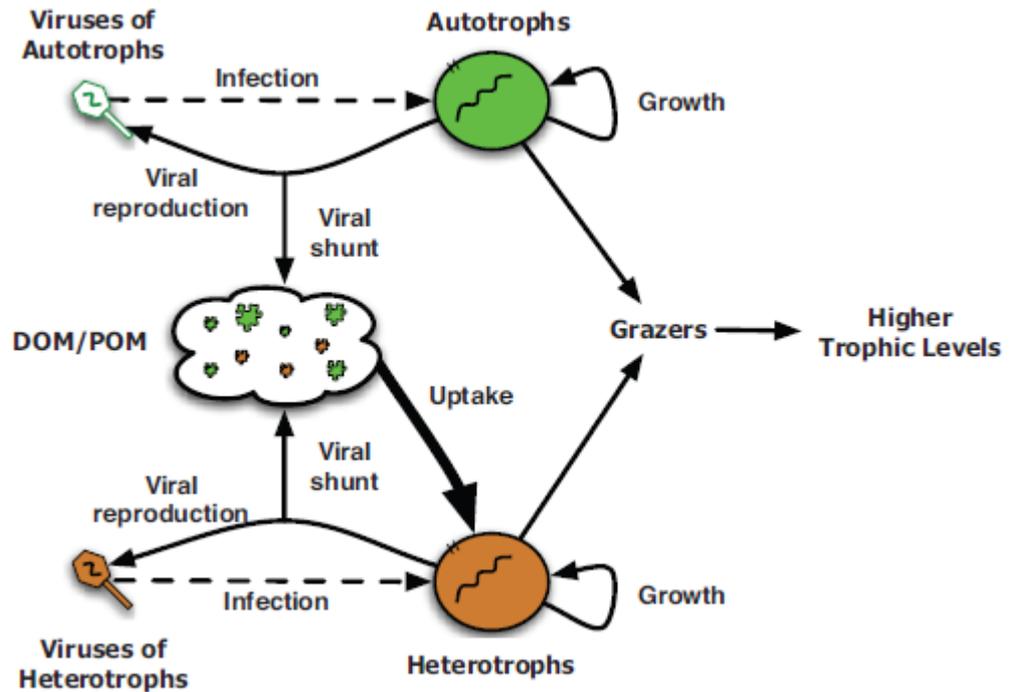


Figure 1.10 Cellular material is released into the environment by microbes upon lysis. The cellular material released could be utilised by microbes for subsequent metabolic processes. The thicker arrow is used to denote that the majority of uptake is by heterotrophs. Taken from Weitz & Wilhelm (2012).

Viruses have even been linked to playing a role in shaping the global climate, made possible by inducing the release of dimethyl sulphide (DMS) by infection of *Micromonas pusilla*, *Emiliania huxleyi*, and *Phaeocystis pouchetii*, which is a precursor to dimethylsulphoniopropionate (DMSP) and which is involved in cloud nucleation (Charlson *et al.*, 1987; Brussaard *et al.*, 1996; Hill *et al.*, 1998; Malin *et al.*, 1998). Although these compounds can be released to the atmosphere, they are just as likely to be transformed and degraded by microbial action (Kiene and Bates, 1990). As viruses must diffuse from host to host with contact entirely random, rarer hosts are less susceptible to triggering infection in their population than more common ones (Fuhrman and Suttle, 1993). This phenomenon is commonly referred to as 'kill the winner' (Thingstad and Lignell, 1997; Thingstad, 2000), where phage therefore maintain diversity within the microbial community. This is based on the premise that phage counteract the dominance of a specific bacterial population within the community. Therefore, rarer species are less likely to be infected; however, there may be advantages to being infected continuously. If bacterial growth is limited in oligotrophic environments, some hosts rely on unsuccessful viral infection (halted by genetic incompatibility or a restriction enzyme) to provide them

with genetic material, rich in carbon, nitrogen and phosphorus (Proctor and Fuhrman, 1990; Fuhrman, 1999).

1.3.4. Phage of methylotrophic bacteria

In the early 1970's and 80's research was carried out into the cultivation of bacteria and yeasts on commercial scales for the production of single cell protein (SCP), to be used as animal feed (Senior and Windass, 1980; Windass *et al.*, 1980). This operation utilised bioreactor systems with high capacities, with one of the most well researched organisms being *Methylophilus methylotrophus* (Windass *et al.*, 1980). This organism was chosen for its high optimum growth temperature (~40 °C), rapid growth on methanol and its high yield (Anthony, 1982). The product of this process was high-quality (~ 70% cell protein) and capable of outputting a projected 70,000 T yr⁻¹, due to a high conversion rate of methanol to SCP (Kelly, Ardley and Wood, 2015). With an estimated £100 million invested into this research in the 1970s and 80s, it was proven to be technically feasible, but was however discontinued due to its perceived uneconomical viability at that time. However, this research also included the first characterisation studies on phage that could infect methylotrophs. The first ever report of a phage to utilise a methanol-assimilating methylotroph as its host was reported by Oki *et al.* (1972). They isolated three separate phage morphologies and reported the methods employed for isolation and the analysis carried out, which included electron micrographs. Ichikawa *et al.* (1977) derived four methylotrophs from the terrestrial environment, from which 11 phage were isolated. Tyutikov *et al.* (1980) reported on the first ever phage of methane-oxidising bacteria from plants and the environment, which took the form of two different particle morphologies with short or long noncontractile tails (Figure 1.11). These are the only reports of phage which infect obligate methylotrophic bacteria.

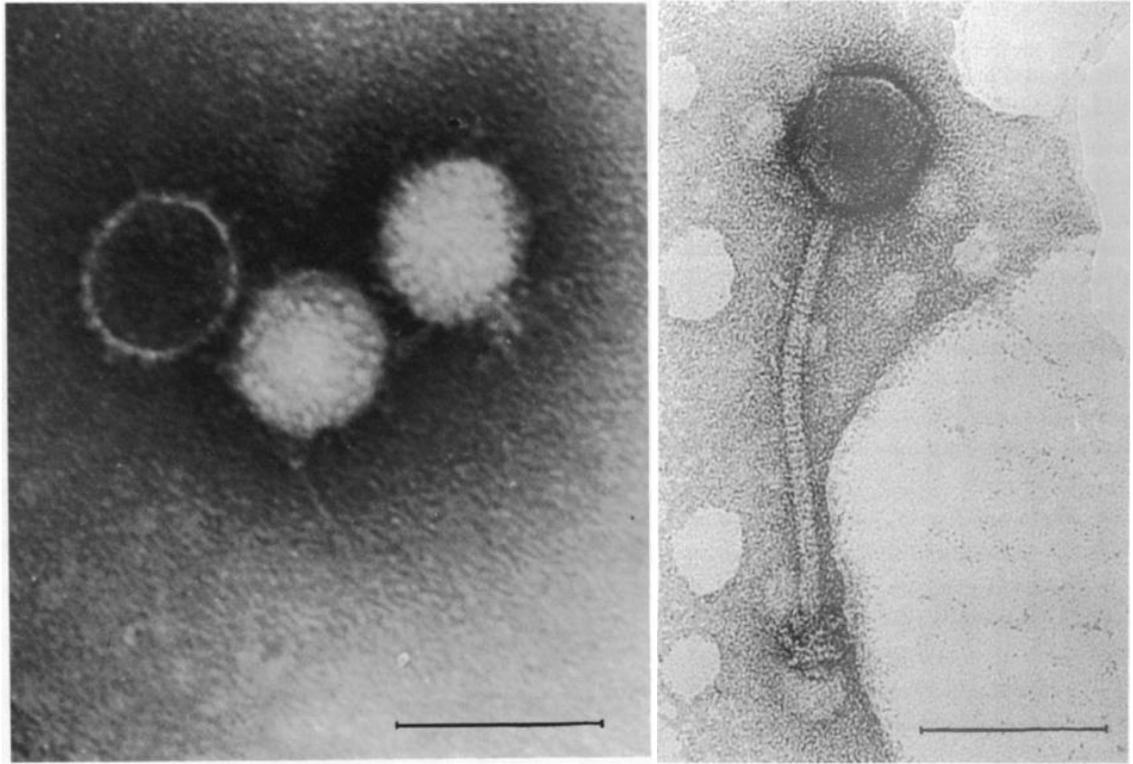


Figure 1.11 Figures taken from Tyutikov *et al.* (1980), indicating the two different virus morphologies isolated from methanotrophic bacteria (Bar = 0.1 μm).

Apart from these few reports, there is no research on the impact that methylotrophic phage have on marine methylotrophs, and on the greater marine environment. Although viruses of abundant bacterioplankton with methylotrophic capabilities have been identified, such as SAR11 (Zhao *et al.* 2013), research on their wider impact on biogeochemistry and ecology has only been proposed. This leaves us with no knowledge of the role of phage in methanol cycling in the oceans, but because of their significant contributions to biogeochemical cycling and impacts on the ecology of marine systems, it is clear that steps must be taken to fully understand their role.

1.4. Project aims and objectives

The aims and objectives of this project were:

1. To isolate and characterise a lytic virus which infects an obligate methylophilic marine bacterium and subsequently establish a robust virus-host system. This would allow a platform for further downstream experiments exploring the unknown infection dynamics and influences upon a methylophilic host, providing valuable insight to the impacts on microbial methanol utilisation.

This aims to test the hypothesis that a methanol-utilising bacterial host can be infected by a lytic virus and will have a decreased capacity to utilise methanol during infection. This is due to a redirection of metabolic processes towards virion production, thereby decreasing the methanol uptake rates of infected hosts. The first experimental chapter (Chapter 3) details the approach utilised to culture naturally occurring obligate methylophilic bacteria from the Western Channel Observatory (WCO). This is followed by a description of the various established methodologies to isolate and characterise a virus-like-particle from the WCO associated with an obligate methylophilic bacterium. The presence of the host and the infecting virus at the WCO is important to the relevance of any findings to come from a virus/host system derived from this project.

2. To further characterise seasonal changes associated with microbial methanol uptake in combination with an assessment of virus abundance dynamics throughout the water column of the WCO.

This is a more holistic aim than the first and hypothesises that the microbial methanol uptake rates will decrease due to increased virus infections during periods of increased virus-like-particle abundance. The second experimental chapter (Chapter 4) approaches this with a one-year timeseries analysis of the methanol uptake rates, virus abundance and microbial community derived from seawater samples in the WCO, with bi-weekly sampling of Station L4 and seasonal sampling of Station E1. This methanol uptake, virus abundance and microbial community data is put into context with additional standard biological and physicochemical

measurements also taken as part of the ongoing timeseries sampling at the WCO, thereby yielding additional insight to the study.

3. To define for the first time the distribution and diversity of the *xoxF5* gene clade in the marine environment and take a spatiotemporal approach by incorporating available cruise samples with the one-year WCO timeseries sampling.

We hypothesise that *xoxF5* clade sequences will be detectable and widespread throughout all the different sampled regions of the marine environment. The second experimental chapter (Chapter 4) addresses the temporal changes in *xoxF5* gene diversity by utilising the DNA samples taken as part of the WCO timeseries of this study. The third experimental chapter (Chapter 5) realises the analysis of the spatial diversity by utilising available cruise samples from the Atlantic Meridional Transect (JC039) to gain insight to the various different provinces of the Atlantic, and also samples taken from a cruise in the Barents Sea shelf region of the Arctic Ocean (JR16006, Barents Sea).

Chapter 2

Materials and methods

2. Materials and methods

2.1. Bacterial strains

Table 2-1 List of organisms used in this study

Strains	Description/genotype	Reference/source
<i>Methylophaga marina</i>	Wild-type strain	Janvier <i>et al.</i> , 1985
<i>Methylorubrum extorquens</i> AM1	Wild-type strain	Murrell Lab Collection
<i>Methylocella silvestris</i> BL2	Wild-type strain	Murrell Lab Collection
<i>Methylophaga</i> AH1 L4	Wild-type strain	Murrell Lab Collection
<i>Marinibacterium anthonyi</i> La6	Wild-type strain	Murrell Lab Collection
<i>Methylobacillus</i> MM2	Wild-type strain	Murrell Lab Collection

Table 2-2 All media components used when making the individual mediums.

Medium		Media Components (Per 1000 mL)
Marine Ammonium Mineral Salts (MAMS) ¹	Before Autoclaving	20 g NaCl
		10 ml (NH ₄) ₂ . SO ₄ (10 g / 100 mL stock)
		10 ml CaCl ₂ . 2H ₂ O (2 g / 100 mL stock)
		10 ml MgSO ₄ . 7H ₂ O (10 g / 100 mL stock)
		10 ml Na ₂ MoO ₄ . 2H ₂ O (0.2 g / 100 mL stock)
		10 ml FeSO ₄ . 7H ₂ O (20 mg / 100 mL stock) ²
		1 ml Na ₂ WO ₄ . 2H ₂ O (3 g / L stock)
	After Autoclaving ³	1 ml SL10 trace metals (See Table 2-3)
		50 µl Na ₃ VO ₄ + Na ₂ SeO ₃ (0.1 µg / mL stock)
		10 mL / L Phosphate Buffer (See Table 2-3)
Marine Basal Media (MBM) ¹	Before Autoclaving	20 g Sea Salts
		250 mL Basal Media (See Table 2-3)
		1 M HCl (Adjust pH to 7.10)
	After Autoclaving	50 mL FeEDTA (50 mg / 100 mL stock)
		1 mL Vitamin Solution (See Table 2-3)
		50 µL LaCl ₃ (2.45 g / 100 mL stock)

¹ 1.5 g agarose is added to the medium solutions before autoclaving, in a situation where agar plates are being made.

² This was made fresh every time the media was made.

³ Vitamin solutions were added to both mediums once cooled sufficiently (~50 °C).

Table 2-3 All media reagents used when making the individual mediums.

Media Reagents	Reagent Components
Phosphate Buffer (0.4 M, pH 7.0)	Na ₂ HPO ₄ .12H ₂ O (28.6 g / 200 mL)
	KH ₂ PO ₄ (10.9g / 200 mL)
SL-10 Trace Metal Solution (Widdel, Kohring and Mayer, 1983)	10 mL HCl (25% or 7.7 M)
	1.5 g FeCl ₂
	70 mg ZnCl ₂
	100 mg MnCl ₂
	6 mg H ₃ BO ₃
	190 mg CoCl ₂ .6H ₂ O
	2 mg CuCl ₂ .2H ₂ O
	24 mg NiCl ₂ .6H ₂ O
	36 mg Na ₂ MoO ₄ .2H ₂ O
	Dilute to 1 L in Milli-Q
Vitamin Solution (pH 4.0) (Kanagawa, Dazai and Fukuoka, 1982)	400 mL Milli-Q
	10 mg Thiamine hydrochloride
	20 mg Nicotinic acid
	20 mg Pyridoxine Hydrochloride
	10 mg p-Aminobenzoic acid
	20 mg Riboflavin
	20 mg Calcium pantothenate
	1 mg Biotin
	2 mg Cyanocobalamin
	5 mg Lipoic acid
	5 mg Folic acid
Diluted to 1 L in Milli-Q	
Basal Media	150ml 1M Tris HCl solution pH 7.5 (add 18.171 g Tris Base in 150 ml, use HCl to bring to pH)
	87 mg K ₂ HPO ₄
	1.5 g NH ₄ Cl
	Diluted to 525 ml in Milli-Q
TE 10:1 Buffer (pH 8.0)	10 mM Tris
	1 mM EDTA
10 × TBE Buffer (pH 8.0)	108 g Tris
	55 g Boric Acid
	40 L 0.5 M EDTA

2.2. Cultivation and maintenance of strains

Marine Ammonium Mineral Salts (MAMS) was used as the growth medium for *Methylophaga marina*, *Methylophaga AH1* and other bacteria isolated from Station L4 (50° 15.00' N, 4° 13.02' W) and Station E1 (50° 02.00' N, 4° 22.00' W), which are located in the Western English Channel (WEC), and was prepared according to Goodwin et al. (2001) as shown in Table 2-2. Agar plates were incubated in gas-tight chambers/containers at 25°C. Headspace methanol was made available to agar growing organisms by pipetting 500 µL of 1 M methanol on to cotton wool, which was then kept in an unsealed 35 mm petri dish, to allow passive diffusion of methanol into a gaseous phase, this was replenished every few days. On occasion the volume was increased, and the largest volumetric container used was a 1 L Quickfit conical flask (Thermo Fisher Scientific, MA, USA), containing 200 mL of media, which was sealed with a SubaSeal (Sigma-Aldrich, USA) stopper and incubated in an orbital shaker (Model SI500, Stuart-Equipment, UK) at the standard 25°C, whilst shaking at 150 RPM.

2.2.1. *Methylophaga spp.*

Methylophaga marina and *Methylophaga AH1* were routinely cultured and maintained in 120 mL glass serum vials containing 30 mL MAMS medium and 10 mM methanol and sealed with grey butyl rubber seals. A new culture would be seeded with a single colony from an agar plate or 3% inoculum from a previously inoculated culture in exponential growth phase (optical density of 0.2). Optical density was determined at 600 nm using a BioPhotometer (Eppendorf, Germany). Cultures were incubated at 25°C, whilst shaking at 150 RPM until growth was observed. Growth could be semi-quantitatively gauged by measuring the optical density of 500 µL of a culture at 600 nm (OD₆₀₀) using a BioPhotometer (Eppendorf, Germany) spectrophotometer.

2.2.2. *Marinibacterium anthonyi La6*

Marinibacterium anthonyi La6 was grown using 30 mL MBM medium and 10 mM methanol and sealed with grey butyl rubber seals. Due to the slower growth cycle of *Marinibacterium anthonyi* La6 when using MBM minimal medium, Marine Broth 2216 (MB) was also utilised (20% w/v; Fisher-Scientific UK Ltd, UK), and amended with sodium chloride (15.56 g L⁻¹).

2.3. Timeseries sampling and process

The Western Channel Observatory (WCO) consists of station E1 (50° 02.00' N, 04° 22.00' W; depth 75 m) and L4 (50° 15.00' N, 04° 13.02' W; depth 50 m) which are located off of the coast of Plymouth, UK in the Western English Channel (Figure 2.1). The WCO is a well-established time series representing both open shelf (E1) and coastal waters (L4) within range of Plymouth, allowing frequent *in situ* sampling, throughout the year.



Figure 2.1 Map of the coast of Plymouth highlighting the sampling stations of the Western Channel Observatory (WCO), with stations L4 and E1 indicated with a red-filled circle. Image from the WCO website: www.westernchannelobservatory.org.uk/.

Seawater sampling was carried out, weather dependent, in the WCO at stations L4 and E1 on a weekly basis by the crew of the research vessel *Plymouth Quest*. Samples were collected using Niskin bottles, mounted on a conductivity-temperature-density (CTD) profiler, which allows for varying depths to be sampled (typically 3 m, 10 m, 25 m and 50 m at L4, and a maximum depth of 60 m at E1). Water was collected from the surface of a Plymouth Sound harbour at high tide (50° 21'48.3 N, 04° 08'48.6 W) and sediment was also collected from station L4, by using a 0.1 m² box corer, at a depth of 53 m.

Seawater collected from the sampling station used for the timeseries analysis was utilised to determine environmental variables of interest in this project (Figure 2.2).

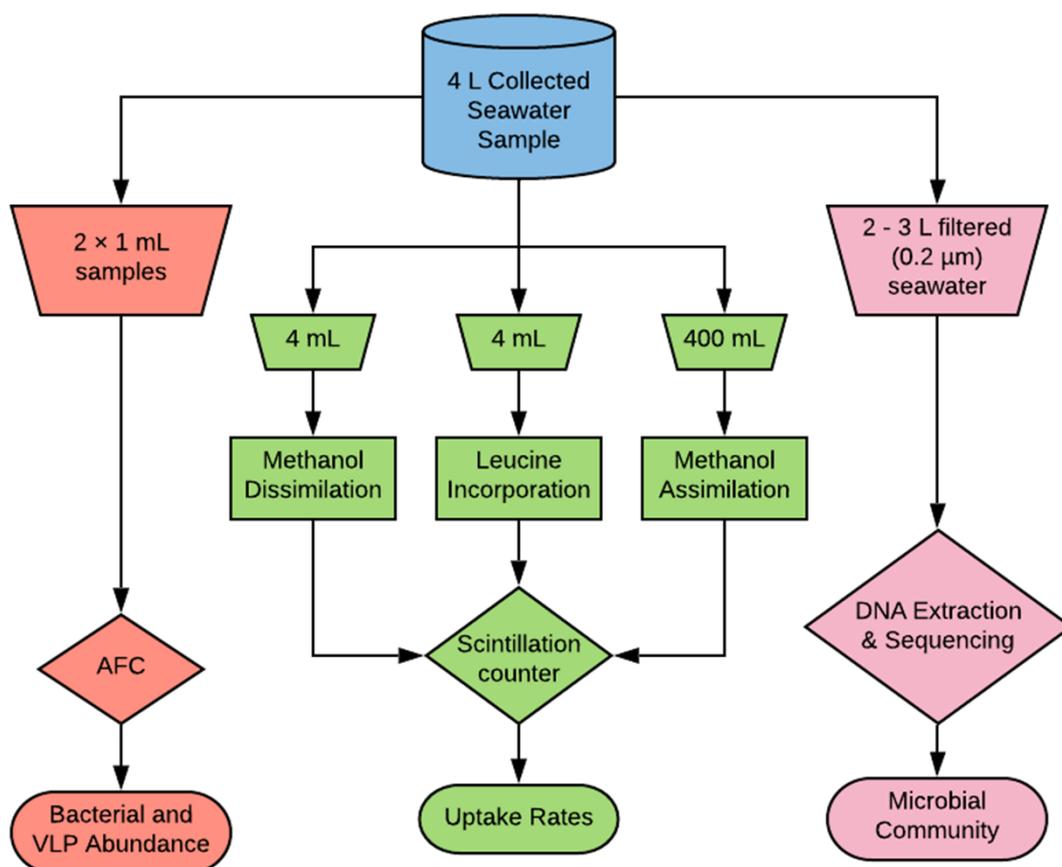


Figure 2.2 Flow diagram to represent the processing of collected seawater and the different volumes utilised for determining environmental variables of interest in the timeseries analysis. Bacterial and virus-like-particle abundance (red), microbial uptake rates of methanol and leucine incorporation (green), and microbial community (pink).

2.3.1. Supplemental timeseries and cruise data

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

- Supporting data for the English Channel was provided by the Western Channel Observatory, including: seawater temperature, salinity and fluorescence values, chlorophyll *a* concentrations, flow cytometry data (including *Synechococcus*, picoeukaryotes, nanoeukaryotes, coccolithophores, phaeocystis, cryptophytes and high/low nucleic acid bacteria) and nutrient concentrations (including nitrate, nitrite, silicate, phosphate and ammonia).

- Supporting data for the Atlantic and Arctic Ocean cruises were provided by the British Oceanographic Data Centre as part of the Atlantic Meridional Transect programme and the Changing Arctic Ocean Cruise (JC039 and JR16006, respectively), which includes: sea surface temperature; fluorescence derived chlorophyll *a* values; salinity, flow cytometry (total bacterial abundance) and nutrients (including ammonia, phosphate, silicate, nitrate and nitrite).
- Sequencing sample preparation was carried out by the author at the University of Exeter, but actual sequencing was carried out by Dr Karen Moore (Exeter Sequencing Service, University of Exeter).

Chlorophyll-a

Chlorophyll-a water samples were taken discretely along cruise transects, with water collected into 5 L Nalgene carboys using a CTD rosette system. Triplicate samples of 200 mL were analysed onboard using a Trilogy Fluorometer (Turner Designs, CA, USA), which was pre-calibrated using spinach chlorophyll-a standards (Sigma Aldrich, Germany).

Nutrients

Nutrient cruise samples were collected in 50 mL polyethylene vials using a CTD rosette system. Triplicate measurements were conducted using a Lachat *QuickChem 8500* flow injection autoanalyser (Hach Lange, Germany) using the manufacturers recommended methods. Individual stock standard solutions of nitrate, phosphate and silicate were prepared in deionised water immediately prior to cruises from oven dried salts (60°C). The stock solutions were used daily to prepare mixed working standard solutions each day using deionised water and calibration solutions prepared by the instruments autodiluter facility using OSIL Low Nutrient Sea Water for dilution (Ocean Scientific International Ltd., UK). Nutrient samples from the WCO were analysed at the Plymouth Marine Laboratory using a SEAL analytical AutoAnalyzer II, segmented flow colorimetric auto analyser, in accordance with the protocols outlined in Woodward and Rees (2001).

Table 2-4 Acknowledgements of supplemental data used in data chapters 4 and 5.

Sampling programme	Supplemental data	Collected/processed by:
Western Channel Observatory (WCO)	Temperature, salinity, depth and fluorescence (CTD Profiler)	Crew of the RV <i>Plymouth Quest</i> and Dr James Fishwick
	Chlorophyll- <i>a</i>	Denise Cummings
	Nutrients	Dr Malcolm Woodward, Lisa Al-Moosawi and Carolyn Harris
	Plankton and bacteria abundance	Dr Glen Tarran and Rebecca May
Research Cruises (JC039 and JR16006)	Bacteria abundance	Elaine Mitchell (JR16006)
	Temperature, salinity and depth (CTD Profiler)	Crew of JR16006 and BODC
	Chlorophyll- <i>a</i>	Andrew Orkney and Prof Heather Bouman (JR16006)
	Nutrients	Tim Brand and Dr Sian Henley (JR16006)
	Sterivex filters and methanol uptake rates	Dr Stephanie Sargeant and Dr Joanna Dixon (JC039)

2.4. Isolation of methylotrophic bacteria

To isolate new methanol-utilising methylotrophic bacteria, one litre of seawater was collected in a sterile carboy from the surface of station L4 in the Western Channel Observatory (WCO). 50 mL seawater enrichment cultures were set up with the addition of 1% (v/v) MAMS and a separate enrichment with 1% (v/v) MBM media and an array of methanol (MeOH) concentrations and lanthanum concentrations (by adding lanthanum chloride, LaCl₃), as described below:

5 mM MeOH	10 mM MeOH	50 mM MeOH
5 mM MeOH/5 µM La ³⁺	10 mM MeOH/5 µM La ³⁺	50 mM MeOH/5 µM La ³⁺
5 mM MeOH/10 µM La ³⁺	10 mM MeOH/10 µM La ³⁺	50 mM MeOH/10 µM La ³⁺

Lanthanum was used in this enrichment process, due to its recently highlighted

importance in the growth of methylotrophs utilising *xoxF* (Keltjens *et al.*, 2014; Howat *et al.*, 2019). Non-inoculated blank controls were set up for each methanol/lanthanum concentration used, and all enrichments and controls were incubated at 25°C (50 rpm) for 10 days, serial dilutions (10^{-1} , 10^{-2} and 10^{-3}) of these enrichments were then spread onto relevant agar plates (MAMS/MBM). These agar plates were incubated at 25°C with headspace methanol (Section 2.2) in a sealed chamber for 10 days. Colonies were purified by re-streaking onto fresh agar plates, and then growth on methanol was confirmed by inoculating into 30 mL liquid MAMS/MBM media with the relevant methanol and lanthanum concentrations of the enrichment the strains were purified from. This accounted for any growth that may have been occurring on trace organic compounds in the agar, or on the agar itself. If the purified cultures had an optical density of $OD_{600} \geq 0.15$ after 24 hours, these strains were investigated further.

2.5. Analytical Flow Cytometry (AFC)

2.5.1. *Sample preparation and storage*

Samples being prepared for flow cytometry could be analysed immediately or could be stored for extended periods (approx. 6 months) of time before being analysed. Typically, 1 mL samples were fixed by adding 10 μ L of (50%) glutaraldehyde (TEM-Grade, Sigma-Aldrich, UK) (0.5% final concentration), vortex-mixing and leaving for 15-30 minutes (4°C). The samples would then be flash frozen in liquid nitrogen, before being stored in a freezer (-80°C). A Becton Dickinson FACScan (Becton-Dickinson BioSciences, UK) was used for all flow cytometric analyses. Data acquisition was performed using the CellQuest Software (Version 3.3; Becton-Dickinson, UK).

2.5.2. *Bacterial numbers analysis by flow cytometry*

When analysing samples for bacterial numbers, a dilution series would be created resulting in a final sample volume of 0.5 mL per AFC vials (Becton-Dickinson BioSciences, UK) and stained with 50 μ L of SYBR-Green I (1 μ L mL⁻¹ diluted in K-citrate; 2.3 g per 25 mL of milli-Q water, 0.1 μ m filtered each time) (Invitrogen, Thermo-Fisher, USA), before being vortex-mixed and left for 1 h in the dark (room

temperature). Samples were analysed for 1 minute at a low flow rate (20 – 50 $\mu\text{L min}^{-1}$) under set machine parameters (discriminated using green fluorescence vs side scatter), if the event rate was too high, then a more diluted sample would be used until the optimum event rate (200 – 600 events s^{-1}) was achieved. Flow rate was determined by measuring the weight of an AFC tube containing 500 μL of milli-Q, removing the outer sheath from the AFC sample injection port and allowing the milli-Q to be taken up for 1 minute. The post-weight of the vial was also measured, and the pre-weight was subtracted from the post-weight and dividing the value by the time and multiplying this value by 1000.

2.5.3. *Viral-like particle analysis by flow cytometry*

A slightly extended protocol was used for looking at viral-like particles and follows the procedure laid out by Brussaard (2006). For determining viral-like particle abundance, samples were removed from the freezer and thawed relatively quickly in a tub containing tap water. A dilution series of the samples of interest would then be set up (0.5 – 1 mL per tube) using TE 10:1 Buffer (pH 8.2; autoclaved when made and 0.1 μm -filtered when used). A working stock of SYBR Green I (5 μL of SYBR Green per 995 μL of 0.1 μm -filtered milli-Q water) was then added to each of the diluted sample vials (5 μL per 500 μL sample for seawater samples), it was important to work in a dimmed room due to the light sensitivity of the stain. Dilutions varied between samples (100 \times for environmental samples and >1000 \times for bacterial cultures), to maintain an optimal event frequency of 200 – 600 events s^{-1} . All samples, including a stained TE-buffer control, were capped and incubated at 80 $^{\circ}\text{C}$ in the dark for 10 minutes, with one TE 10:1 Buffer control not incubated but kept in the dark for the same period. The samples would then all be removed and allowed to cool in the dark for approximately 5 minutes before analysing. The ideal running conditions, which yield the most accurate data, involved using sterile milli-Q water as the sheath fluid, running the sample at an event rate between 200 – 600 events s^{-1} and at a flow rate between 20 – 50 $\mu\text{L min}^{-1}$ for one minute. Virus-like particles were discriminated using green fluorescence against side scatter. Contamination between samples was mitigated by wiping the uptake needle each time with a moist tissue and changing the tissue regularly. When carrying out virus work, the machine was kept as clean as possible, by flushing the system with FACSClean, FACSRinse (Becton-Dickinson BioSciences, UK) and milli-Q water, before and after each use. The discrimination of different virus subgroups (VLP 1, VLP 2 and VLP 3) was based

on groupings observed in scatter plots when comparing side scatter against green fluorescence (Figure 2.3). These plots were set up using the CellQuest software (Becton Dickinson, USA).

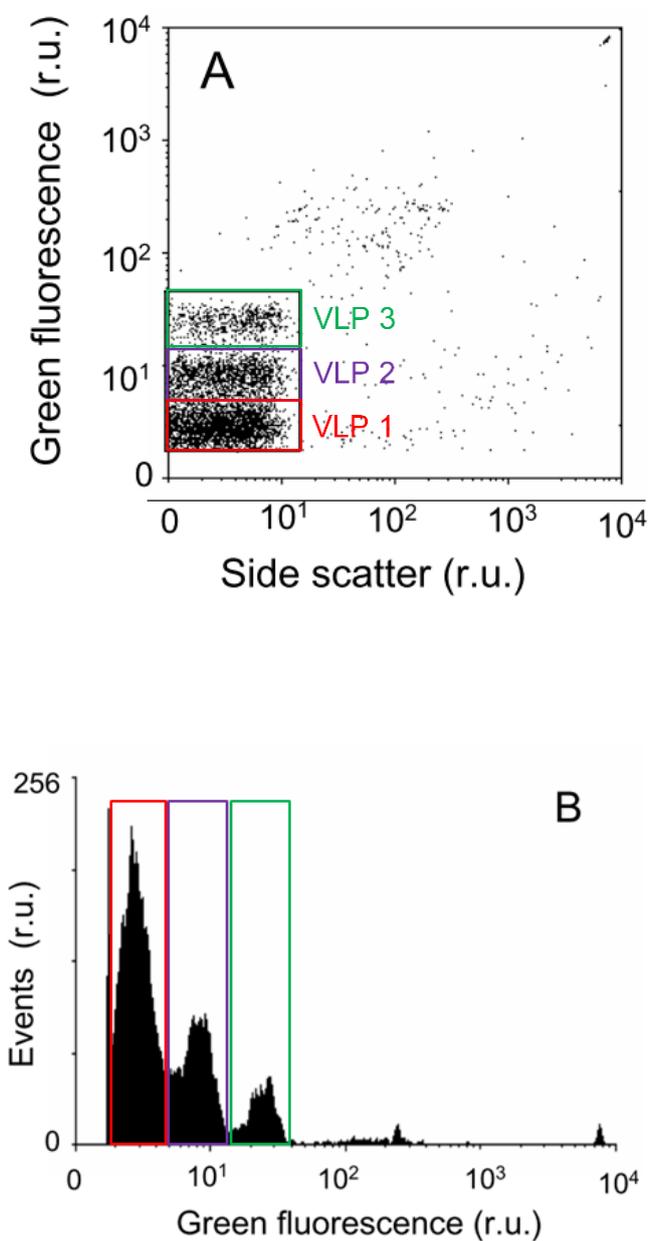


Figure 2.3 Flow cytometric distribution of virus subgroups (VLP1, 2 and 3) of a natural seawater sample at 70 m depth (taken from Brussaard (2006)). The plots indicate: (A) side scatter against green fluorescence and (B) events against green fluorescence.

2.6. Virus Techniques

2.6.1. *Filtrate*

Seawater was used to attempt to infect methanol-utilising marine bacterial strains with naturally occurring viruses, and was filtered through 0.22 µm filters (surfactant-free cellulose acetate, SFCA, membrane; Sigma-Aldrich, UK) to remove bacteria, whilst retaining as many virus-like particles (VLPs) as possible. The filtrate was then stored at 4°C (Figure 2.4). The SFCA filter was used for all filtrate in experiments, unless otherwise stated, based on findings in Chapter 3 which compared three 0.2 µm low-protein binding membrane filters for retention of VLPs. The membranes compared were surfactant-free cellulose acetate (SFCA), glass fibre cellulose acetate (GFCA; Sarstedt, Germany) and non-pyrogenic polyethersulfone (PES; Sarstedt, Germany). Filtrate used in experiments were all sampled from station L4 seawater from various depths (3 m, 10 m, 25 m and 50 m), station E1 (3 m and 60 m), Plymouth Sound harbour water (50° 21'48.3 N, 04° 08'48.6 W, collected at high tide) and pore water derived from sediment (as described later in Section 2.6.6).

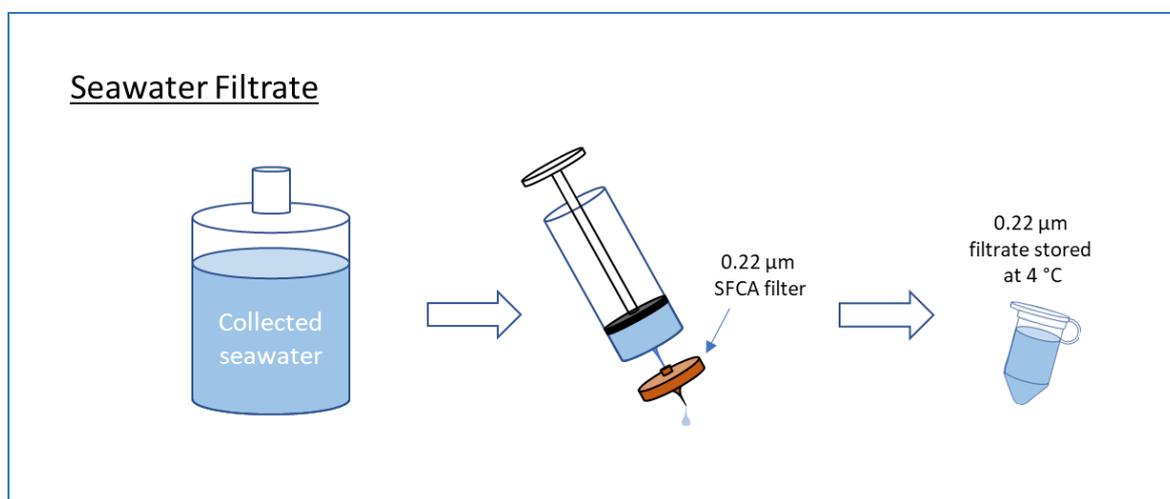


Figure 2.4 Flow visualisation of the preparation of seawater for use as filtrate in experiments.

2.6.2. Incubation variables

In addition to the variation in filtrate being tested against different bacterial strains (*M. marina* and station L4 *Methylophaga sp.* strain isolates) for a viral infection using plaque assays and spot tests, these two methods were also carried out at different incubation temperatures (17.5, 25 and 30°C) and the growth phases (early, mid and late exponential) when filtrate was added. Varying the incubation conditions was carried out due to the unknown conditions required to propagate a measurable viral infection with a methanol-utilising bacterium.

2.6.3. Growth of methanol-utilising methylotrophic bacteria in agar

Typically, plaque assay and spot testing methods have the carbon compounds added to the molten soft agar medium (0.4%, maintained at 43°C) to allow the bacteria to grow successfully. However, due to the volatile nature of methanol, it could not be added to the agar media and remain in a liquid phase because of the heating steps involved in assay set up. Therefore plaque assay and spot testing methods were carried out with headspace methanol as the carbon source (Section 2.5.3).

2.6.4. Plaque assay

To isolate novel viruses, seawater/benthic samples were filtered (0.2 µm SFCA; Sigma-Aldrich, UK) and the filtrate was retained for use in plaque assays as potential lysate (lysis is indicative of a viral infection). The locations targeted for filtrate sampling for use in assays were chosen due to the perception that viruses should exist in the same locations as the hosts they infect. Therefore, 100 µL of filtrate was placed into the first of a series of 1.5 mL microcentrifuge tubes, and 90 µL of filtrate was placed into the remaining microcentrifuge tubes and inverted. A serial dilution (10^{-1} – 10^{-5}) was created by taking 10 µL from the initial tube, placing it into the next tube and gently mixing by inverting 5 times. 10 mL of bacterial culture (exponential growth phase determined using optical density) was centrifuged (3000 g, 10 min^{-1}) in a 15 mL tube. The supernatant was removed from the tube, and the resulting pellet was re-suspended in 1.5 mL of fresh media (+ 300 µL for each control assay). Three control plates were also set up with all plaque assays, which included: agar and bacterial cells where growth should occur (positive control); agar and

filtrate where no growth should occur as a result of the filtrate only (negative control), and agar only where no growth should occur to indicate that agar is not contaminated (additional negative control). 300 μ L of re-suspended culture was combined with each of the filtrate dilutions and inverted mixed twice, before being left for 30 minutes at 4°C. Following this adsorption step, 500 μ L was removed from each vial in the dilution series and placed into a 7 mL vial. 2.5 mL of molten soft agar (0.4%, 43°C) was placed into each vial and vortex mixed for 1 second, before being quickly and gently poured on to the surface of a petri dish containing 1% agar media. The top agar was then spread across the surface of the pre-set bottom agar by sliding the plate in a circular motion and avoiding the formation of any bubbles. The top agar was then left to set for 30 minutes at room temperature and the plate incubated at 25°C with headspace methanol, and checked daily for the formation of circular clearings, or 'plaques', of non-growth in the top layer, which are indicative of a lytic interaction occurring.

2.6.5. Spot testing

Due to the time required in setting up plaque assays, a spot testing approach was used, with the advantage that a single plate can be used to test multiple filtrates. This would allow an increase in the number of bacterial strains being tested against filtrate (derived from environmental samples and from phage enrichment, as described in Section 2.6.6) for a lytic viral infection. 500 μ L of bacterial culture was grown up to mid exponential phase ($OD_{600} \geq 0.2$) and quickly vortex-mixed with 2.5 mL molten 0.4% agar, before being poured on to a 1% set agar plate (using the same media used in the culture). The top agar was then spread across the surface of the set bottom agar by sliding the plate in a circular motion and avoiding the formation of any bubbles, then left to set for 10 minutes in the bench. Once set, a pen was used to split the plate into fractions before the application of the filtrate. 20 μ L of filtrate was pipetted onto the bacterial lawn within the outlined sections and left until the droplet was absorbed into the top agar (typically 20 min); the plate was then incubated overnight at 25°C with headspace methanol. The plate was then monitored visually for clearings in each of the sections of the top agar; this method was used to compare multiple potential lysates against a single bacterial host.

2.6.6. *Phage enrichment in liquid culture*

The number of potential infection events occurring in a plaque assay approach is often limited; however, these events can theoretically be increased using a phage enrichment approach. Additionally, the same filtration and concentration methods used on environmental samples can be applied to the filtrate derived from phage enrichment. Therefore, an exponentially growing bacterial culture was grown up, and 30 mL of fresh media was inoculated with 1% of the pre-grown bacterial culture and 50 μ L of filtrate (0.2 μ m SFCA membrane, Sigma-Aldrich, UK), the freshly inoculated culture was then left to incubate at 25°C (150 rpm) (Phage Enrichment Method 1 - Figure 2.5). A positive control (media inoculated with culture only) and negative control (media inoculated with only filtrate) were also incorporated with all phage enrichments. Uninhibited growth should occur in the positive control, and no growth should occur in the negative control as a result of the filtrate being added to the media. Sub-samples were taken at desired time points, using a needle and syringe. These sub-samples were taken to check for any increase in the VLP abundance using AFC (Section 2.5.3), and a separate sub-sample was also taken for use in later enrichments, based on the findings of the AFC analysis. This additional sub-sample would be centrifuged (5,000 g; 10 min) and 0.2 μ m filtered (SFCA), before storing at 4°C until used as the filtrate in a plaque assay or spot test.

A variation of the phage enrichment method was used, as described by (Twest and Kropinski, 2009), where the volumes of the filtered seawater (0.2 μ m; SFCA) and the media were changed to 15 mL each, with 1% inoculum from an exponentially growing bacterial culture (Phage Enrichment Method 2 - Figure 2.5). Controls were also altered to accommodate the different approach, where the positive control switched the filtered seawater for deionised milliQ water, and the negative control included no culture inoculum. An additional phage enrichment approach was also carried out involving filtrate, produced by separating pore water from sediment as described by Cervený et al. (2002). Briefly, the top layer of sediment in a box grab was sampled into a 50 mL sterile tube. 20 g of the sampled sediment was then added to a 250 mL volumetric flask, and 50 mL of sterile MAMS media (10 mM MeOH) was also added. This sediment/media mixture was then incubated at 25°C for three hours (150 rpm). The resulting slurry was then centrifuged (8,000 \times g, 10 min⁻¹) and 1 mL of the supernatant was sub-sampled for virus-like particle abundance (Section 2.5.3) before and after being filtered through a 0.2 μ m SFCA

filter. This filtrate was then stored at 4°C until used as described previously for phage enrichment and a plaque assay (Phage Enrichment Method 3 - Figure 2.5).

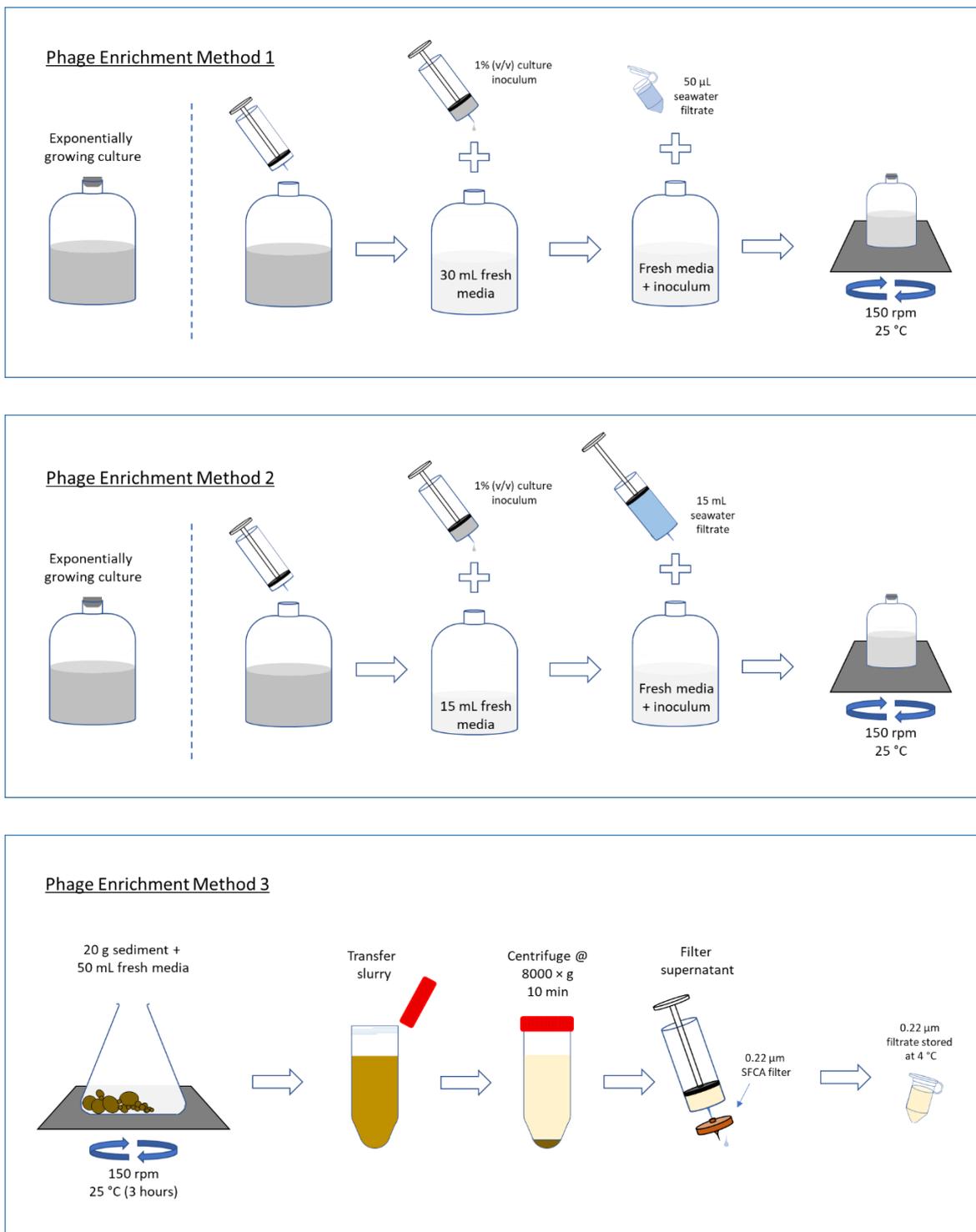


Figure 2.5 Flow diagram of the various phage enrichment methods utilised in experiments.

2.6.7. Prophage induction by ultraviolet, antibiotic and thermal stress

Induction of prophages existing within bacterial genomes was attempted using ultraviolet (UV) irradiation to damage host DNA, and potentially inducing a lytic response by any susceptible prophages. An agar plate with bacteria was prepared using the same method as Section 2.6.4, and 500 μL of bacterial culture was grown up to mid exponential phase ($\text{OD}_{600} \geq 0.2$) and quickly vortex-mixed with 2.5 mL molten 0.4% agar, before being poured on to a 1% set agar plate (using the same media used in the culture). The top agar was then spread across the surface of the set bottom agar by sliding the plate in a circular motion and avoiding the formation of any bubbles, then left to set for 10 minutes in the bench. This petri dish containing cells was then irradiated at 1500, 1000 and 500 $\times 100 \mu\text{J cm}^{-2}$ (Spectrolinker XL-1000 UV crosslinker, Spectroline, USA). These plates were then incubated with headspace methanol as per Section 2.2, and compared against a control plate over the following few days for any clearings, indicative of a viral lytic response.

A lysogenic bacterial species with a prophage can be induced by stressing the host with an antibiotic. This can be carried out by growing up a bacterial culture to exponential phase and inoculating 30 mL of fresh media with 1% of the pre-grown bacterial culture, bacterial growth was then monitored by spectrophotometer until an $\text{OD}_{600} \geq 0.2$ was reached. Once this optical density was achieved, 30 μL of mitomycin C (1 mg mL^{-1} ; Sigma-Aldrich, UK) was added to the bacterial culture and incubated for 30 minutes (25°C, 150 RPM). After the initial incubation period mitomycin C was removed from the culture by washing twice with fresh media by centrifugation (7500 g, 10 min^{-1}) and re-eluting the resulting cell pellet with 30 mL of fresh media. The culture was then incubated and monitored hourly for a drop in optical density or a change in bacterial abundance determined by AFC (Section 2.5.2).

A further lysogenic induction assay was also carried out using thermal stress as described in Fogg et al. (2011), briefly, a bacterial culture was grown to exponential phase and then 1% was used to inoculate 30 mL fresh media, bacterial growth was monitored by spectrophotometer until an $\text{OD}_{600} \geq 0.2$ was reached. The culture was then incubated in a water bath at 42°C for 30 min^{-1} . Following this heating step, the culture was removed and incubated (25°C, 150 RPM) for 2 hours. The culture was sampled for bacterial abundance determined by AFC. *Marinibacterium anthonyi* La6

was grown using Marine Broth 2216, as described in Section 2.2.2, due to its slower growth rate.

2.6.8. *Tangential Flow Filtration (TFF)*

To concentrate or remove viruses from seawater, a tangential flow filtration unit (Vivaflow 200 Units, Sartorius AG, Germany) was used with two polyethersulfone (PES) membrane pore sizes available - 100,000 MWCO and 0.2 μm . A TFF unit works differently from a standard filter (filtering based on particle size), instead back pressure permeates particles through a membrane, based on the molecular weight of the particles and the force exerted on the particles by a cross flow ('permeate' will be termed 'filtrate' for simplicity). The filtrate and retentate are separated, allowing a sample to be concentrated, diluted or filtered of particles based on the size of TFF unit membrane. The filtration unit was set up, used, washed and stored according to the manufacturer's instructions.

2.6.9. *PEG Precipitation and CsCl gradient purification*

Virus particles can be concentrated by using polyethylene glycol (PEG) precipitation, this concentrate was subjected to a caesium chloride gradient purification step. PEG 6000 (Fisher-Scientific, USA) was added to filtrate to 10% w/v (3 g/30 mL), and then Sodium Chloride (NaCl) was added to the PEG/filtrate mixture to 1 M concentration (1.75 g/30 mL). This mixture was then gently dissolved by repeated inversion at room temperature. The solutions were then left to stand on ice for 2-3 hours before being centrifuged for 20 min^{-1} at 8,000 \times g (4°C). The supernatant was discarded and the vial and resulting pellet were left upside down to dry for 15 min^{-1} , at room temperature. The pellet was then resuspended in 250 μL of 1 \times PBS buffer and stored at 4°C until used.

Purification of the resulting concentrated viruses was then conducted by caesium chloride (CsCl) gradient centrifugation. The PEG concentrate was made up to 12 mL with 1 \times PBS buffer, and subdivided in 4 \times 3 mL samples in 15 mL tubes. To these individual tubes, CsCl was added in the following amounts: 1.1989, 1.3417, 1.5185 and 1.7029 g. These tubes were then gently mixed until the CsCl had dissolved and, starting with the densest concentration first, were carefully layered over one another in an ultraclear ultracentrifugation tube (Beckman Coulter Inc.,

USA), so as the layers would not mix. The tube was then centrifuged at 25,000 rpm for 2 hours (15°C) in an ultracentrifuge (Optima L-100 XP Centrifuge, SW 40 rotor; Beckman Coulter Inc., USA).

2.7. Virus dilution and reduction experiments

2.7.1. *Virus reduction*

To determine the impact of viruses on methanol utilisation by microbes, the abundance of viruses in seawater was altered whilst retaining the same microbial abundance. Initially this was attempted by using a method described by Wilhelm, Brigden and Suttle (2002), whereby surface seawater from station L4 (Section 2.3) was collected and 900 mL was pre-filtered through a 100 kDa TFF filter (Section 2.6.8). 300 mL of the unfiltered seawater was then filtered through a 0.22 µm (PES) filter, whilst being filtered the reservoir volume was maintained at 300 mL using the 100 kDa TFF filtrate, effectively retaining the bacterial portion of the assemblage. The retentate was sampled for VLP and bacterial abundance (Section 2.5), and samples were also taken for methanol dissimilation/assimilation (Section 2.11) and bacterial production (Section 2.12). Due to unexpected reductions in bacterial abundance, this method was also amended with constant resuspension of the retentate with a pipette during filtration and also adding 3 mL bovine serum albumin (BSA, 1% w/v in 1 × phosphate buffered saline solution) to the filter for 1 hour before flushing with milli-Q water in an attempt to limit cells adhering to the filter.

2.7.2. *Virus dilution*

A dilution approach was also utilised, which was adapted from Evans *et al.* (2003). Briefly, surface seawater was collected from station L4 and pre-filtered through a 0.2 µm (PES) filter. The filtrate was then transferred into four 1 L Durans and combined with unfiltered seawater to make up dilutions of 20, 40, 70 and 100%, ending up with a dilution series of unfiltered seawater diluted with 0.2 µm. At each level of dilution, triplicate 200 mL Duran bottles were filled by siphoning, and triplicate unfiltered/undiluted bottles were also filled. These steps were then repeated with seawater filtered through a 100 kDa TFF filter (Section 2.6.8), to achieve a virus concentration gradient. Bottles were incubated at *in situ* temperatures for 24 hours and samples taken at T₀ and T₂₄. Samples were taken

for VLP and bacterial abundance, as described in section 2.5, and 1 mL samples were also taken for dissimilation and bacterial production rates as described in Sections 2.11.3 and 2.12, respectively.

2.8. Transmission Electron Microscopy (TEM)

The presence of virus-like particles in filtrate was confirmed by TEM. 20 μL of a filtrate was aliquoted onto a silicone sample prep grid (Agar Scientific Ltd., UK), and a copper loading grid (Formvar/carbon 200 mesh Copper; Agar Scientific Ltd., UK) was gently placed onto the sample droplet (copper side up) for 15 minutes in the dark. It was then briefly placed onto a droplet of milli-Q water (20 μL), then placed onto a 20 μL droplet of Uranyl Acetate (2% w/v; Agar Scientific Ltd., UK) to stain for 15 minutes in the dark, before being removed and placed onto another droplet of milli-Q water as a final wash step. The loaded and stained sample grid was then loaded into a 1400 TEM (JEOL Ltd., Japan) for visualising.

2.9. Extraction of nucleic acids

2.9.1. *Environmental DNA extraction*

2 – 3 L of seawater was filtered through a 0.22 μm Sterivex polyethersulfone filter (Millipore), from which the environmental DNA was then extracted. If the filter was frozen, then it was first defrosted before beginning the following steps. Once thawed, the smaller outlet of the filter was sealed with blu-tack and 500 μL of TE 10:1 buffer was added to the filter. The other end of the filter was also sealed and then the filter was thoroughly vortexed for 2 minutes, before the entire Sterivex filter was sonicated in ice-water for 15 minutes. The TE/retentate mixture was removed from the filter to a 2 mL Eppendorf using a syringe. An additional 500 μL of TE 10:1 buffer was added to the filter, sealed again and vortexed for 30 sec and added to the first aliquot. This was repeated with 200 μL , making the final mixture up to 1.2 mL. The Eppendorf tube was centrifuged (Maximum RPM; 2 min^{-1}), the supernatant was removed, and the pellet was re-suspended in 467 μL of TE 10:1 buffer. Then 30 μL of 10% (w/v) sodium dodecyl sulfate (SDS) and 3 μL of Proteinase K (20 mg mL^{-1}) was added and incubated for 1 hr at 37°C. An equal volume of phenol:chloroform:isoamyl

alcohol (25:24:1) was added, vortexed briefly and centrifuged (16,000 RPM; 5 min⁻¹). The supernatant was transferred to a sterile 2 mL Eppendorf tube and the previous step repeated once more and followed up once again using chloroform:isoamyl alcohol (24:1). The final supernatant phase was transferred to a sterile Eppendorf tube, along with a tenth of the supernatant volume of NaOAc (3 M; pH 7) and two and half volumes of 100% ethanol, and then incubated for 20 minutes at room temperature. The tube was then centrifuged (16,000 RPM; 30 min⁻¹), the supernatant transferred to a sterile Eppendorf tube, the pellet washed in 120 µL of ice-cold 70% ethanol and then centrifuged (16,000 RPM; 20 min⁻¹). The resulting supernatant was then removed; the tube was air dried and the pellet re-suspended in 50 µL of TE 10:1 buffer.

2.9.2. Genomic DNA extraction

DNA of a high molecular mass was extracted for use in sequencing and other downstream molecular processes using the CTAB (cetyltrimethylammonium bromide) method of Doyle & Doyle (1987). The isolate of interest was grown up to 50 mL, centrifuged (10,000 g; 10 min⁻¹; 21°C) in a 50 mL tube and the resulting pellet was re-suspended with 5 mL re-suspension buffer (20 mM Tris; 2mM EDTA; pH 8), 60 µL of lysozyme (100 mg mL⁻¹; Sigma-Aldrich) and incubated for 1 h at 37°C. 7 µL of RNase A and 375 µL of Proteinase K (10 mg mL⁻¹) was added and incubated for 15 minutes at 37°C. 780 µL of N-laurylsarcosine (10% w/v in re-suspension buffer) was added and incubated at 60°C for 1 h, whilst gently rotating. 1,012 µL of 5 M NaCl and 803 µL of warm CTAB (10% v/v in 0.7 M NaCl) was then added and incubated for 15 minutes at 60°C. An equal volume of phenol:chloroform:isoamyl alcohol (25:24:1) was then added, rigorously shaken and incubated for an additional 10 minutes at 60°C. The mixture was again thoroughly mixed and centrifuged (8,000 g; 5 min⁻¹), before transferring the supernatant to a new tube. This process was repeated twice, with 5-minute incubation steps at room temperature each time, and then once again using chloroform:isoamyl alcohol (24:1). Approximately double the volume of ice-cold 100% ethanol was added to the final aqueous phase to precipitate the nucleic acids, incubated for at least 1 h (or overnight) at -20°C, before centrifuging (17,000 g; 20 min⁻¹; 4°C). The pellet was washed twice in 1 mL 70% (v/v) ethanol, air dried and re-suspended in 600 µL of Tris 5 mM pH 8.5 buffer.

2.9.3. Pulsed Field Gel Electrophoresis (PFGE)

PFGE was performed according to the modified procedure set out by Sandaa, Short and Schroeder (2010), which was adjusted to obtain viral nucleic acid. The method was a variation of the more widely used agarose gel electrophoresis, where the electric field only flows in one direction, however with PFGE the orientation of the electric field switches periodically. This key difference allows larger nucleic acid fragments, as well as whole genomes, to be size fractionated effectively.

For PFGE analysis, viral filtrates were prepared from enrichment bacterial cultures and concentrated using tangential flow filtration. Equal volumes of viral concentrate and molten 1.5% low melting-point agarose (prepared in TE 10:1 Buffer) were combined, mixed and dispensed into 80 μ L plug molds (Bio-Rad Laboratories) at -20°C for 2 minutes until set. The plugs were removed from the molds and incubated in 5 mL lysis buffer (250 mM EDTA, pH 8, 1% SDS, 1 mg mL⁻¹ Proteinase K) overnight in the dark at 30°C. The plugs were then washed three times with TE 10:1 buffer for 30 minutes at room temperature. A 1% agarose gel in 1 × TBE Buffer was made up using the casting stand provided with the CHEF-DR II (Bio-Rad) apparatus. Molecular weight standard plugs were trimmed to the correct size and placed into the wells on either side of the set gel, along with the plugs containing the viral nucleic acid. The wells *containing the plugs and standards were overlaid with the remaining molten 1% agarose*. The gel was removed from the casting stand and placed into the electrophoresis module along with 2 L of 1 × TBE buffer, which was kept at 14°C using the cooling module. The gel was run under the electrophoretic conditions of 6 V cm⁻¹, with switch times of 2.5 seconds and a run time of 20 hours. Once the run was completed, the gel was removed and stained overnight with ethidium bromide (10 μ L per litre of milli-Q water).

2.10. Nucleic acid manipulation techniques

2.10.1. Quantification of DNA

Quick quantification of DNA amount and quality was conducted using a Nanodrop-1000 spectrophotometer (NanoDrop, ThermoFisher Scientific). Another method utilised for determining the DNA concentration within different gel bands, was to use

the GeneTools (Syngene, UK) software to compare a known quantity of DNA (DNA ladder) on an agarose gel to a known volume of DNA loaded on to the gel (See Section 2.10.3).

However, for preparation of DNA amplicons to be sequenced using a MiSeq sequencer, more accurate DNA quality and quantity analysis had to be undertaken at the Exeter Sequencing Centre, University of Exeter. For this purpose, the QuantiFluor ONE dsDNA System (Promega Corp., WI, USA) was used for fluorescent quantitation of DNA, in combination with the GloMax Explorer Microplate Reader (Promega Corp., WI, USA). This was carried as per the manufacturer's instructions. Additional analysis of samples before sequencing was carried out using a 2200 TapeStation (Agilent Technologies Inc., USA), with D1000 Reagents and D1000 ScreenTape (Agilent Technologies Inc., USA) as necessary. Protocol used according to the manufacturer's manual instructions.

2.10.2. *Standard polymerase chain reaction (PCR)*

PCR reactions were set up in 20 μ L volumes in a DOPPIO (VWR) thermal cycler. Typical reactions contained 10 \times CoralLoad Buffer and Taq DNA Polymerase (QIAGEN), including 2 mM dNTPs, 6 μ M forward and reverse primer (See Table 2-5) and 30 – 50 ng DNA. General thermos cycler settings used were: initial denaturation (95°C for 5 min⁻¹); 30 cycles of denaturing (95°C for 45 s), annealing (55°C for 45 s) and elongation (72°C for 1 min⁻¹); and a final elongation (72°C for 3 min⁻¹). Negative controls were performed using template-free reactions.

When PCR amplicons were to be used for Next-Generation Sequencing (i.e. Illumina MiSeq), a high-fidelity DNA polymerase was used (Q5 High-Fidelity DNA Polymerase, New England Biolabs Inc., MA, USA).

Table 2-5 PCR primers used in this study, and the positive control species used for each of the genes.

Name	Target Gene	Sequence (5' – 3')	Control Species	Reference
27F	16S rRNA	AGAGTTTGATCMTGGCTCAG	<i>Methylophaga marina</i>	Lane, 1991
1492R		TACGGYTACCTTGTTAGGACTT		
515F	16S V4 Region ⁴⁵	GTGYCAGCMGCCGCGGTaA		Apprill <i>et al.</i> , 2015
806R		GACTACNVGGGTWTCTAaT		
1003F	<i>mxoF</i>	GCGGCACCAACTGGGGCTGGT	<i>Methylocella silvestris</i> BL2	Neufeld <i>et al.</i> , 2007
1555R		CATGAABGGCTCCCATCCAT		
xoxF1F	<i>xoxF1</i>	TAYGCCGAYGGCAAGSTGST	<i>Methylocella silvestris</i> BL2	Taubert <i>et al.</i> , 2015
xoxF1R		CCGTCRTARTCCAYTGRTCGAA		
xoxF2F	<i>xoxF2</i>	GGCYTAYCAGATGACBCCNTGG	None available.	
xoxF2R		GCCTTRAACCAKCCRTCCA		
xoxF3F	<i>xoxF3</i>	GGHGAGWCCATSACVATGGC	<i>Methylocella silvestris</i> BL2	
xoxF3R		TCCATSGTKCCGTAGAA		
xoxF4F	<i>xoxF4</i>	TTYCCHAAAYAACGTNTAYGC	<i>Methylobacillus</i> MM2	
xoxF4R		GGRTTRCCHGTHCCGTAGTA		
xoxF5F	<i>xoxF5</i>	GAYGAVTGGGAYTWYGACGG	<i>Methylocella silvestris</i> BL2	
xoxF5R		GGYTCVTARTCCATRCA		

⁴ Primers used for amplicons to be sequenced using Illumina MiSeq required sequencing extensions for annealing to the Nextera flow cell as described in the Earth Microbiome Project (<http://www.earthmicrobiome.org/>) and is further described in Section 2.10.6.

⁵ These primers have a phosphorothioate oligonucleotide on the 3' end indicated by lowercase.

2.10.3. *Agarose gel electrophoresis*

Fragments of DNA were separated using a 1% (w/v) agarose gel, containing ethidium bromide (1 μ L per 100 mL molten agarose solution) in 1 \times TBE Buffer. GelPilot 1 kb and 100 bp+ DNA (QIAGEN) ladders were used when estimating DNA fragment sizes. Typically, gels were run for 40 minutes at 100 V. The separated DNA fragments were visualised using a UVP trans-illuminator (DBA Analytik Jena, US) or imaged using a G:BOX (Syngene, UK).

2.10.4. *DNA purification*

Amplified DNA resulting from PCR was purified using the QIAquick PCR purification kit (QIAGEN, Germany) and the QIAquick Gel Extraction kits (QIAGEN, Germany) was used for DNA extraction from excised agarose gel slices, both according to the manufacturer's recommendations.

In some cases, an additional ethanol precipitation step was utilised where NaCl contamination was present after the gel extraction procedure. Ethanol precipitation involved combining 2.5 volumes of cold ethanol (95% v/v with milliQ) with the volume of sample in a micro-centrifuge tube and storing at -20°C overnight. The sample was then centrifuged (17,000 \times g, -20°C) for 30 minutes. The supernatant was then removed by careful pipetting, before adding 1 mL of 70% ethanol and centrifuging for an additional 5 minutes to wash the pellet, this step was then repeated once more. The ethanol supernatant was then removed again, and the pellet allowed to air dry for 20 minutes before being eluted in the required volume of buffer.

2.10.5. *Sanger sequencing of PCR amplicons and analysis*

Nucleotide sequences were manually analysed using the CodonCode Aligner software (Version 5.0.2) package (CodonCode Corporation, MA, USA). 16S rRNA gene amplicons were sent for sequencing at MRC PPU DNA Sequencing and Services (University of Dundee, UK), using a 3730 DNA Analyzer. Sequence

chromatograms were analysed using CodonCode Aligner (Version 7.1.2, CodonCode Corporation, MA, USA) to assess sequence quality. Forward and complementary sequences were aligned using Multiple Sequence Comparison by Log-Expectation (MUSCLE) tool (<https://www.ebi.ac.uk/Tools/msa/muscle>, EMBL-EBI, UK). Sequences were then aligned using the BLAST program, with the nucleotide database (<https://blast.ncbi.nlm.nih.gov/blast/Blast.cgi>, NCBI, USA).

2.10.6. *Primers used for Illumina sequencing*

For DNA amplicons to be sequenced using Illumina sequencing technologies (MiSeq), the DNA amplicons required regions for Nextera XT (Nextera XT DNA Library Preparation Kit, Illumina, CA, USA) indices to attach to (known as adapter sequences), which were added on to the *16S (V4)* and *coxF5* primer sets (Table 2-6).

Table 2-6 Indicates the sequence that was attached to the 5' end of all primer sets used.

Primer	Sequence (5' – 3')
Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG
Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG

In addition to the primer extensions required for the Nextera indexing, a phosphorothioate oligonucleotide was placed on the 3' end of each primer sequence to limit endonuclease degradation. Both primer sets were also phased, using the phasing amplicon sequencing approach described by Wu *et al.* (2015), which yields many benefits to Illumina sequencing outputs. In short, this principle revolves around having replicate primer sets, with varied oligonucleotide spacers added (in this case limited to 5) in between the Illumina adapter sequence and the primer sequences (Table 2-7).

Table 2-7 Spacer sequences added in between Illumina adapter sequences and the specific gene primer sets. The same spacer sequences were used for *16S V4* and *xoxF5*.

Primer Sequence	Primer Direction	
	Forward	Reverse
1	No addition	No addition
2	T	T
3	CA	AC
4	ACACA	CTT

2.10.7. *Next Generation Sequencing (NGS) of PCR amplicons*

DNA extracted from time series, Arctic and Atlantic samples were amplified using primers targeting the V4 region of the *16S rRNA* gene and *xoxF5* (Table 2-5), both of which were designed with overhang adapter sequences for compatibility with the Illumina indexing and sequencing adapters (Section 2.10.6). PCR amplicons were then ran on an agarose gel (Section 2.10.3), and bands of the expected size were removed and purified (Section 2.10.4). Nextera XT indices (See Section 2.10.6) were attached to the gel purified PCR amplicons by a limited cycle PCR programme, and the resulting PCR amplicons were purified further using AMPure XP magnetic beads (Beckman Coulter, CA, USA). The final prepared libraries were purified again using AMPure XP beads, quantified and normalised before being pooled together. The final sequencing preparation was carried out by the Exeter Sequencing Service (Exeter, UK), which involved denaturing the library pools with NaOH, diluting with hybridisation buffer and heat denaturing before loading on to a reagent cartridge. The pools were finally sequenced on an Illumina MiSeq (Illumina Inc., CA, USA).

2.10.8. *Bioinformatic analysis of NGS amplicons*

16S V4 rRNA and *xoxF5* sequences were processed using the second iteration of the divisive amplicon denoising algorithm (DADA2; version 1.6.0) pipeline with R (Version 3.4.1). The DADA2 algorithm (Callahan *et al.*, 2016) is an alternative to other widely used algorithms which cluster sequencing reads into operational taxonomic units (OTUs) based on a fixed dissimilarity threshold (typically 97%) (Caporaso *et al.*, 2010). The DADA2 method is sensitive to single base-pair

differences between sequences (amplicon sequence variants or ASVs) and thus offers higher resolution taxonomy than previous methods, enabling differentiation between ecotypes (Eren *et al.*, 2013). The DADA2 pipeline was used to inspect sequence quality profiles, carry out filtering/trimming, removal of primer artefacts and chimeras, and to assign taxonomies against the SILVA database v132 (Quast *et al.*, 2013). Singletons and doubletons were also removed as part of the pipeline, except when alpha diversity analysis was carried out. Forward and reverse reads of the *16S V4 rRNA* amplicon reads were trimmed and filtered, where forward reads were truncated at 250 nucleotides and reverse reads at 200 nucleotides (forward reads trimmed at 240 nucleotides for *xoxF5* amplicons), based on the first 10,000 reads of each dataset. A parametric error model was produced separately for the *16S V4* and the *xoxF5* amplicon data, these models were then used in the DADA2 algorithm with the forward and reverse reads to infer ASVs, the read pairs were then merged. Following merging, sequence lengths were checked, and *xoxF5* reads ranged in length between 215 and 372 nucleotides, which was out with the expected range (possibly due to non-specific binding), and were therefore trimmed of sequences smaller than 320 and larger than 350 nucleotides. Identifiable chimeras within each sample were also removed. Taxonomic assignment of the ASVs inferred from the *16S V4 rRNA* reads were performed against the SILVA (Version 132) database using the naïve Bayesian classifier method (Wang *et al.*, 2007). However, assignment of *xoxF5* ASVs required a manually curated database (provided by Martin Taubert), derived from the *xoxF* database of Taubert *et al.* (2015) which was built by investigating genomes of methylotrophs and non-methylotrophs within the NCBI (Bethesda, MD, USA; <https://www.ncbi.nlm.nih.gov/nucleotide>) nucleotide database for gene sequences encoding PQQ-dependent dehydrogenases. This database consists of 233 total sequences (50 *mxoF*, 24 *xoxF1*, 5 *xoxF2*, 15 *xoxF3*, 25 *xoxF4* and 109 *xoxF5*), the 109 *xoxF5* sequences consisted of 87 full sequences to the species level and an additional 22 partial sequences to genus level.

Within the DADA2 pipeline, the DECIPHER package (Version 2.6.0; Wright, 2016) was used for multiple sequence alignment. The phangorn package (Version 2.4.0; Schliep, 2011) was used to construct phylogenetic trees using the maximum

likelihood tree estimation function, which first constructs a neighbour joining tree and then fits a maximum likelihood tree using the neighbour-joining tree as a starting point. For visualisations, the phyloseq package (Version 1.22.3; Mcmurdie and Holmes, 2013) was used, which allows the synthesis and organisation of different data types gathered and produced during the time series, and to then combine them into a single data object for manipulation. These objects were then used in conjunction with the ggplot2 package (Version 3.0.0; Wickham, 2009) to produce figures.

2.11. Microbial uptake rates

2.11.1. *Microbial methanol uptake rates*

Methanol can be utilised by microbes, and the rate of utilisation was determinable using ^{14}C -labelled methanol to investigate this assimilation and dissimilation of methanol. These two pathways allow the microbial community to reduce methanol to formaldehyde, and then either assimilate carbon into biomass, or to dissimilate to carbon dioxide for energy production. The assimilation and dissimilation rates were determined using the protocol described by Dixon et al. (2011) for environmental samples.

2.11.2. *Microbial methanol assimilation into biomass*

As different iterations of the protocol have been designed based on the samples being analysed, a general method is laid out in the text below. Samples of interest were decanted into acid washed (5% HCl acid) sterile polycarbonate bottles (305 mL of seawater samples), with three 'live' replicates and two 'killed' controls. Killed control vials were fixed with 15.25 mL of TCA (100% Trichloroacetic Acid, Sigma-Aldrich; seawater samples) or 100 μL of glutaraldehyde (25%; Sigma-Aldrich). All samples and controls then received 50 μL of ^{14}C -labelled methanol (specific activity 57.1 mCi mM^{-1} , concentration 0.1 mCi mL^{-1} ; American Radiolabelled Chemicals Inc., St Louis, MO, USA). Samples were then all briefly vortex-mixed and incubated

in the dark for approx. 3 hours at *in situ* seawater temperatures. The incubation period was then terminated by filtering the samples onto filters (47 mm 0.2 µm Supor-200 membrane filters; PALL, Life Sciences), twice rinsing the filters with milli-Q water and fixing with ethanol (80%) and left to air dry overnight in 6 mL polyethylene vials (Meridian Biotech. Ltd., UK). The final step before analysis of seawater samples involved suspending the air-dried filters in 3 mL of scintillation fluid (ProSafe FC+, Meridian Biotech. Ltd., UK). All samples were then analysed on the scintillation counter (Tri-carb 2910 TR, Perkin-Elmer).

2.11.3. *Microbial methanol dissimilation to carbon dioxide (CO₂)*

Microbial methanol assimilation rates were determined using ¹⁴C-methanol, which has been described previously (Dixon, Beale and Nightingale, 2011; Sargeant *et al.*, 2016). The methanol dissimilation rates of a sample of seawater could be determined by the addition of ¹⁴C-labelled substrate to 1 mL of sample in a screw-cap micro tube (Biosphere SC Micro Tube, 2.0 mL; Sarstedt AG & Co., Germany). Sample 'killed' controls were fixed with 100 µL of glutaraldehyde (2.5% final concentration) and all samples incubated in the dark for 1 h (at *in situ* temperature for seawater samples). All samples would then be 'spiked' with 50 µL of ¹⁴C-labelled methanol, and then incubated again under the same conditions for 1 h. The sample incubations would then be terminated by adding the following: 0.5 mL SrCl₂.6H₂O (1 M); 0.02 mL NaOH (1 M) and 0.1 mL Na₂CO₃ (1 M). All samples were briefly vortex-mixed and stored overnight in the dark. The samples would then be briefly vortex-mixed and centrifuged (16,000 × g, 10 min⁻¹), the supernatant aspirated and the pellet sample was washed twice with 1 mL of ethanol (80%). Following the final aspiration step the resulting pellet was eluted in 0.5 mL of NaOH (1 M, pH 11.7) and 1 mL of scintillation fluid (ProSafe FC+), briefly vortex-mixed and stored overnight in the dark. The 2 mL vials were then placed into 20 mL of liquid scintillation counting vials (Perkin-Elmer) and analysed in the liquid scintillation counter (Tri-carb 2910 TR, Perkin-Elmer).

2.11.4. Determination of methanol uptake rates

The 'killed' controls accounted for the cumulative abiotic processes, which includes other factors such as: chemiluminescence; interaction with sample tube walls; cross contamination; counting error and any chemical reaction with dissolved organic matter.

'Killed' control counts (d_c) were subtracted from sample counts (d) and were corrected for sample volume (v) and incubation period (t) per hour.

The scintillation counter analyses the disintegrations per minute (DPM) of a sample, these counts are then converted into Curies (Ci; 1 DPM = 4.5×10^{-13}), before then converting to nCi, as shown in Equation 1.

$$\left(\frac{d - d_c}{v t}\right) \times 4.5 \times 10^{-13} \times 10^9 = a \quad \text{Equation 1}$$

This product, a (nCi mL⁻¹ h⁻¹), was then utilised to determine the rate of methanol uptake (m , nmol L⁻¹ h⁻¹) by using the specific activity of ¹⁴C-methanol (specific activity 57.1 mCi mmol⁻¹), Equation 2.

$$a \times \left(\frac{1}{\text{specific activity}}\right) \times 1000 = m \quad \text{Equation 2}$$

$$BGE_M = \left(\frac{\text{assimilation}}{\text{assimilation} + \text{dissimilation}}\right) \times 100 \quad \text{Equation 3}$$

Equation 1 and Equation 2 were both used for determining methanol dissimilation and assimilation rates. The resulting rates could then be used to determine the bacterial growth efficiency for methanol (BGE_M) using Equation 3, which was

derived from Del Giorgio and Cole (1998), and is the amount of new biomass produced per unit of organic substrate assimilated.

2.12. Leucine incorporation rates

Incorporated ^3H -leucine into bacterial protein in seawater samples was determined by utilising the method outline by Smith and Azam (1992). A final concentration of 25 nM (calculated from the specific activity of 149 Ci mmol^{-1} and a concentration of 1 mCi mL^{-1} , American Radiolabelled Chemicals Inc., St. Louis, MO, USA) was added to 1.7 mL seawater samples in sterile 2.0 mL screw cap microcentrifuge tubes (Biosphere SC Micro Tube, 2.0 mL; Sarstedt AG & Co., Germany). Killed controls were conducted for all sampling time points by adding 85 μL trichloroacetic acid (100%; TCA) for a 5% final concentration (Smith and Azam, 1992). Triplicate samples were taken, with an additional killed control, and incubated at *in situ* seawater temperatures in the dark.

Following one-hour, incubated samples were removed, and assays terminated by adding 85 μL TCA (5% final conc.), before vortex mixing. The sample tube was then centrifuged for 10 minutes ($16,000 \times g$), and the liquid was aspirated off, making sure to aspirate just beneath the liquid surface and on the side of the tube that would have been closest to the centre of the centrifuge, thereby avoiding accidental aspiration of the invisible pellet. This pellet was washed twice with the addition of 1.5 mL TCA (5% final conc.) and mixed by vortexing. The final addition to the centrifuge tube was 1 mL of scintillation fluid (ProSafe FC+) and then mixed by vortexing, and then left overnight in the dark. The sample microcentrifuge tubes were then placed within 20 mL liquid scintillation counting vials and analysed in the liquid scintillation counter. Incubation times were kept to a minimum, as a radiotracer can become distributed throughout a cell by passive diffusion and can lead to over estimation of incorporation rates (Robarts, 1998).

2.12.1. *Determination of bacterial leucine incorporation rates*

Bacterial leucine incorporation rate was determined by taking the disintegrations per minute reading determined from the scintillation counter and inputting the value in to Equation 4.

$$\frac{(d - d_c) 4.5 \times 10^{-13}}{SA} = a \quad \text{Equation 4}$$

Where d is the disintegrations per minute (DPM) of the sample, and d_c is the DPM of the killed control. DPMs are then converted to Curies (Ci) by multiplication with 4.5×10^{-13} (1 DPM = 4.5×10^{-13} Ci). SA is the specific activity (Ci mmol⁻¹) of the ³H-leucine stock. The value of a , is then input into Equation 5.

$$\left(\frac{a}{t v}\right) \times 1000 = BLI \quad \text{Equation 5}$$

Where t is the incubation period (h), and v is sample volume (mL). The BLI is the bacterial leucine incorporation rate (pmol Leu L⁻¹ h⁻¹). Conversion factors were then applied to the BLI rate to convert into rates of bacterial carbon production.

2.12.2. *Conversion factors and bacterial production rates*

As bacterial production rates cannot be measured directly, estimations must therefore be inferred, and this can be carried out using related metabolic processes (Fuhrman and Azam, 1980; Ducklow, 2000). The method used in this research revolved around the measurement of radiolabelled ³H-leucine converted into protein (Simon and Azam, 1989a). As bacteria typically incorporate leucine into protein without prior transformation to other amino acids (Kirchman, K'nees and Hodson, 1985), and protein constitute half of the dry weight of bacterial cells (Robarts, 1998). However, one of the major limitations of this approach is the requirement of a conversion factor from leucine incorporation rates into bacterial carbon production

(Kirchman and Ducklow, 1993). This conversion factor requires measurements of different cellular components, which can be very difficult to measure routinely, therefore literature values are typically applied, termed a “theoretical conversion factor” (Kirchman and Ducklow, 1993). A theoretical conversion factor (TCF) of 1.55 kg C mol Leu⁻¹ determined by Simon and Azam (1989) from seawater collected off of the coast of California, USA, has been applied to a range of contrasting environments, including oligotrophic gyres, the Mediterranean and other productive coastal regions (Kirchman and Ducklow, 1993; Ducklow, 2000; Hoppe *et al.*, 2002; Lamy *et al.*, 2010; Dixon and P D Nightingale, 2012; Laghdass *et al.*, 2012; Joint and Smale, 2017). Alternatively, “empirical conversion factors” (ECF) can also be estimated using natural bacterial assemblages, and comparing the uptake of radiolabelled tracers with the increase in bacterial biomass production (Kirchman and Ducklow, 1993). An averaged ECF (0.88 kg C mol Leu⁻¹) determined by Calvo-Díaz and Morán (2009) is comparable with another averaged empirically derived conversion factor of 0.73 kg C mol Leu⁻¹ used by Dixon, Beale and Nightingale (2011) with leucine incorporation rates determined for the same sampling location used in this project (Station L4, Plymouth, UK). As the TCF of 1.55 kg C mol Leu⁻¹ has been used previously and recently in studies conducted in the WCO and is directly comparable with other studies within the marine environment, this was the conversion factor chosen to convert leucine to carbon (Sargeant, 2013; Joint and Smale, 2017).

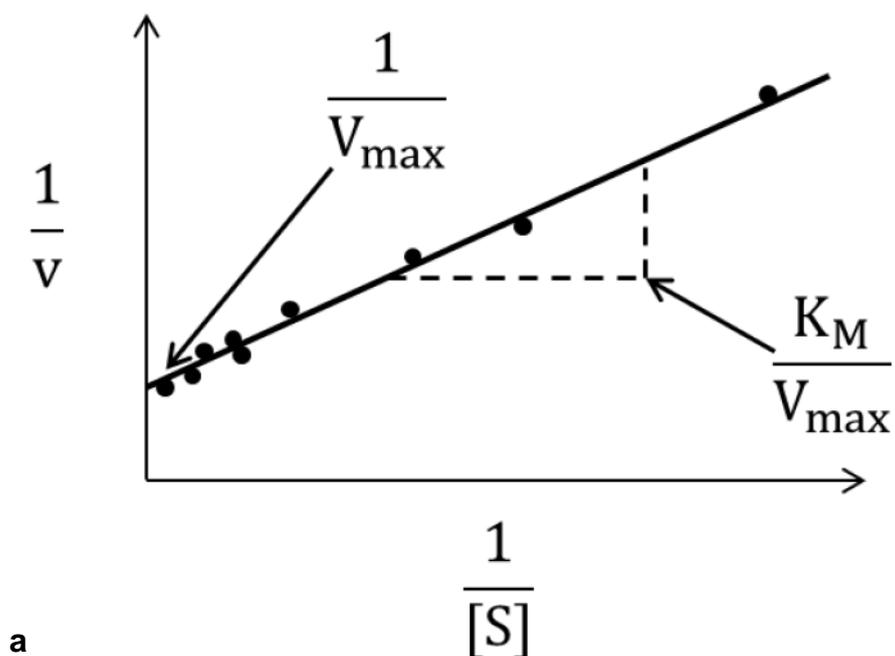
2.12.3. *Enzyme kinetics*

Michaelis-Menten kinetics has typically been used to describe single-substrate enzyme reactions and the uptake of organic compounds in water (Wright and Hobbie, 1965; Iriberry *et al.*, 1985). At higher substrate concentrations, the reaction rate will reach a theoretical maximum (V_{max}), where enzyme active sites are all utilised. This then determines the value of the Michaelis constant (K_m), which is described as the substrate concentration at half of the V_{max} value. Michaelis-Menten kinetic experiments were described using the Lineweaver-Burk transformation plot (Figure 2.6a), as suggested previously for the Western Channel Observatory (WCO), where a mixed population is present (Dixon, Beale and Nightingale, 2010;

Sargeant, 2013). Methanol kinetic experiments were undertaken at station L4 in the Western Channel Observatory in autumn 2017 (October) and spring 2018 (April) to highlight the variation of enzyme kinetics of methanol dissimilation rates, as has been previously described by Sargeant (2013), and using the Lineweaver-Burk transformation plot, both the V_{\max} and K_m values were calculated for methanol dissimilation (Figure 2.6b).

Table 2-8 V_{\max} and K_m values of methanol dissimilation rate kinetic experiments conducted at station L4 in the Western Channel Observatory.

Month	Year	V_{\max} (nmol L ⁻¹ h ⁻¹)	K_m (nM)
<i>October</i>	<i>2017</i>	3.5	145
<i>April</i>	<i>2018</i>	55.6	894



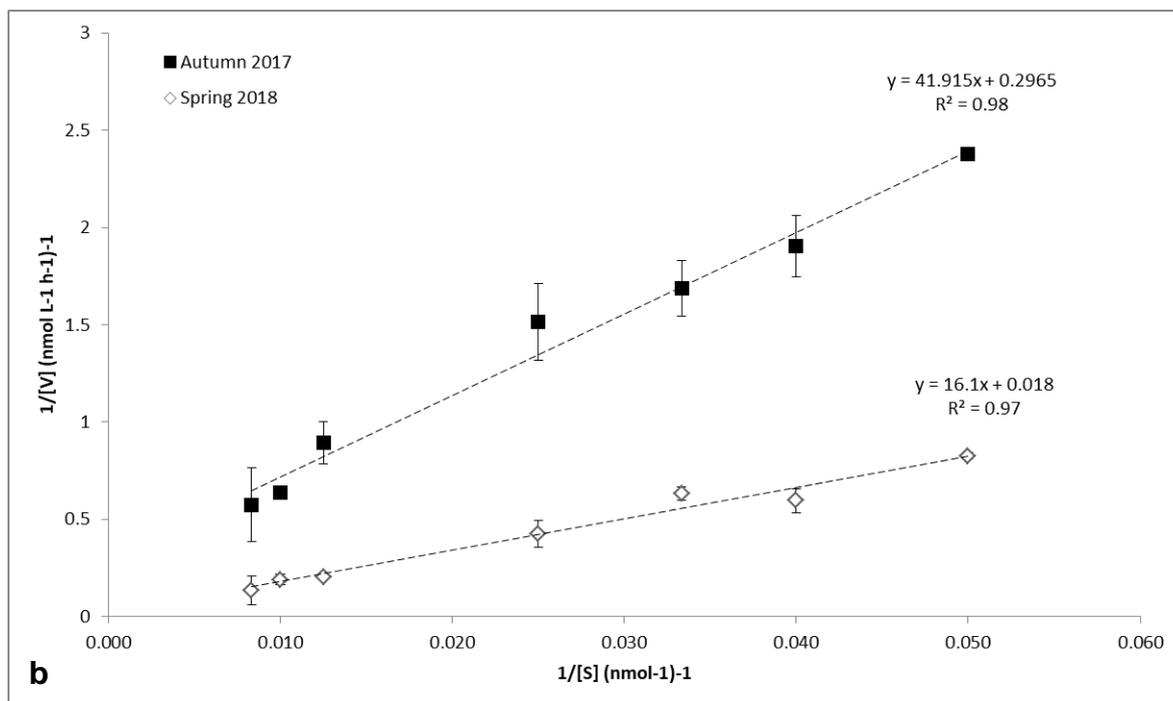


Figure 2.6 a) A model Lineweaver-Burk transformation plot and b) Lineweaver-Burk transformation plot of methanol dissimilation rates at station L4 in autumn 2017 and spring 2018.

2.13. Statistical methods

Microbial methanol uptake rates, environmental variables and MiSeq sequencing data from the WEC were imported into statistical package PRIMER v7 (Clarke and Gorley, 2015) and statistical analyses were performed as summarised below.

2.13.1. *Data pre-treatment*

Single missing data points were replaced with averages values determined by the preceding and subsequent data value. All environmental variables and sequence data were log transformed before calculation of Euclidean distance and Bray-Curtis similarity matrices. Additionally, normalisation of environmental variables was carried out to mitigate variation driven by different unit and scales.

2.13.2. *BEST: BIO-ENV and BV-STEP*

BEST is a multivariate statistic utilised to best explain sample or community patterns, by using environmental variables or species (BIO-ENV and BV-STEP, respectively) (Clarke and Gorley, 2015). BIO-ENV was used to determine relationships between different environmental variables and microbial uptake rates of methanol. However, BV-STEP is designed to determine what statistical support groups in the bacterial community had for driving methanol uptake rates.

2.13.3. *Non-metric multidimensional scaling (NMDS)*

Non-metric multidimensional scaling (NMDS) was utilised to investigate variables in time and space, with plots giving a visual representation of similarity. These work by translating information from pairwise 'distances' among different variables into a configuration mapped on a plot, where distances between points reflect community or environmental resemblances (Clarke and Gorley, 2015). Clusters were then overlaid on to plots to identify statistically distinguishable groups amongst the points.

2.13.4. *RELATE*

The RELATE test was employed as a means of testing for seasonal cyclicity and trends in community patterns in resemblance matrices.

Seriation was tested by taking the data and comparing on a linear sequence, with the closest samples being closer in similarity to each other and the samples furthest from one another being indicative of the largest difference among samples. This was applied to time series samples to statistically test inter-annual variability.

Similarly, to seriation, cyclicity compared adjacent time series samples found to be most similar to one another, however on a seasonal context which would progress throughout the year and gradually become more similar to the beginning of the year.

Chapter 3

Virus-like particles associated with an obligate methyloroph

3. Virus-like particles associated with an obligate methylotroph

3.1. Introduction

Methanol is an oxygenated volatile organic compound prevalent in the oceans and troposphere, which can be utilised by methylotrophic microbial organisms in the ocean (Dixon, Beale and Nightingale 2011). Extensive research has been carried out to characterise the molecular pathways of methylotrophic bacteria, which can uptake one-carbon compounds (such as methanol) as a sole carbon and energy source (Anthony, 1982; McDonald and Murrell, 1997; Anthony and Ghosh, 1998; Skovran *et al.*, 2011). *Methylophaga marina*, Gram-negative bacterium, which is a type strain that has been used extensively as the 'work-horse' system for investigative studies of methylotrophs in marine environments due to its extant nature and ease to culture (Janvier *et al.*, 1985; Dumont, 2010). These methylotrophic organisms are involved in the global methanol budget, and microbial uptake has been indicated as the main methanol loss process occurring in the oceans (Heikes, 2002). However, Sargeant *et al.* (2016) determined that microbial methanol utilisation rates were independent of *in situ* methanol concentrations, which suggests that other variables are likely driving spatial variability in methanol utilisation.

One such variable is believed to be viruses, which are known to be highly abundant in aquatic and marine environments (Bergh *et al.* 1989). It is generally accepted that viruses influence all marine organisms in some way, either through direct infection or indirectly (Section 1.3.3), with bacteria being one such group of organisms shown to have a complex affiliation with their infecting viruses (Clokier *et al.*, 2011; Breitbart *et al.*, 2018). Phage play a substantial role in oceanic biogeochemistry and bacterial population dynamics, frequently mediated by the lytic infection cycle and the subsequent release of organic matter (Poore *et al.*, 2004). This provisioning role could prove important to methylotrophs, as the release of cellular material may also provide a significant *in situ* source of methanol, potentially explaining unknown sources highlighted by the literature (Dixon, Beale and Nightingale 2011a; Beale *et al.* 2013). This has previously been shown to occur in phytoplankton, where

methanol has been produced by decaying detritus, and additionally has shown that methanol is retained within cells (Mincer and Aicher, 2016).

In addition to the well-characterised lytic cycle, temperate phage can lysogenise the host and integrate in to the bacterial genome, allowing replication with progeny cells until an environmental trigger event results in conversion to the lytic cycle (Weinbauer and Suttle, 1999). However, the exact interplay between host and virus is still a topic of intense debate, especially with lesser researched modes of infection, including pseudo-lysogenic (Cenens *et al.*, 2013) and chronic infections (Garoff, Hewson and Opstelten, 1998; Howard-Varona *et al.*, 2017), which are still disputed (Weinbauer, 2004; Clokie *et al.*, 2011; Breitbart *et al.*, 2018). As a robust lytic system gives reproducible results which are easy to detect, it can be seen as typically easier to investigate, and lytic viruses have been isolated from methylotrophic bacteria previously (Oki, Nishida and Ozaki, 1972; Ichikawa, Tahara and Hoshino., 1977; Lin, Wang and Wu, 1983). However, this research was carried out in the 1970s and '80s and only particle morphology analysis was attempted (Figure 1.11), with no further work carried out since the viruses of obligate methylotrophs.

In this chapter the assumption was made that viruses will likely infect methylotrophs, therefore the relationship between methanol-utilising methylotrophs and their associated viruses was investigated using a variety of culture-dependent approaches, culminating in the discovery of virus-like particles associated with *Methylophaga marina*. Additionally, the intention was to establish a robust virus-host system to be utilised in further chapters. Environmental samples were taken from station L4 (2 m, 10 m, 25 m, 50 m depths and benthic samples), to isolate methanol-utilising methylotrophs and to be used to induce infection in methylotrophic lab strains. Methodologies are further described in Chapter 2.

3.2. Isolation of methanol-utilising methylotrophic bacteria

Surface seawater sampled from station L4 was enriched with different concentrations of methanol and lanthanum to select for methylotrophic growth of organisms extant within the Western English Channel. These enrichments were then spread on to agar plates, and incubated to allow colony formation (Figure 3.1; Section 2.4). 204 colonies were picked from the agar plates and isolated further by streaking on to specific medium agar plates (90 isolates on MAMS agar plates and 114 on MBM agar plates), and of these isolations, 9 strains grew to a specified optical density threshold ($OD_{600} \geq 0.15$ after 24 hours), which was determined to be optimum for designing future repeatable experiments. All these strains were investigated further by amplification and sequencing of the *16S rRNA* gene. Eight strains had identities of $\geq 98\%$ to previously characterised *Methylophaga* species, with one strain having a lower identity of 90%. The majority of the isolate sequences indicated significant similarity with the lab strain, *M. marina* strain 222 (Section 2.2.1). DNA from these isolates were used in PCR assays to test for the presence or absence of two *xoxF* gene clades (using primers for *xoxF4* and *xoxF5* listed in Table 2-5) indicated to be important in marine one-carbon (C1) cycling and involved during growth on methanol (Taubert *et al.*, 2015).

All tested strains had the *xoxF5* functional gene and strains 3, 7 and 18 had the *xoxF4* functional gene. Strains 3, 7 and 18, contained the *xoxF4* gene and phylogeny aligning with *M. marina* strain 222 (99%). However, these strains differ based on the enrichment conditions, with strains 3 and 7 isolated from enrichments with the same media and methanol concentrations (MAMS and 10 mM, respectively), but strain 7 was isolated using 5 μ M lanthanum and strain 18 was able to grow on both MAMS and MBM media using a lower methanol concentration of 5 mM. As mentioned previously, lanthanum was used in the enrichment process, due to its recently highlighted importance in the growth of methylotrophs which utilise the *xoxF* gene, a homolog of *mxoF* (Keltjens *et al.*, 2014; Howat *et al.*, 2019). Due to the high sequence identity, characterisations and enrichment conditions, strains 4 and 5 were thought to be the same at the strain level. Strain 10 was isolated from an enrichment using MBM media and lower methanol and lanthanum concentrations (5 mM and 5 μ M, respectively), but phylogeny didn't align to any of the more

common alignments of *M. marina* strain 222 and NBRC (Table 3-1; reference strains from the NCBI database), and not to the species level with *Methylophaga* sp. In general, it has been recommended that when delineating between bacterial species, a 16S rRNA gene sequence similarity of <97% represents a new species (Stackebrandt and Goebel, 1994; Janda and Abbott, 2007), therefore due to the differences in phylogenetic alignment, this indicates that strain 1 was potentially a different species, however without further genomic information this can only be speculated at. Phylogenetic analysis of the various isolates was carried out to compare 16S rRNA gene sequences against other *Methylophaga* species and representatives from the Gammaproteobacteria class (Figure 3.2). All isolates group with the *Methylophaga* species likely indicating that they are also *Methylophaga* species, and all of the isolates (except L4-10) group next to *M. marina* (highlighted in red), likely indicating that they are strains of *M. marina*. Isolate L4-10 appears to relate closer to *Methylophaga nitratireducenticrescens* strain JAM1, so was most likely not an *M. marina* strain, however isolate L4-1 still groups with the *M. marina*.

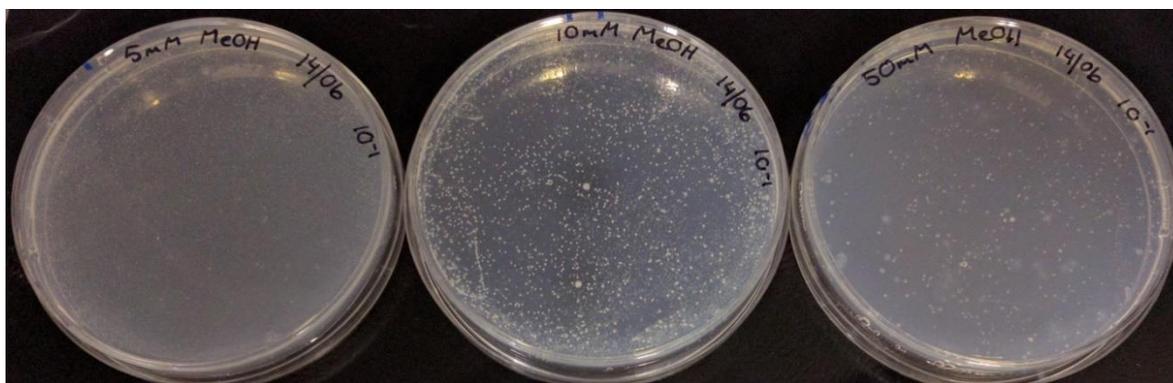


Figure 3.1 Agar plates spread with methanol seawater enrichments and incubated for one week at 25°C. These specific plates include agar plates spread with methanol enrichments of increasing methanol concentration from left to right (5 mM, 10 mM and 50 mM methanol).

Table 3-1 The enrichment conditions that each strain was cultivated from, phylogeny, and some basic characterisation of the strains isolated from station L4 in the Western Channel Observatory. Seawater samples were collected from the surface of station L4 on 01/06/16 (SST = 12.8°C and salinity = 35 PSU).

Strain	Enrichment Conditions			Phylogeny	Presence of functional genes		Identity (%)
	Media (MAMS/MBM)	MeOH Conc.	La ³⁺ Conc.		xoxF4	xoxF5	
L4[1]	MAMS	10 mM	-	<i>M. marina</i> strain 222	-	+	90
L4[2]	MAMS	10 mM	-	<i>M. marina</i> strain NBRC	-	+	99
L4[3]	MAMS	10 mM	-	<i>M. marina</i> strain 222	+	+	99
L4[4]	MAMS	50 mM	-	<i>M. marina</i> strain 222	-	+	99
L4[5]	MAMS	50 mM	-	<i>M. marina</i> strain 222	-	+	99
L4[6]	MAMS	5 mM	5 µM	<i>M. marina</i> strain NBRC	-	+	98
L4[7]	MAMS	10 mM	5 µM	<i>M. marina</i> strain 222	+	+	99
L4[10]	MBM	5 mM	5 µM	<i>Methylophaga</i> sp.	-	+	99
L4[18]	MAMS/MBM	5 mM	5 µM	<i>M. marina</i> strain 222	+	+	99

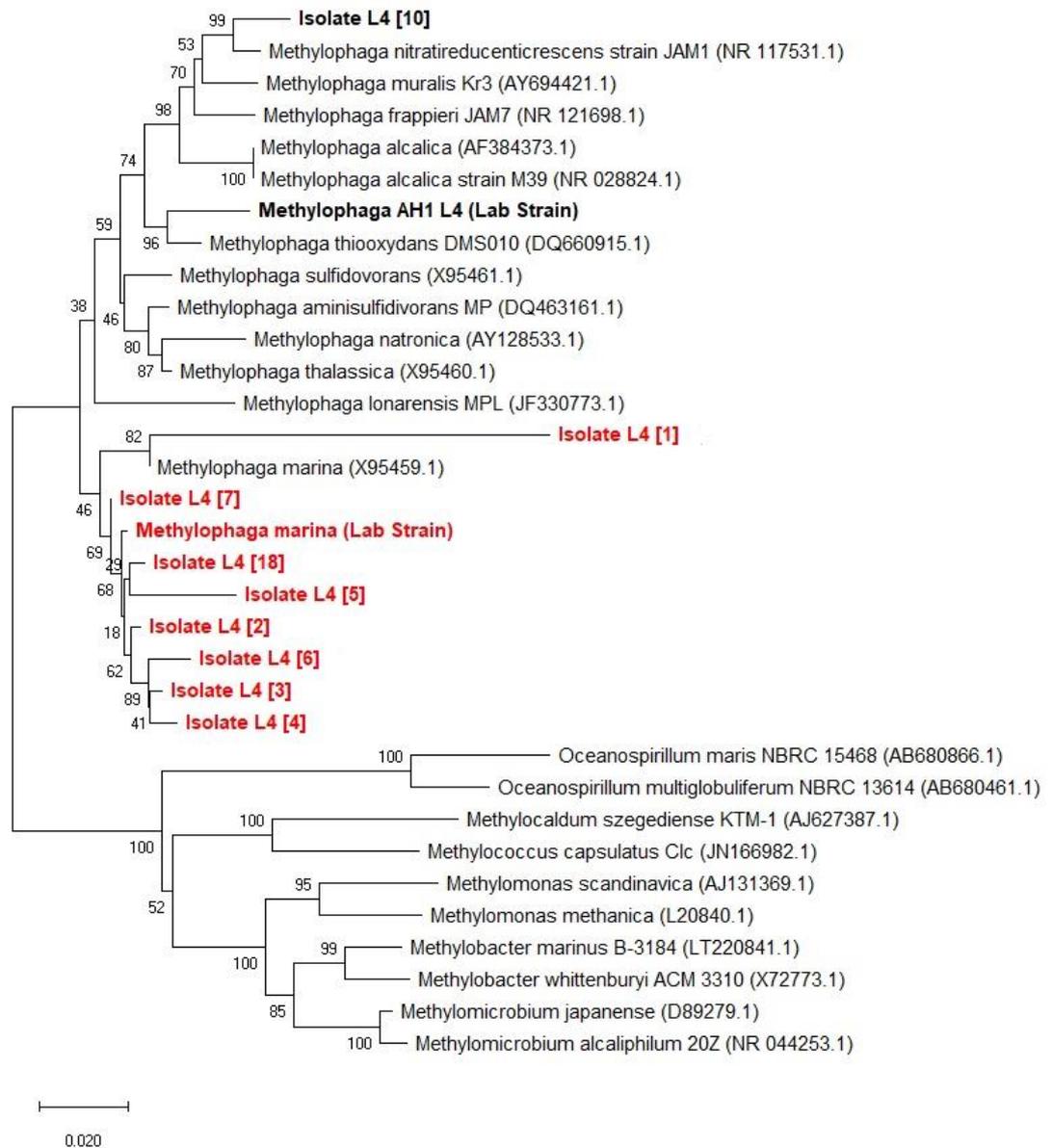


Figure 3.2 Phylogenetic analysis of 16S rRNA gene sequences of L4 isolates and lab strains (both in bold) and sequences of members of the genus *Methylophaga* was determined using an unrooted neighbour-joining phylogenetic tree. The bootstrap consensus tree was inferred from 2,000 resamples, and the numbers at branches indicate the percentage of those resamples. The scale bar is indicative of the nucleotide substitutions per position. 16S rRNA sequences from Gammaproteobacteria representatives are also shown for comparison. Numbers in parentheses are GenBank accession numbers. Sequences were aligned using the MUSCLE algorithm before the tree was produced using MEGA X (Kumar *et al.*, 2018).

3.3. Virus-like particle (VLP) abundance

Abundance of virus-like particles (VLPs) in filtrate was compared between different filters to ensure that a sufficient number of VLPs were passing through the filter for downstream phage isolation and enrichment studies (Section 2.6.1). Viruses can be isolated from an environmental water sample by passing through a filter (typically 0.2 µm), which removes the bacterial fraction and retains the viral fraction. However, although viruses may be physically small enough to pass through pores in a filter membrane, particle orientation or the wrong filter can retain much of the virus portion due to membrane surface adsorption characteristics. Therefore, three different low-protein binding filters were compared to determine which filter would provide the best recovery of VLPs (Figure 3.3) - surfactant-free cellulose acetate (SFCA) membrane, glass fibre cellulose acetate (GFCA) membrane, and non-pyrogenic Polyethersulfone (PES) membrane.

When comparing the filtered samples against an unfiltered sample, the filters retained 30 – 42% of the total VLPs (SFCA - 30.2%, GFCA – 32.7% and Non-pyro PES – 41.5%). Associated percentage errors are reported with actual values (Figure 3.3). Across the different virus subgroups (VLP 1, VLP 2 and VLP 3, as described by Brussaard (2006)), the SFCA membrane filter retained the least VLPs (VLP 1 - 26.9%, VLP 2 - 27.1% and VLP 3 - 82.5%) with the GFCA and non-pyro PES membranes retaining ~10% more VLPs of the VLP 1 and VLP 2 subgroups. Although the SFCA filter allowed the most VLPs to pass through for all virus subgroups, the VLP 3 subgroup was the most retained in all filters, with 82.5 – 91.0% of VLPs not passing through the filter membranes. Although there was a difference in total VLP retention between the SFCA and the GFCA filters, this was not statistically significant ($n = 4$, $P = 0.64$). The VLPs from the VLP1 subgroup were least retained using the SFCA membrane (26.9%), and the most was retained by the GFCA membrane for VLP 3 (89.1%). Therefore, as the passage of VLPs was greatest through the SFCA filters, this was the filter chosen for all future filtrate preparation.

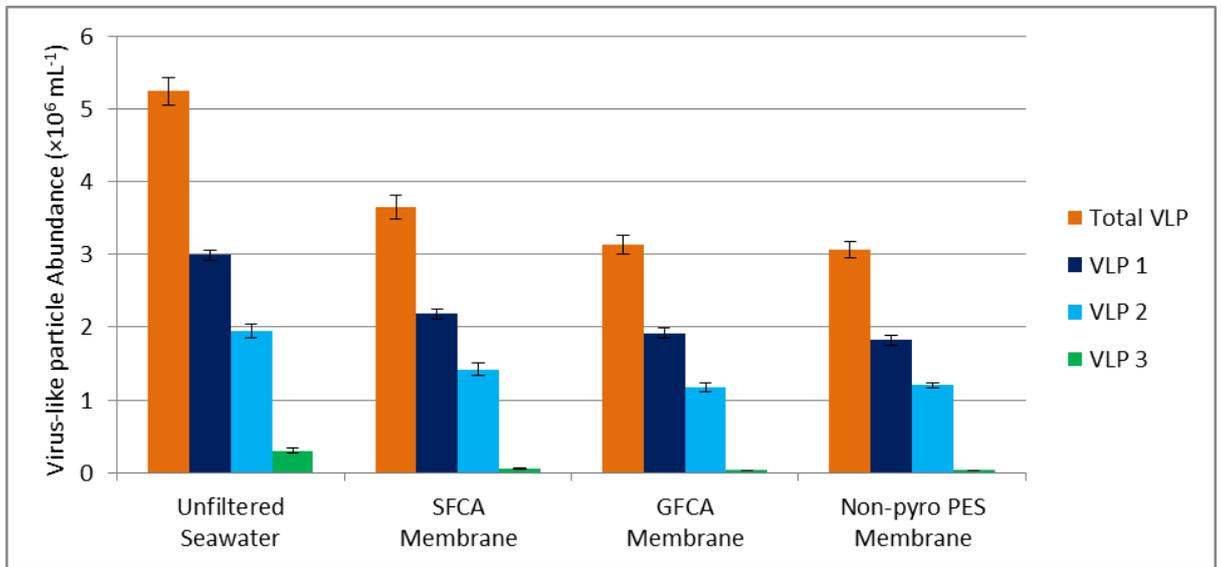


Figure 3.3 Comparison of virus-like particle groupings (VLPs) in unfiltered surface seawater collected in January 2017 from station L4, and filtrate from three different filters. SFCA; surfactant-free cellulose acetate membrane, GFCA; glass fibre cellulose acetate membrane, Non-pyro PES; Non-pyrogenic Polyethersulfone membrane.

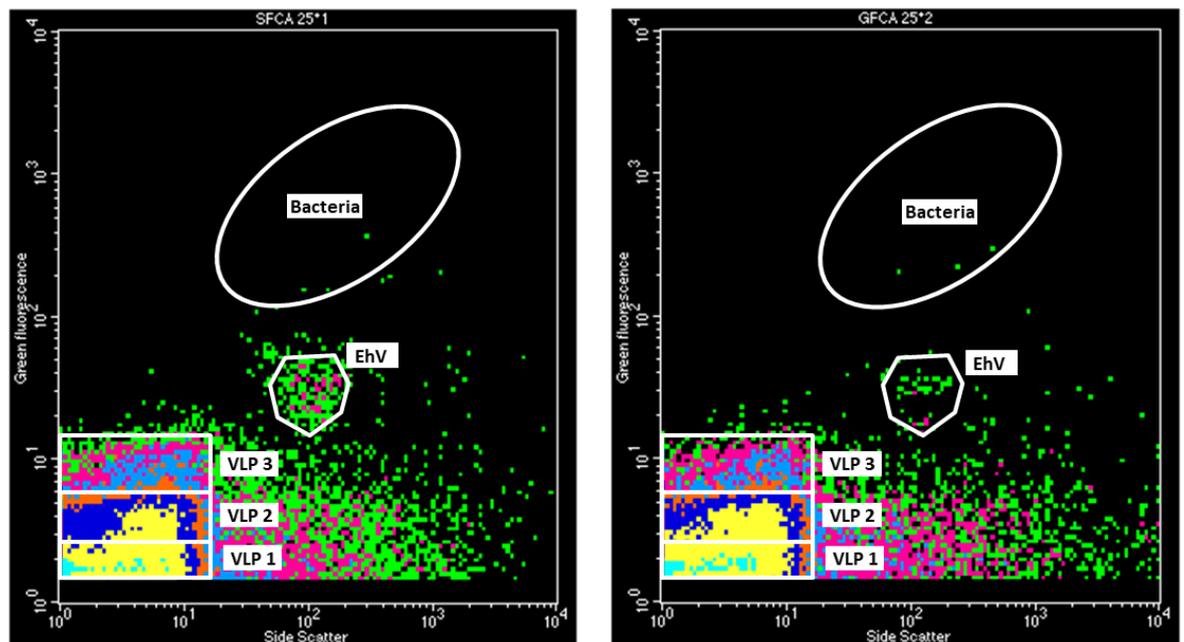


Figure 3.4 AFC scatter plots of samples taken from the SFCA (left) and GFCA (right) filter, with gates added on indicating VLP 1, 2 and 3. Additional gates are used for estimating *Emiliana huxleyi* virus (EhV) abundance and bacterial abundance.

3.3.1. Abundance of virus-like particles from environmental samples

With the aim of infecting a methanol-utilising methylotroph with a virus, samples to be used as potential lysates were taken from two sampling locations (station L4 and Plymouth Sound harbour) and different water column depths (3 m, 10 m, 25 m and 50 m), as well as pore water derived from sediment (sediment sample prepared as described in Section 2.6.6). In each sample VLP abundances were determined using AFC (Section 2.5.3; Figure 3.5). These sampling sites were chosen because the *M. marina* type strain was initially isolated from coastal waters (Janvier *et al.*, 1985), and as the earlier findings in Section 3.2 indicated, this was also a site where this strain can be isolated, further emphasising why this location made sense to investigate.

Highest total VLP abundance was found in the sediment sampled from station L4 (3.9×10^7 VLPs mL⁻¹), followed by the harbour sample (3.2×10^7 VLPs mL⁻¹). Total VLP abundances of $4.8 - 5.4 \times 10^6$ VLPs mL⁻¹, were discovered across all L4 water column samples, and the highest total VLP abundance within the water column was found at 10 m depth. Amongst all samples, the proportions of the subgroups making up the total VLP abundance was in descending order VLP 1, 2 and 3, except in the sediment sample where VLP 2 was the dominant virus sub group (Figure 3.5, 2.3×10^7 VLPs mL⁻¹). The VLP 1 sub group which makes up 34% of the total VLP abundance for sediment varies between 51 and 57% for all water column depths and makes up 71% for the harbour water (Figure 3.6). VLP 3 makes up the lowest proportion of the total VLP abundance: 1% (harbour seawater), 5% (sediment) and 6 - 7% (water column depths).

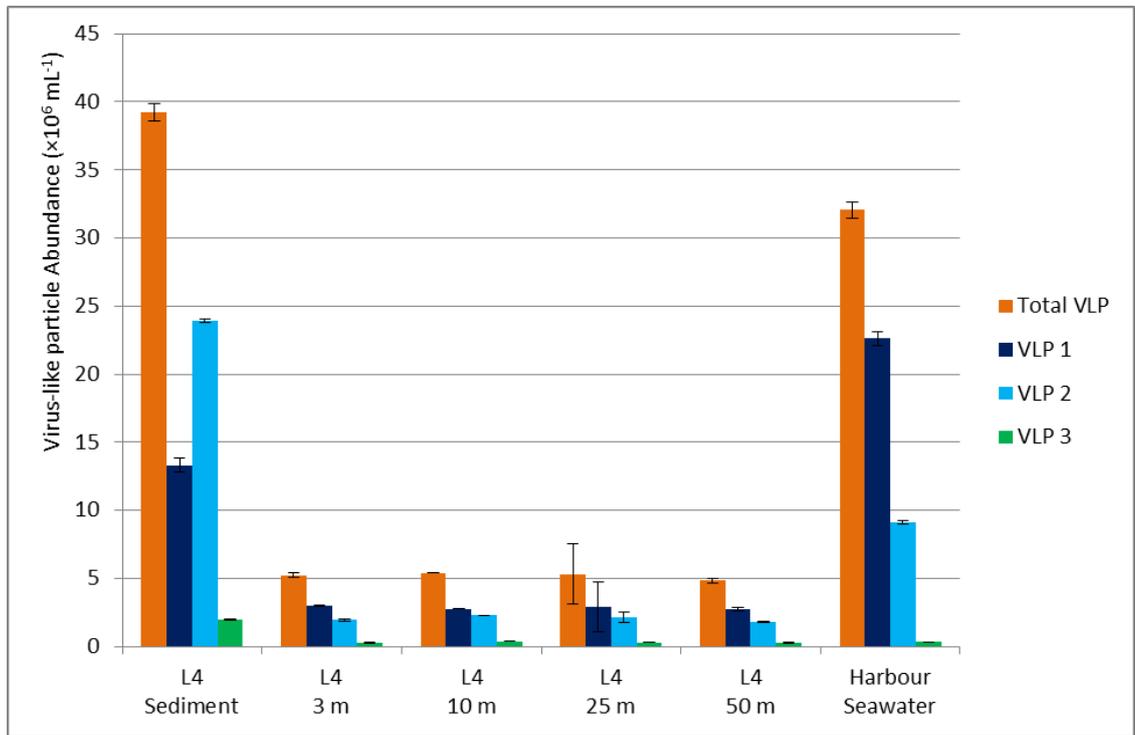


Figure 3.5 Abundance of total virus-like particles (VLP) and the three sub fractions (VLP 1, 2 and 3) in samples collected from the different depths and sediment pore water of station L4 and a harbour in the Plymouth Sound ($50^\circ 21'48.3 \text{ N}$, $04^\circ 08'48.6 \text{ W}$).

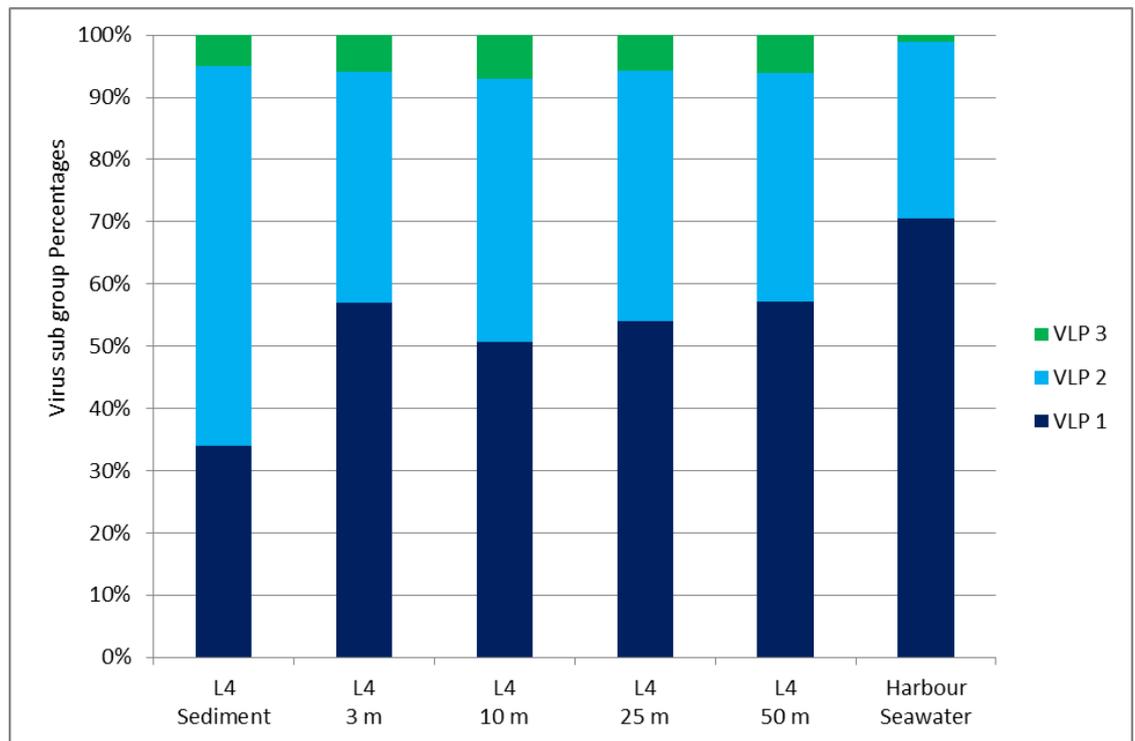


Figure 3.6 Percentages of each of the individual VLP subgroups which make up the total VLP abundance for each of the samples analysed.

3.4. Growth Rates

After sequencing the 16S rRNA PCR products led to the aligning of the sequences with *Methylophaga* sp. (Table 3-1), the growth curves of the strains isolated from station L4 were compared. The strains were grown in the media relevant to their enrichment conditions, with equal methanol and lanthanum concentrations (10 mM and 5 µM, respectively), and entered exponential phase of growth after an incubation period of 12 hours, with growth typically plateauing to stationary phase after 18 hours (Figure 3.7). This matched the growth curve of the *M. marina* type strain when grown using the same methanol and lanthanum concentrations and MAMS media. For downstream experiments, it was important to be working with isolates which grew at the same rates under similar conditions, thereby simplifying future experimental set ups. Additionally, avoiding the bias of physiological growth conditions was important as infection rates in bacteria are influenced by host growth rates (Middelboe, 2000). Once growth rates were known for strains, they could be used in virus inoculation assays and compared to determine susceptibility to virus infection, and thus further characterisation of the infection cycle.

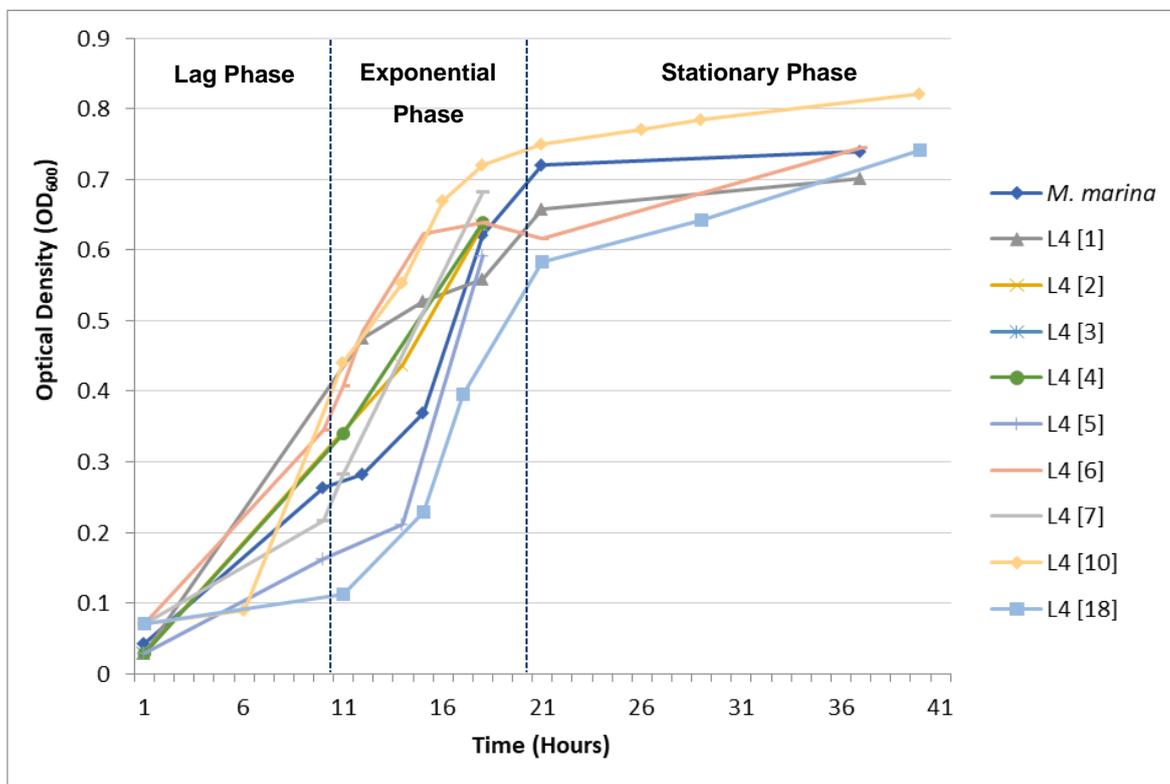


Figure 3.7 Optical density comparison of the growth rates of *M. marina* and the enrichment isolates from station L4. Isolates and *M. marina* were grown using MAMS media (except strain 10, which was grown using MBM media) and the same methanol and lanthanum concentrations (10 mM and 5 μ M, respectively).

3.5. Phage Inoculation

3.5.1. *Growth of methanol-utilising methylotrophic bacteria in agar*

Plaque assay and spot testing methods are techniques where a host organism is grown in an agar matrix and a potential lysate is applied to determine if a lytic relationship exists. As both methods involve bacteria grown in agar media, a slightly different approach had to be developed for methanol-utilising methylotrophic bacteria (Section 2.2). Three bacteria culture/soft agar volume mixtures were compared to determine which would result in the best bacterial growth in soft agar with headspace methanol available as the carbon source: 1) 100 μ L bacterial culture + 2.9 mL soft agar overlay media, 2) 300 μ L bacterial culture + 2.7 mL soft agar overlay media and 3) 500 μ L bacterial culture + 2.5 mL soft agar overlay media. All three mixtures grew successfully after 24 hours using this method, but mixture 3 yielded the highest turbidity - indicative of growth. Therefore, this was the mixture

ratio of bacterial culture and soft agar media used for further plaque assays and spot testing.

3.5.2. *Plaque assay*

Visual inspection of the plaque assay plates revealed that all *Methylophaga* strains grew uninhibited in the soft agar of plates, looking clear within the first 10 hours, but eventually the soft agar would turn milky white, indicating successful growth of the strains to a density which would highlight any plaques of inhibited growth forming (Figure 3.8). Three control plates were always set up, which included: agar and bacterial cells only (positive control); agar and filtrate only (negative control) and agar only (additional negative control).

Seawater filtrate sampled from station L4 (3 m, 10 m, 25 m and 50 m, Section 3.3.1) was the potential lysate initially used in plaque assays with all *Methylophaga* strains, which all yielded no plaque formation. Surface seawater (2.0 L) from station L4 was also concentrated to 20 mL using tangential flow filtration and additionally, 0.2 μ m filtered pore water from station L4 sediment, which was indicated to have 7 \times more total VLP abundance (Figure 3.5), were both also used as a potential lysate, however these also did not yield any plaques for all tested *Methylophaga* samples. The phase of growth of inoculating bacterial strains was also altered from the standard mid-exponential growth phase to using early- and late-exponential inoculating cultures, to determine if this would produce a successful infection. In addition, the incubation temperature used once the assay had been set up was also altered (17.5, 25 and 30°C). These different incubation variables were carried out using *M. marina*, resulting in 36 plaque assays (including dilutions of each variable) and 9 spot tests, but as with the previous assays, neither produced plaque clearings in the top agar layer.



Figure 3.8 Agar plates where cells have been added into the soft agar (left) and where no cells have been added (right), to indicate the difference in turbidity between plates.

3.5.3. Phage enrichment and spot testing

Since no plaque clearings were observed in any of the plaque assays using filtrate from station L4, phage enrichments were used in conjunction with spot testing to attempt to propagate the potentially small lytic phage populations of interest which may not be identifiable using a plaque assay approach. Phage enrichments were carried out using two methods (Section 2.6.6), where the filtrate used in both methods was sampled from station L4: 3 m; 10 m; 25 m; 50 m depths; and sediment pore water (Section 2.6.6). *Methylophaga* sp. strains (L41, 2, 4, 5, 6, 7 and 18) were tested against each of these filtrates (35 enrichments in total), with sub-samples taken from individual enrichment vials at different time points (Hours 0, 16 and 40) to be used in spot testing assays (105 spot tests in total). This was an initial experiment to produce filtrate to be used in spot test assays with *Methylophaga* strains. Filtrate (0.22 μm) derived from all enrichments was used for spot testing (Section 2.6.5), a quicker method when compared to a plaque assay. Spot testing did yield limited success, with eight clearings against *Methylophaga marina* and *Methylophaga* AH1 bacteria using different filtrate (Figure 3.9 and Table 3-2). However, clearings were also present in the negative control tests; therefore, the spot tests were repeated. This repeat did not yield any clearings with the same

filtrate/bacteria combinations or in a plaque assay for each combination, and a 'pickette' sample taken from the clearings did not indicate any increase in VLP numbers when compared with a negative control taken from a region where no clearing occurred. These clearings were thought to be false positives.

Sediment pore water (collected from L4) was used for phage enrichment (using the Twest and Kropinski (2009) method described in Section 2.6.6), using *M. marina* and *Methylophaga* L4 strains 1 and 6, with sub samples taken for VLP abundance and optical density. Both *Methylophaga* strain isolates did not have any increases in VLP abundance at any point of the enrichment and optical density of these phage enrichment cultures was also not affected when both were compared to a positive control. The phage enrichment involving *M. marina* produced a significant peak in total VLP abundance after 17 hours (20×10^8 VLP mL⁻¹, Figure 3.10), with the VLP 1 sub fraction making up the greatest proportion (89%) of this abundance (VLP 2 – 11%, VLP 3 – 0.01%). Additionally, the optical density of the positive control enrichment reached 0.8 after 20 hours; however, the optical density of the phage enrichment was lower at the same time-point (0.4).

Table 3-2 Different conditions and locations of filtrate that were obtained, which resulted in regions of no growth (Figure 3.9).

Filtrate			Bacterial Inoculum (mid-exponential)
<i>Location</i>	<i>Depth (m)</i>	<i>Date</i>	
Station L4	3	19/10/16	<i>M. marina</i>
Station L4	3	12/12/17	<i>Methylophaga</i> AH1
Station L4	3	12/12/17	<i>M. marina</i>
Station L4	10	12/12/17	<i>M. marina</i>
Station L4	25	12/12/17	<i>M. marina</i>
Filtrate (<i>derived from a previous enrichment</i>)			Bacterial Inoculum (mid-exponential)
<i>Bacterial Culture</i>	<i>Sampling time point</i>		
<i>Methylophaga</i> AH1	17 hours		<i>Methylophaga</i> AH1
<i>M. marina</i>	17 hours		<i>M. marina</i>
<i>M. marina</i>	20 hours		<i>M. marina</i>

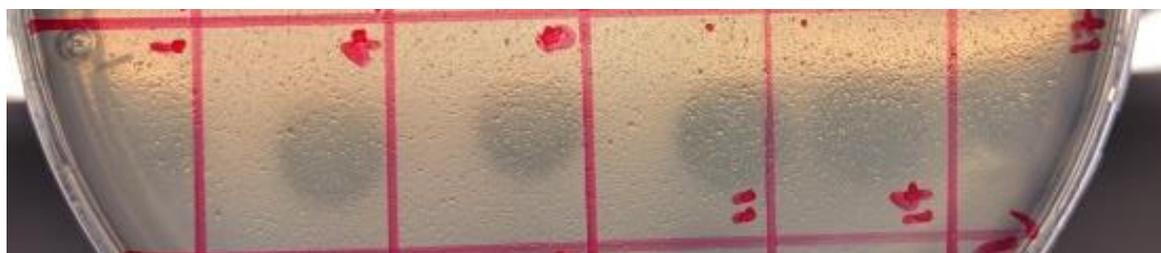


Figure 3.9 Sectioned plate from a spot test, indicating regions of no growth within different sections, where filtrate has been spotted on to the top agar.

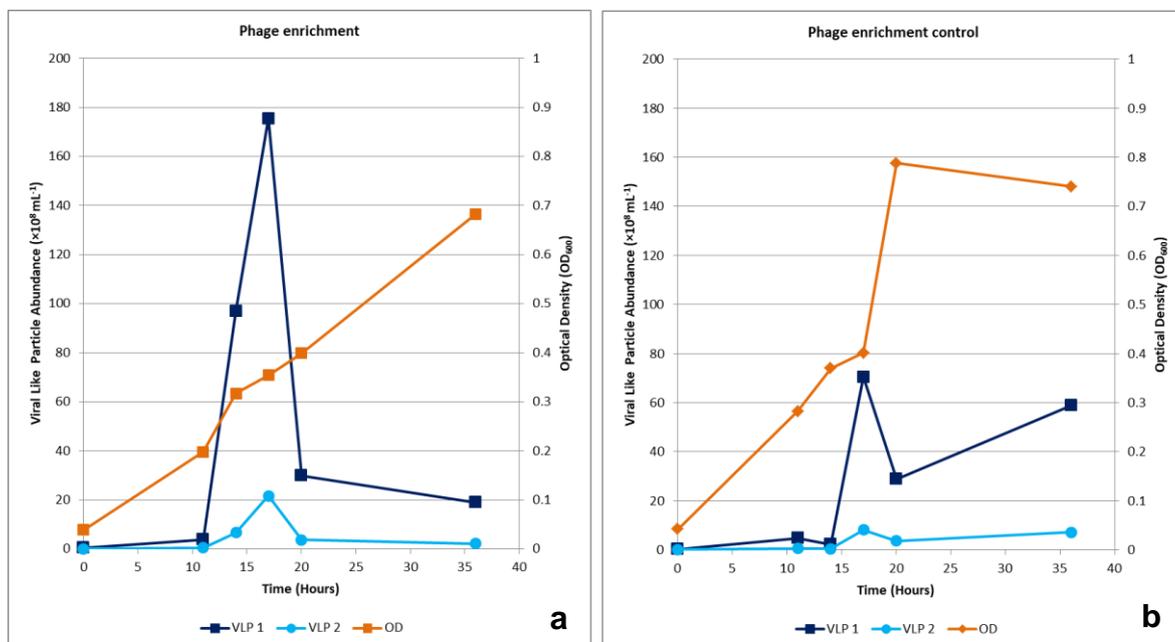


Figure 3.10 Abundance of VLP 1 and 2, and the optical density (OD) in a phage enrichment (a) and a positive control enrichment (b) using *M. marina*. 50 μL of sediment pore water was sampled from Station L4.

However, the positive control enrichment also indicated a peak at this same time point (T17, $7.8 \times 10^8 \text{ VLP mL}^{-1}$). Thus, it was important to repeat this experiment with the same filtrate and additionally use the filtrate from hour 17 to see if the results could be replicated (filtrate samples taken as described in Section 2.6.6). In the repeated experiment (Experiment 2) bacterial abundance was determined using flow cytometry, and the sampling frequency was increased to every hour between time point 11 and 18, to gain greater clarity of how the abundances were changing.

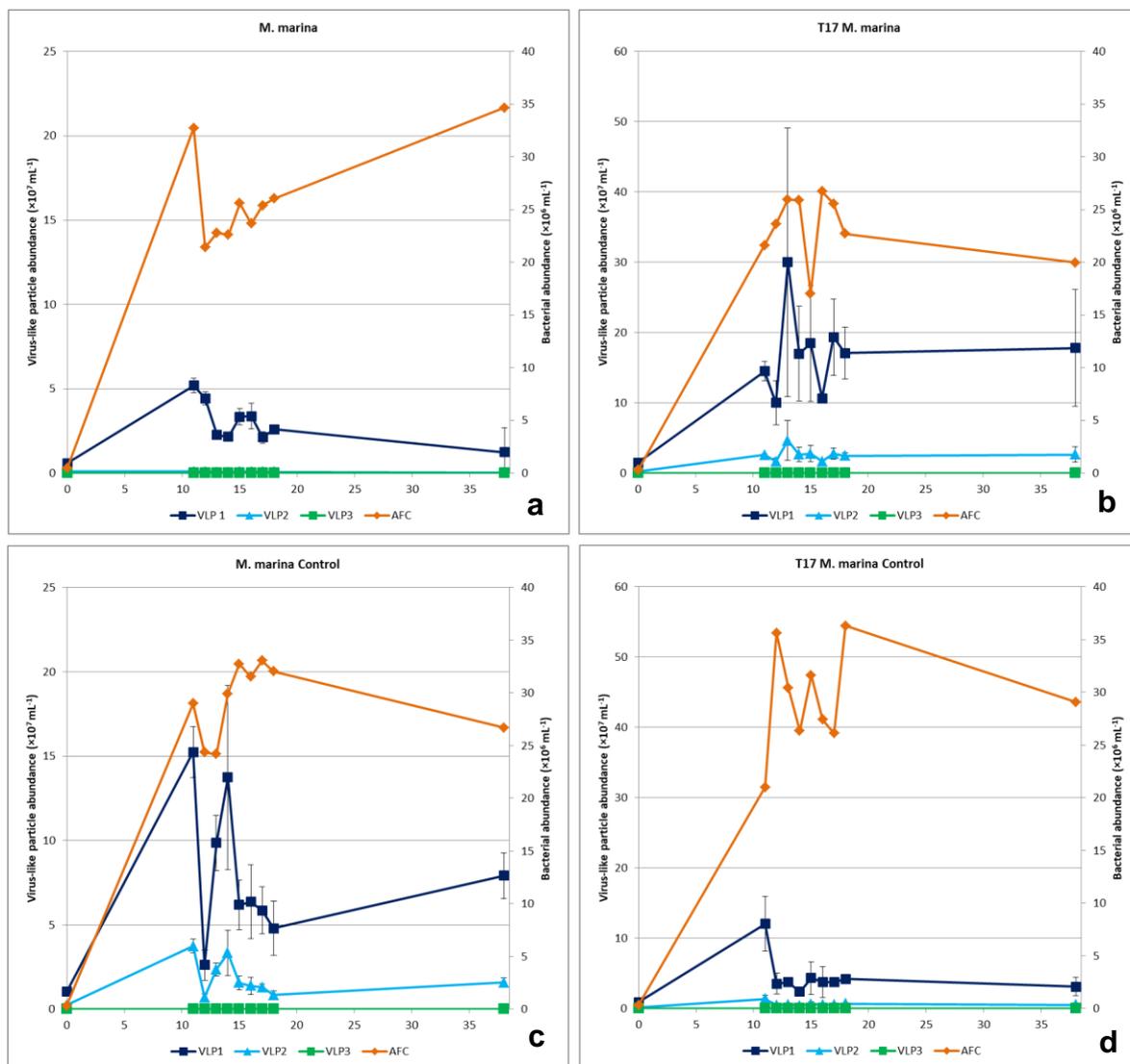


Figure 3.11 Abundance of VLP 1, 2 and 3, and bacteria in a *M. marina* phage enrichment using the same sediment filtrate as for Figure 3.10 and filtrate from hour 17 in Figure 3.10 (a and b respectively) and a corresponding positive control for each enrichment (c and d, respectively).

In phage enrichment experiment 2 (Figure 3.11a) the results indicated that the VLP 1 peak occurred at hour 11 ($5.2 \times 10^7 \text{ VLP mL}^{-1}$) and peaks again at hour 15, but then dropped off in abundance, compared to VLP 2 which dropped from an initial abundance of $1 \times 10^6 \text{ VLP mL}^{-1}$ to $0.3 \times 10^6 \text{ VLP mL}^{-1}$, additionally the VLP 3 concentration did not change. The VLP abundances did not increase to the same amount as before, or have a single large peak, and the VLP 1 and 2 concentrations at all time points were less than the positive control (Figure 3.11c). The phage enrichment using filtrate from hour 17 (Figure 3.11b and d) had a higher VLP

abundance than the positive control enrichment, but peaked earlier at hour 13 (3×10^8 VLP mL⁻¹). In all enrichments and controls, the VLP 1 virus subgroup was the most abundant group. In all enrichment attempts, after hour 11, the bacterial abundance appears to fluctuate (between $1.7 - 3.6 \times 10^7$ cells mL⁻¹). This variability in the bacterial abundance did not reflect the standard growth curve expected for *M. marina* (Figure 3.7), and was perceived to be a side effect of increasing the frequency of sub-sampling from a single vial. In previous cultures and enrichments this was not seen as an issue, but the increased number of sampling time points was potentially affecting the growth curves and/or any potential infection cycles. Therefore, the phage enrichment was repeated (Experiment 3), but with biological triplicates set up and 'sacrificed' at each time point, thereby removing the potential influence of sub sampling. The bacterial growth curve in the positive control enrichment (Figure 3.12a) resembled the previous baseline growth curves (Figure 3.7), with cells entering exponential growth phase after 12 hours and plateauing after 17 hours. Bacterial abundance (Figure 3.12a) peaked at 4.2×10^6 cells mL⁻¹ at hour 18; however cell numbers were an order of magnitude less than the positive control enrichment, which peaked at 3.6×10^7 cells mL⁻¹ after 17 hours. Additionally, the total VLP abundance (Figure 3.12b) within the phage enrichment was lower than in the control enrichment. Abundance of the different virus groups (Figure 3.12c and d) indicates that the VLP 1 sub group was the dominating fraction in experiment 3 throughout the 36 hours in both the enrichment and the control, with VLP 3 making up <1% of the total VLP abundance at all time points (Figure 3.12e and f). Results from the first 20 hours indicated that the total VLP abundance (Figure 3.12b) remained similar in each enrichment, and after T20 the total VLP abundance in the control enrichment began to increase (from 5.5 to 10.5×10^8 VLP mL⁻¹). However, due to the large standard deviation error bars associated with the final two time points, this increase may not be statistically significant.

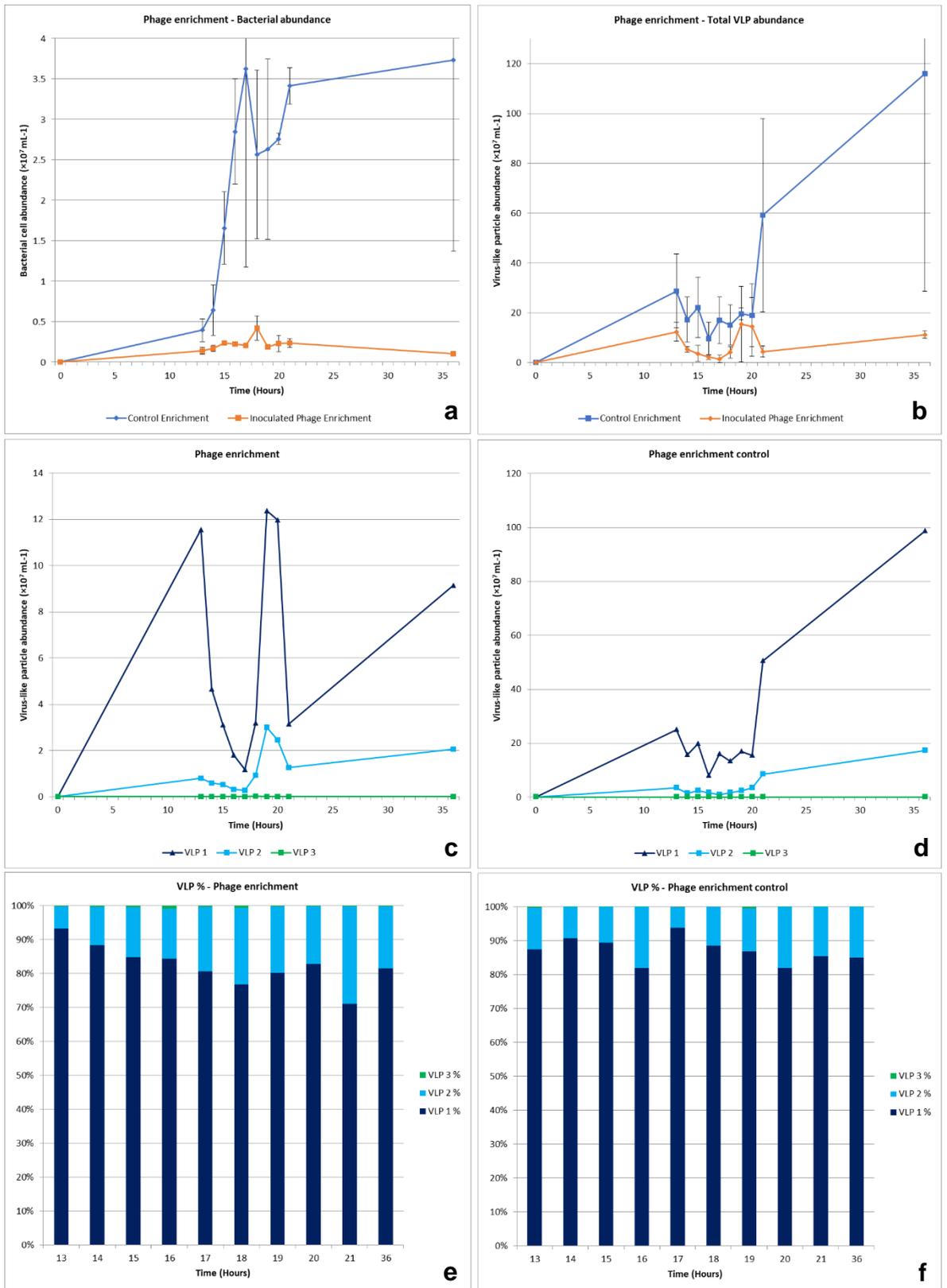


Figure 3.12 Results of phage enrichment experiment 3, over 36 hours, utilising a sacrificial approach for each sampling time point. Total virus-like particles (VLP) (a) bacterial cell number (b) positive control enrichment (d). The VLP abundance for

each of the individual VLP subgroups in the phage enrichment (c) and in the positive control culture (d). The percentages of the total VLP abundance for each of virus subgroup in the phage enrichment (e) and in the positive control (f).

3.6. Transmission Electron Microscopy (TEM)

Transmission Electron Microscopy (TEM) was used to examine particles to determine size and any visible morphological features of identifiable VLPs within filtrate from an environmental sample as a positive control (Figure 3.13a), and then filtrate derived from the *M. marina* phage enrichment experiments 2 (Figure 3.11), particularly focussing on filtrate from hours 13, 16 and 17 (Figure 3.13b-f). 30 mL filtrate from hour 16 (derived from an *M. marina* phage enrichment inoculated with 50 μ L filtrate from experiment 2) was concentrated using PEG precipitation (Section 2.6.9) before being prepared along with other samples for TEM (Section 2.8). Figure 3.13a indicates a *Myoviridae*-like virus particle in an environmental sample from Station L4, with a head and tail size of 100 nm and 305 nm, respectively. Filtrate from hour 13 (Figure 3.13b) indicated spherical VLPs, with diameters of 22 ± 2 nm, appearing to be nucleic acid dense within the capsid and have an outer membrane. Filtrate from hour 16 and 17 (Figure 3.13c and d) also indicate particles, but with larger particle sizes (40 ± 4 nm), which have stained externally, but do not appear to be nucleic acid dense internally when compared against the particles in Figure 3.13a and b. This appears to indicate different particle morphologies, with particles of 40 nm dominating in hours 16 and 17 of the enrichment, compared to the particles present at hour 13. Particles visualised in filtrate from hours 13 and 17 (Figure 3.13b, c and d) appear to be enveloped. Filtrate from hour 16, treated by PEG precipitation, shows particles concentrated together when compared against samples with the same level of magnification. The electron micrographs suggest that virus-like particles are present within filtrate, which has been derived from phage enrichments of *M. marina* and can be concentrated using PEG precipitation.

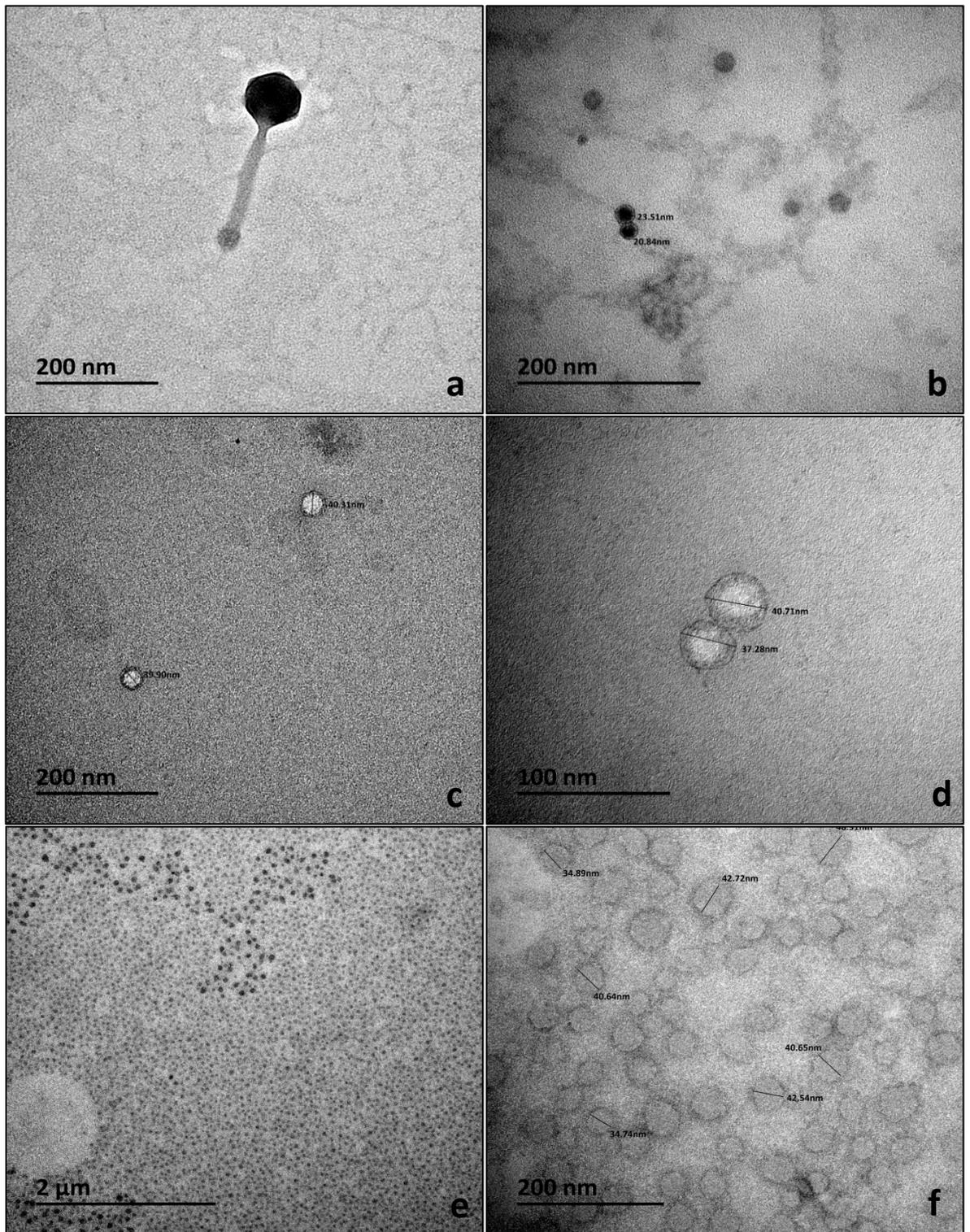


Figure 3.13 Electron micrograph of a Myoviridae-like virus particle imaged from an environmental sample collected from station L4 (a). Micrograph of virus-like particles (VLPs) visualised in filtrate derived from phage enrichment experiment 2 after 13 hours (b), 17 hours (c and d) and filtrate from hour 16 of an additional enrichment set up using filtrate from hour 13 in experiment 2 which was concentrated using PEG precipitation (e and f).

3.7. Purification and nucleic acid extraction of particles

As electron micrographs indicated that virus-like particles were associated with *M. marina*, then the nucleic acid of particles needed to be extracted for further genomic characterisation of the particles. Therefore particles within filtrate taken from *M. marina* were precipitated using polyethylene glycol (PEG) and then the precipitate was used in a caesium chloride (CsCl) gradient centrifugation in an effort to purify VLPs (Section 2.6.9). CsCl centrifugation did not yield a visible band of purified particles, therefore the PEG precipitate was used in a pulsed field gel electrophoresis (PFGE) gel. A modified approach was used as suggested by Sandaa, Short and Schroeder (2010), whereby the particles were embedded in a gel 'plug' and the plugs incubated in a lysis buffer overnight to release any nucleic acid present in particles. The plugs were then inserted into a pre-set agarose gel and ran for 20 hours on the PFGE; PFGE preparation steps are further described in Section 2.9.3. The resulting gel (Figure 3.14) was visualised with UV light and indicated that there were no visible bands in the lanes where the plugs had been inserted.

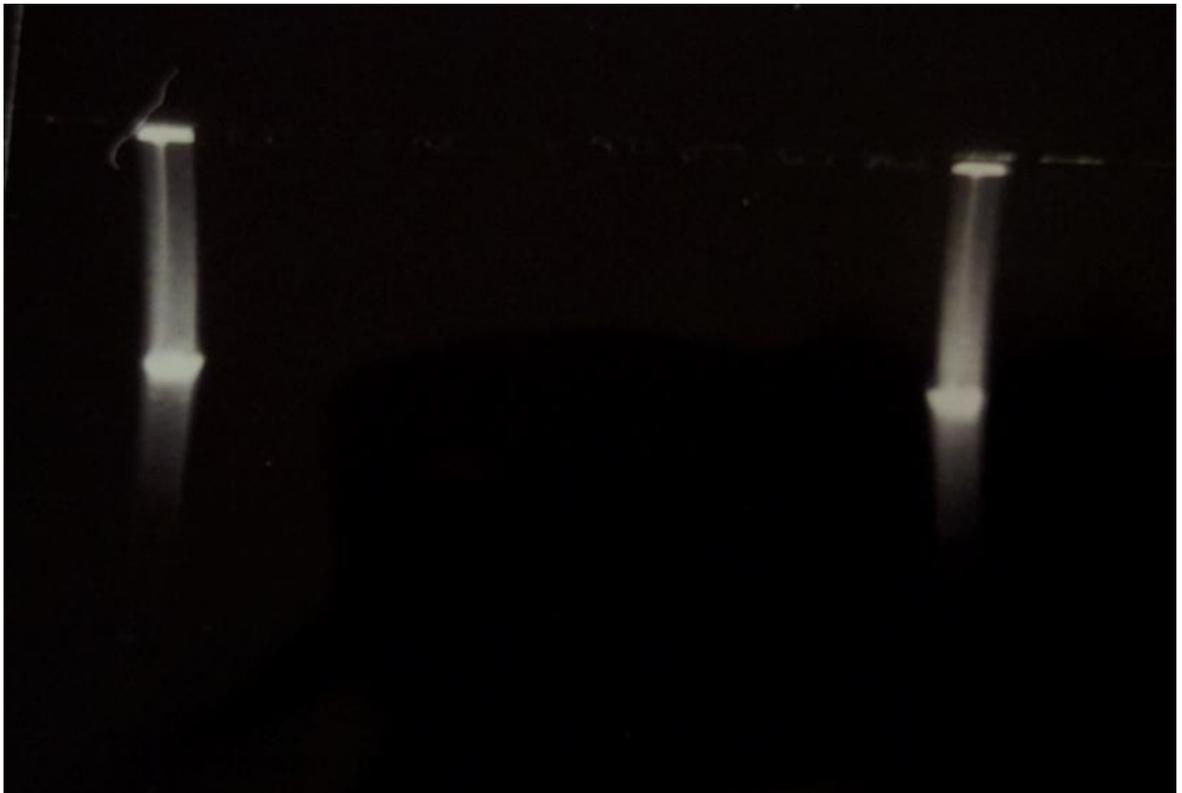


Figure 3.14 The PFGE agarose gel after 20 hours, visualised with UV after an ethidium bromide bath. Streaked bands indicate the PFGE ladders.

A DNA extraction was attempted on replicate gel plugs using a gel extraction method (Section 2.10.4). This yielded DNA concentrations between 2.3 – 7.2 ng/ μ L when quantified with a Nanodrop-1000 spectrophotometer (260/280 ratios: 1.22 – 3.53, and 260/230 ratios: 0.01 - 0.02). However, these DNA concentrations were driven by salt artefacts which were visible in the spectrograms, therefore this was not repeated.

3.8. Prophage Induction

Due to the apparent difficulty in establishing a lytic virus-host system for a methanol-utilising methylotroph, the potential of a lysogenic system was then explored. Lysogenic prophages are viruses which have inserted their genome into the genome of a host and can be induced into a lytic state by various stimuli. The genome of a recently characterised methanol-utilising methylotroph isolated from station L4, *Marinibacterium anthonyi* La6 (Howat *et al.*, 2019), was indicated to have an intact and a questionable prophage present within its genome (Figure 3.15) when it was

ran through the PHASTER (Phage Search Tool Enhanced Release) web tool (Arndt *et al.*, 2016). PHASTER identifies putative prophage regions by searching windows of 60 proteins for 6 or more proteins with phage-related keywords in the GenBank name field. If an integrase protein is identified, then further scanning of the region for potential phage attachment sites are identified. Putative prophage regions are then assigned completeness scores determined by the percentage of genes/proteins within a region, the regions size and the number of genes identified. Completeness scores are then used to put putative prophage regions into three categories: incomplete; questionable and intact (Zhou *et al.*, 2011). The two prophage regions were identified within scaffold 1 (3.15 Mbps) of the *M. anthonyi* La6 (6.79 Mbp) genome. Region 1 lacked an integrase sequence and was determined to be of questionable completeness with 19 identified proteins and was 16 Kbps in length, whereas region 2 was determined to be an intact prophage region with 22 identified proteins and was 27.4 Kbps in length. Region 2 consists of 15 annotated proteins and 11 hypothetical proteins (ranging between 42 - 140 amino acids each). The phage-like protein within region 2 most closely related with other phage terminase sequences (Figure 3.16). The two integrase sequences identified within region 2 group with integrase sequences of phage associated with the *Rhodobacteraceae* family (Figure 3.16). The intact prophage sequences indicated that this system may have an inducible lysogenic virus and is a potentially good system to explore and conduct prophage induction experiments using UV irradiation, antibiotic and thermal induction methods (Section 2.6.7) to stimulate virus production and release. Although a prophage sequence was not identified in the genome of *M. marina* by PHASTER, the web tool is only an indicator; therefore, this strain was also used in induction experiments.

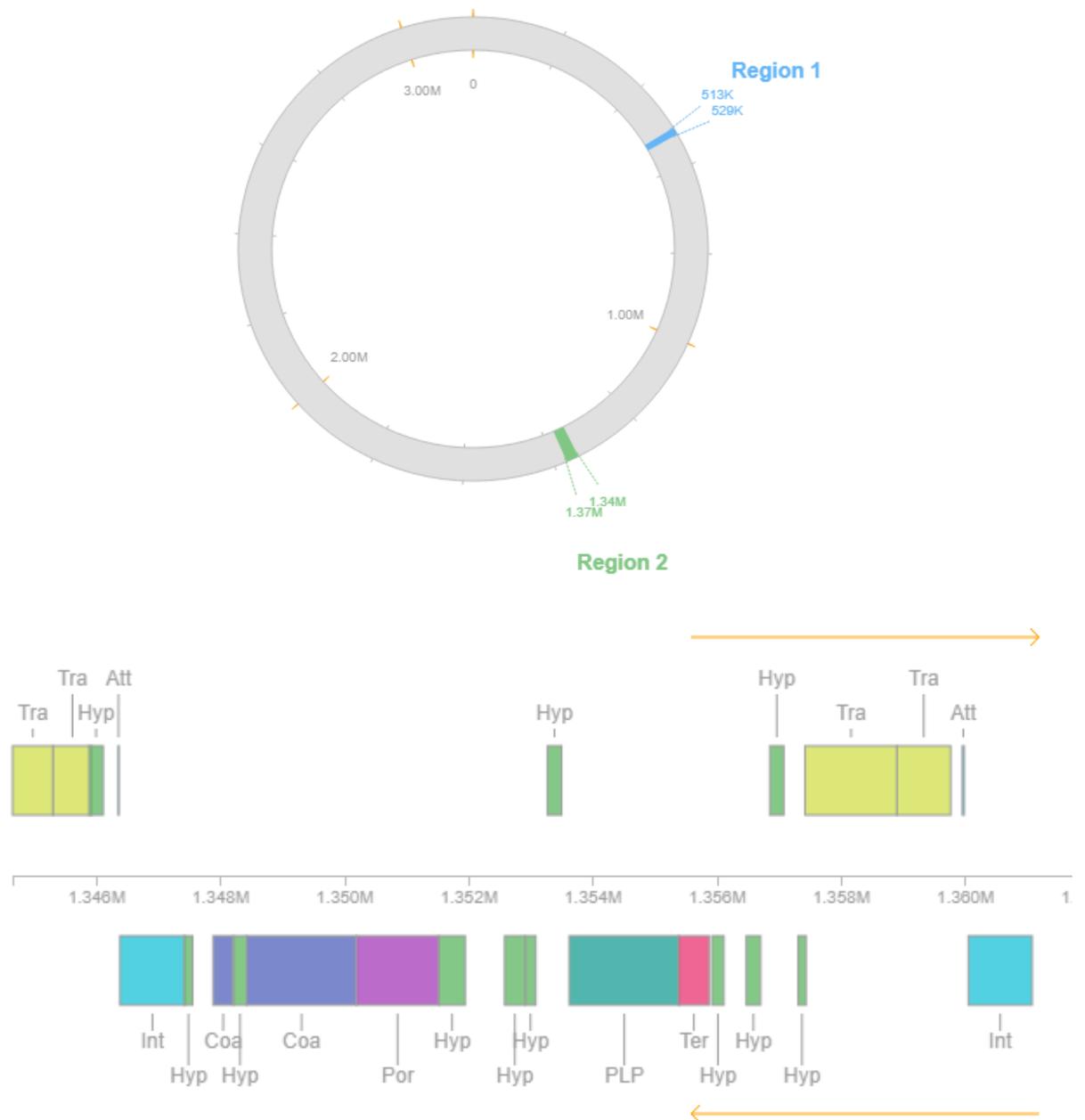


Figure 3.15 Annotation output of the intact prophage sequence, region 2, identified within the genome of *Marinibacterium anthonyi* La6 by PHASTER (Arndt *et al.*, 2016)(<http://phaster.ca/> - Tra; Transposase, Hyp; Hypothetical protein, Att; Attachment site, Int; Integrase, Coa; Head protein, Por, Portal protein, PLP; Phage-like protein and Ter; Terminase). Two attachment site sequences (11 and 29 bps) were removed from the figure but were located on the forward strand at 1.372 Mbps.

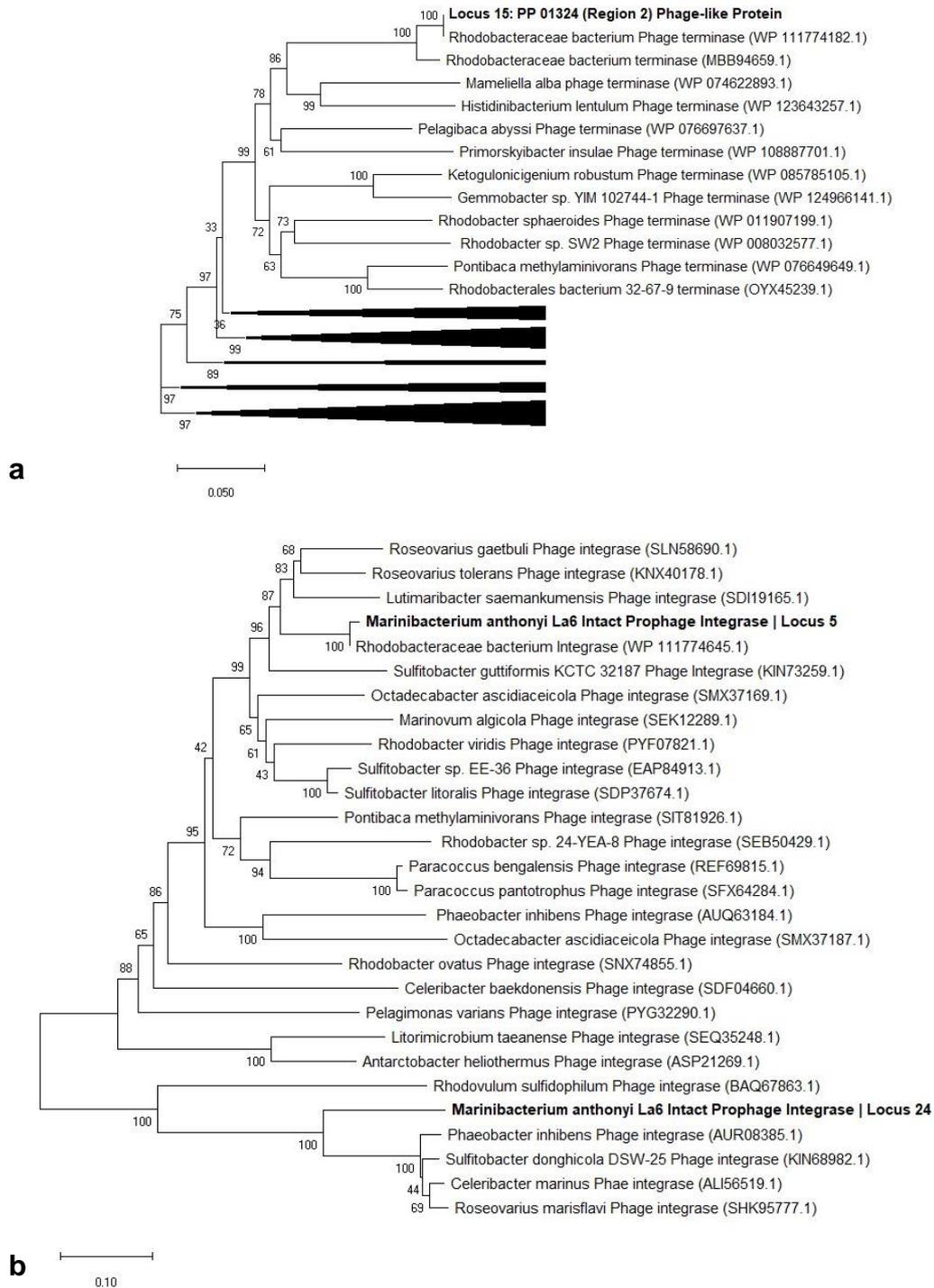


Figure 3.16 Phylogenetic analysis of a) the phage-like protein sequences and b) integrase identified within the *M. anthonyi* La6 genome (in bold) was inferred using an unrooted neighbour-joining phylogenetic tree. The bootstrap consensus tree was inferred from 2,000 resamples, and the numbers at branches indicate the percentage of resamples. The scale bar is indicative of the nucleotide substitutions per position. Sequence homologs of other phage integrase genes were included for comparison after using the NCBI BLAST algorithm. Sequences were aligned using the MUSCLE algorithm before the tree was produced using MEGA X (Kumar *et al.*, 2018).

3.8.1. UV Irradiation

UV irradiation can be used to induce a lytic response in lysogenic viruses, which have integrated into the genome of the host bacterium (Weinbauer and Suttle, 1999; Paul, 2008). Petri dishes containing *M. marina* and *Marinibacterium anthonyi* La6 cells in top agar were prepared (Section 2.6.5) (3 agar plates per strain) and then irradiated at 1500, 1000 and 500 × 100 μJ cm⁻² (Section 2.6.7). Plates were then checked periodically over the following 48 hours for any visible clearings until the top layer was completely opaque, however no clearings were visible in any plates.

3.8.2. Mitomycin C

Another method of inducing a prophage is the addition of mitomycin C (Section 2.6.7). This method was used with *Marinibacterium anthonyi* La6 (*M. marina* was not used due to culturing issues), and the bacteria and total VLP abundance was determined at four different time points after the mitomycin was removed (Section 2.6.7) (Figure 3.17). Changes in bacterial abundance indicated that both the mitomycin C-treated and negative control cultures reached stationary growth phase 2.5 hours after the washing step, (used to remove mitomycin from the treated culture). Total VLP abundance in the treated culture increased by 17% from 1.2 – 1.5 × 10⁶ VLPs mL⁻¹, however this abundance does not increase above that which was observed in the negative control culture. Thus, this method did not appear to induce viruses in this system under these conditions.

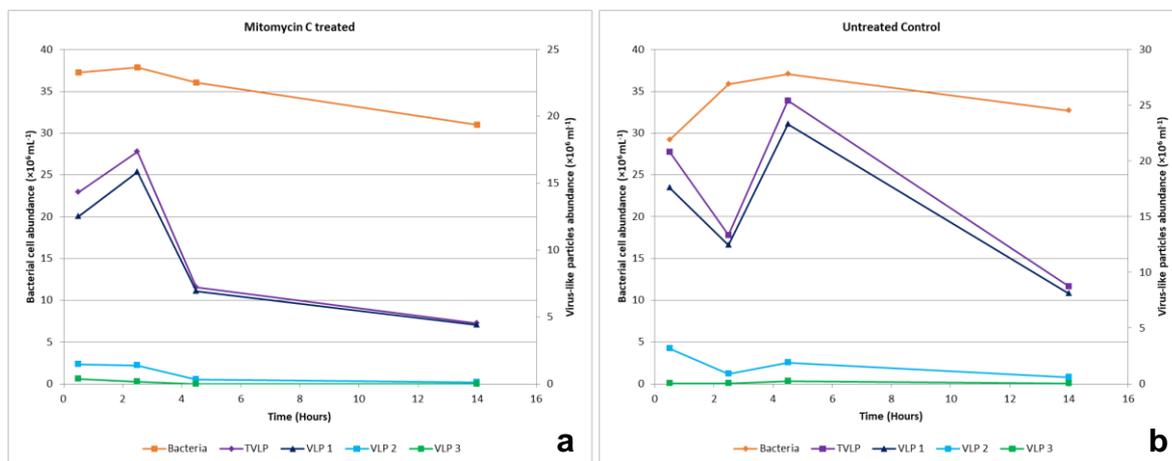


Figure 3.17 Abundance of total VLP, VLP 1, 2 and 3, and the bacterial cell abundance in a culture treated with mitomycin c ($30\mu\text{L}$, 1 mg mL^{-1}) and a negative control culture (a and b, respectively) using *M. anthonyi* La6.

3.8.3. Heat Shock

As a prophage could not be induced in *M. anthonyi* La6 either by antibiotic or UV induction, an alternative approach was to induce using heat shock of bacterial cells, this is another method which has been used to induce prophages and was attempted with *M. marina* and *M. anthonyi* La6, as described in Section 2.6.7. The results from an initial experiment indicated that when a 30 mL culture of *M. anthonyi* La6 (Figure 3.18a) was subjected to heat shock, there was no increase in any VLP fraction, or perceived significant decrease in bacterial cells as would occur due to cell lysis. No error data was available, as this was a single experiment without replicates.

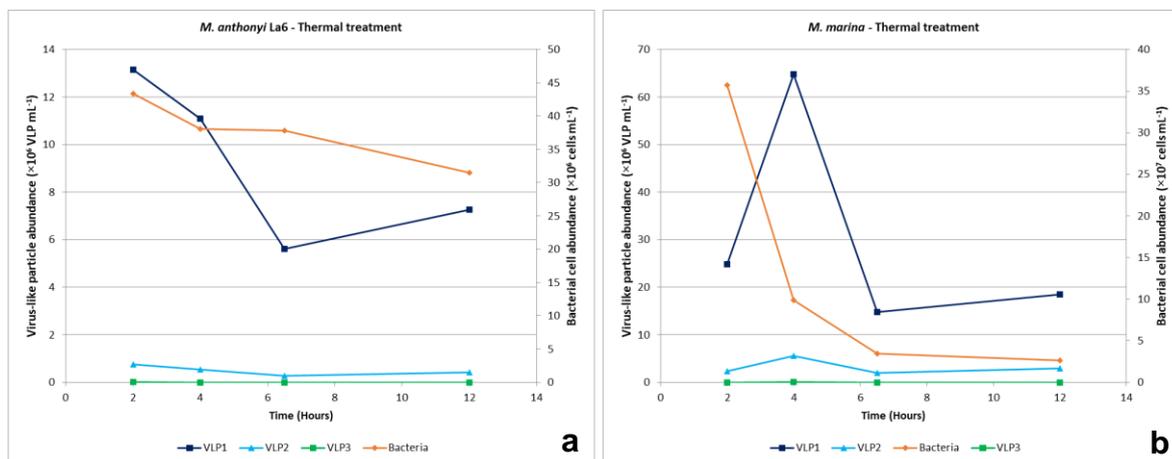


Figure 3.18 Initial bacterial cell abundance and VLP1, 2 and 3 abundance results from a heat shock experiment of a) *M. anthonyi* La6, and b) *M. marina* with samples taken over the 10 hour period after heat treatment.

However, *M. marina* indicated a 90% drop in bacterial abundance during the 5 hours following the 42°C heating step (from 3.6 to 0.3 × 10⁷ cells mL⁻¹) with a subsequent 61% increase in total VLP abundance. To determine if the results of Figure 3.18b were genuine, the experiment was repeated with a control and replicates, and the sampling period after treatment was extended to 50 hours (Figure 3.19).

Bacterial abundance after the 42°C heating step was lower than when compared against the negative control culture at all sampled time points, however this difference was not significant (n = 14, P = 0.73) (Figure 3.19a). The growth curve of the heated culture was similar to the negative control culture and indicated a plateau in cell abundance after 11.5 hours. Total VLP abundance in both cultures suggested that the overall VLP abundance in the negative control was higher than the heat-shock treated culture at all time points except 34.5 hours (Figure 3.19b). VLP 1 made up 99% of the total VLP abundance in the negative control culture, and in the treated culture this lowered to 97 – 99%, except time point 34.5 where VLP 2 increased to 5% of the total VLP abundance (Figure 3.20).

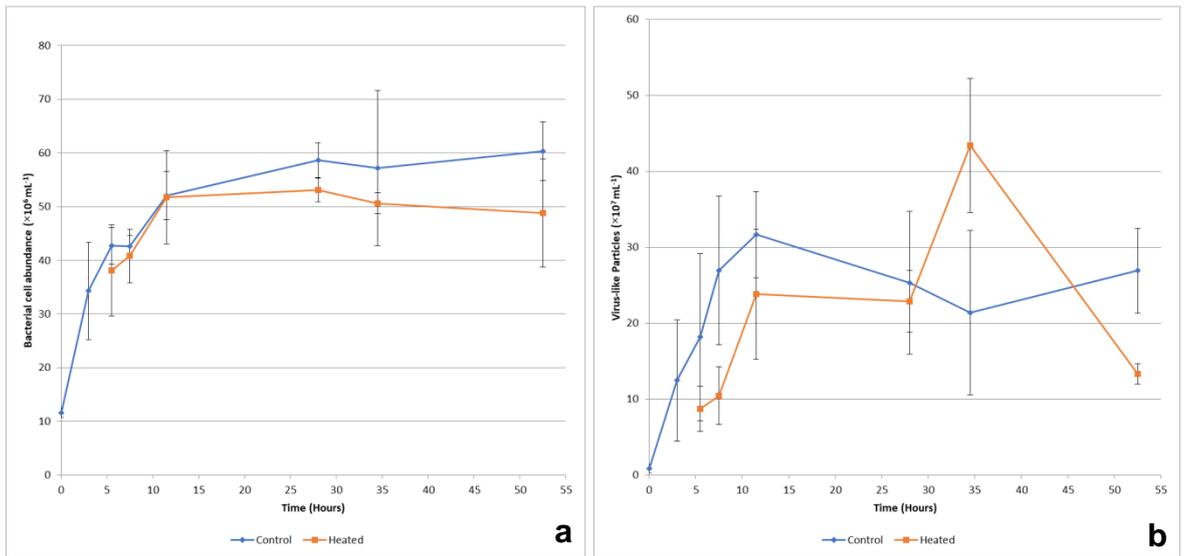


Figure 3.19 Total VLP and bacterial cell abundance in an a) *M. marina* culture shocked thermally at 42°C and b) a negative control culture.

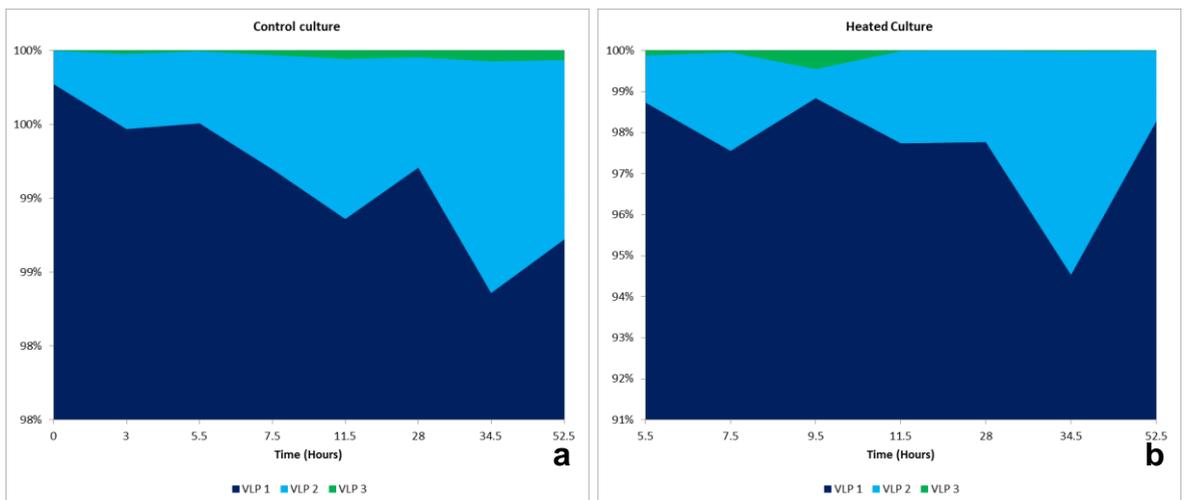


Figure 3.20 Percentage plots of virus groups VLP 1, 2 and 3 derived from samples taken from a) the *M. marina* culture thermal shock experiment, and b) the negative control culture.

3.9. Discussion

3.9.1. Isolation of methanol-utilising methylotrophs

Seawater samples taken from station L4 enriched with methanol (and lanthanum) allowed for the successful isolation of multiple *Methylophaga* species. Furthermore,

as isolate L4-1 had only 90% identity to the 16S *rRNA* of *M. marina* 222, this indicates that it is a potentially new species. This is interesting, as previous research by Howat (2017) stated that enrichments with added rare earth elements (REEs; such as lanthanum) increased the number of unknown sequences, however L4-1 was isolated from a methanol-only enrichment, indicating that novel species can still be derived without the addition of rare earth elements. Although, REEs could have been scavenged from the glassware as has been shown previously (Vu *et al.*, 2016). Regardless, the characterisation of novel methylotrophs was not a focus of this project, and further genomic investigation was not carried out on this isolate. Additionally, the presence of *xoxF4* in *Methylophaga* spp., is interesting as it has previously not been shown in *M. marina*, however *xoxF4* primers have indicated cross-specificity with *xoxF5* genes (Taubert *et al.*, 2015).

3.9.2. VLPs and isolation of methylotrophic phage

As *Methylophaga* spp. and *M. marina* could be enriched and isolated from station L4, this indicated that methanol-utilisers are extant within the Western Channel Observatory and made it a suitable location to sample for viruses which infect these hosts. This media-dependent approach is known to be highly selective for *Methylophaga* (Murrell, *pers. comms.* and Howat, 2017), but has also been used successfully to isolate other taxonomic groups (Howat *et al.*, 2019). However, *Methylophaga* is not a dominant genera throughout the year and therefore not as abundant as other taxa at station L4 (Sargeant *et al.*, 2016), therefore host abundance is likely a very important factor for the abundance of virulent phage (Murray and Jackson, 1992a).

The comparison of VLP abundance from different sites and depths indicates that sediment and harbour samples have significantly more VLPs than in the open water column, which in the case of sediment samples has been shown before (Drake *et al.*, 1998; Rastelli *et al.*, 2016). Three subpopulations of viruses were discriminated using flow cytometry, and the seasonal dynamics between these subgroups is discussed in more detail in Chapter 4. As with previous research, this study indicated VLP 1 as the more abundant population in environmental water samples

(Chen *et al.*, 2001; Mojica, Evans and Brussaard, 2014; Evans *et al.*, 2017). However, this was not the case with sediment pore water, where VLP 2 was the dominant subpopulation, indicating that pore-water samples are a valid milieu for marine virus isolation studies, likely increasing the diversity of VLPs being investigated (Drake *et al.*, 1998). Especially in regards to methylotrophy, where the benthic environment is an active environment for one-carbon metabolism (Antony *et al.*, 2010; Kalyuzhnaya *et al.*, 2012). The VLP 1 subpopulation is often associated with bacteriophage and was dominant in environmental samples. This is important when combined with the results of the phage enrichment experiments, where VLP 1 made up $\geq 89\%$ of the total VLP abundance (Marie *et al.*, 1999; Goddard *et al.*, 2005; Payet and Suttle, 2008b), as it provides further evidence for a potential phage for the isolated L4 methylotrophs. Although evidence does exist that lytic phage, such as myoviruses, can be derived from the VLP 2 subpopulation (Zhong *et al.*, 2014), VLP 2 is typically associated with pico-cyanobacteria and/or small eukaryotes (Payet and Suttle, 2008b; Personnic *et al.*, 2009). This would suggest that a significant portion of the VLPs enumerated in samples from the phage enrichments were viruses associated with bacteria. Despite this, as plaque assays did not indicate a lytic infection; this suggests that these particles may not be lytic phage of *M. marina*, although the absence of plaque formations and other lysis indicators does not necessarily mean that an infection has not occurred, and the assay could have been impacted by extended latent periods, reduced burst sizes or external factors, such as temperature, pH or light conditions (Abedon and Yin, 2009). The time of the year may also have played a role, due to a likely variability in the viral assemblage at L4 throughout the year (Rodriguez *et al.*, 2000), which is discussed further in Chapter 4. Phage enrichments did illustrate increases in virus-like particle abundance over time when combined with the obligate methylotroph *M. marina*, specifically the bacteriophage-associated VLP1 subgroup made up the largest percentage of virus-like particles and was more abundant in phage-enriched samples than in the environment (Figure 3.6). However, results were inconclusive with control cultures reaching similar VLP abundances as the VLP-enriched culture in some cases (Figure 3.11a), and these findings, combined with a lack of understanding of lytic viruses which infect methylotrophs, meant that a robust virus-

host system could not be established for *M. marina* in this study. This unsuccessful attempt to isolate a lytic virus for this host may have been a result of the low abundance of individual virus strains in natural communities, where even the most abundant account for <1% of the viral assemblage (Wommack *et al.*, 2009), and perhaps could have been overcome by increasing the volume of concentrated seawater from 2.0 L (used in this chapter) to 20 L as suggested by Wommack *et al.* (2009). Regardless of these issues, electron micrographs were successfully taken of VLPs within filtrate (Figure 3.13), supporting the theory of a potential phage system of *M. marina*, despite the unsuccessful attempts to actually isolate a phage. These VLPs appear to be either positively (Figure 3.13 a, b and e) or negatively stained (Figure 3.13 c, d and f), which has been used as an indicator of presence or absence of nucleic acid - however this could also be a result of the unpredictability of the uranyl acetate stain (Ackermann, 2009). The isolated virus-like particles in micrographs appear to be most similar in morphology to either the *Cystoviridae* or *Plasmaviridae* families of phage (Ghabrial *et al.*, 2018; Cotmore *et al.*, 2019), primarily due to the apparent lipid envelope visible around the capsids (Figure 3.13b, c and d).

3.9.3. A potential chronic infection system

M. marina was absent of any intact prophage sequences detectable by PHASTER or VirSorter (Ben Temperton, pers. comm.), and did not react to lysogenic induction experiments. When combined with the absence of an observable lytic response and the assumption that viruses will affect this species to some degree, another mode of infection may be involved. This could be very likely as 85% of available phage genomes which have been explored, are derived from phage that only infect three of the 45 identified bacterial phyla (Holmfeldt *et al.*, 2013), indicating a potentially vast reservoir of unknowns.

Lytic viruses can be prevented from infecting a bacterial cell, due to phage-dependent cell-surface modifications typically associated with temperate phage (prophage), and is known as superinfection immunity (Casjens and Hendrix, 2015; Obeng, Pratama and Elsas, 2016). Although generally attributed to prophage

infections, this has also been demonstrated with carrier state chronic infections, where bacteria and phage exist in equilibrium (Siringan *et al.*, 2014). The results of the phage enrichment experiments suggest that there are VLPs in high abundance in both the enrichment and positive controls, and that the VLP 1 subpopulation is significant in all analysed samples. VLP abundance could be increasing in control cultures as VLPs are not always virus particles, hence “virus-like particles”, which can include biomolecules produced by the host cell such as proteins or exosomes – although typically these are larger than reported here (50-300 nm; Zamora and Aguilar, 2018). In experiment 1 (Figure 3.10), the total VLP abundance is higher in the enrichment compared to the control, which would be expected where an infection is taking place and viruses are being released during cell lysis, which would occur with a lytic infection (Hobbs and Abedon, 2016). However, in experiment 2 (Figure 3.11), the sampling frequency is increased during the period of expected VLP production and the results are inconclusive, as VLP abundance is higher in the positive control culture. This suggests that VLP production occurs regardless of a filtrate being added to a bacterial culture, as would occur in a chronic infection where virions are either extruded (Russel and Model, 2006) or budded (Putzrath and Maniloff, 1977). Based on the images in Figure 3.13b, c and d, the budded chronic life cycle is more likely, due to the visible enveloped membrane around particles. Additionally, experiment 3 indicated that when a phage enrichment is set up to be sampled sacrificially (to avoid sampling from the same vial), the phage enrichment has a much lower growth rate (Figure 3.12a), which is similar to the findings of Putzrath and Maniloff (1977), that newly infected cells with a chronic phage grow slower. Unfortunately, chronic infections are very poorly understood in the marine environment, due to the detection issues associated with traditional approaches and assays (Weinbauer, 2004), which are prominent methods in virus research. Typically, chronic infections and viral release by budding is more associated with animal and human cell infections (Garoff, Hewson and Opstelten, 1998; Sundquist and Krausslich, 2012). However, budding has been shown to occur in unicellular organisms existing in the marine environment (Mackinder *et al.*, 2009).

3.10. Conclusion

Enveloped virus-like particles were visualised using electron microscopy, and the results suggested that the life cycle and virus-host interaction of this phage may be distinct from typically reported lytic and lysogenic host-virus interactions, however much more data is required to corroborate these findings. Although no specific virus isolations were successfully carried out, certain lines of evidence have set a foundation for future work. Ideally, phage enrichment experiments could have been repeated were more time available, with a focus placed on increasing the abundance of virus-like particles used to inoculate enrichments (>20 L), which would be an inexpensive and a relatively simple additional step. This assumes that a lytic virus does exist in the environment which infects this bacterial isolate and was simply not detected by our efforts. This would follow the original principle of increasing the probability of a particle infecting a bacterial cell by increasing the total concentration of virus-like particles in the filtrate being added to the bacterial culture. If a lytic infecting particle can be isolated then this would allow more material for viral genome characterisation, which would give greater insight into the virus structure and infection cycle. Alternatively, single virus genome sequencing could be an option; however, the whole genome amplification step would be reliant on the genome type, which could be a DNA genome (such as *Parmaviridae*) or an RNA genome (such as *Cystoviridae*). This would only be determined from sequencing or can be basically indicated by degrading the nucleic acid using DNase/RNase. Primers have been used successfully to target and further investigate samples for presence/absence of viruses; however, a universal primer set, such as *16S rRNA*, does not exist for viruses. Therefore, methods which do not utilise marker genes could be more suitable, such as: thin-section TEM (to further investigate the relationship between particles and host cells) and/or cryo-EM (for greater characterisation of virus-like particles).

Chapter 4

Seasonal variation in microbial utilisation of methanol in the Western English Channel

4. Seasonal variation in microbial utilisation of methanol in the Western

English Channel

4.1. Introduction

The well-established western channel observatory (WCO) is located in the Western English Channel, and consists of station L4 (50° 15.00' N, 04° 13.02' W; depth 50 m) and station E1 (50° 02.00' N, 04° 22.00' W; depth 75 m), which are situated 13 and 40 km south of Plymouth, UK (Figure 2.1), respectively. L4 is a dynamic site representative of coastal waters with tidal influences, affected by periodic run-off from the nearby rivers Tamar and Plym (Smyth et al., 2015), and is comprehensively sampled on a weekly basis. Further south is station E1, which is a site more representative of open shelf waters (Smyth et al., 2010), and is sampled on a bi-weekly basis, however sampling is more weather dependent. Sampled parameters include *in situ* seawater temperature, salinity, fluorescence, nutrients (nitrate, nitrite, phosphate, silicate and ammonia), bacterial abundances and plankton abundances. Hourly measurements are recorded by moored buoys with an array of sensors at the surface, recording temperature, salinity, oxygen, turbidity and fluorescence and different wave/wind measurements (www.westernchannelobservatory.org.uk).

As part of this PhD, a one-year timeseries was conducted with the objectives of (i) characterising the seasonal variability in microbial methanol uptake rates throughout the water column, and (ii) determine whether any relationships exist with virus abundances, biogeochemical parameters and/or bacterial groups. Sampling was carried out between August 2017 and September 2018, with water samples used to determine the microbial methanol uptake (dissimilation and assimilation) and bacterial production rates. Additionally, samples were taken to investigate the virus and bacterial abundances and additionally DNA was extracted from 0.2 µm filters to investigate the taxonomic diversity of the *16S rRNA* (V4 Region) and *xoxF5* genes at surface and bottom depths. Samples from L4 were taken from different depths (2 m, 10 m, 25 m and 50 m), but assimilation was only determined at surface and 50 m depth. Similarly, at E1, water samples for uptake rates and virus abundances were determined at the surface and 60 m. The seasons in this chapter are defined

as autumn (September – November 2017), winter (December 2017 – February 2018), spring (March – May 2018) and summer (June – August 2018) as defined by Sargeant (2013).

4.2. Results

4.2.1. *Microbial methanol uptake*

Rates of microbial methanol dissimilation (oxidation to CO₂ for energy) at station L4 (Figure 4.1a) between all sampled depths ranged between 0.0 – 9.3 nmol L⁻¹ h⁻¹, and the highest dissimilation rates were measured during winter at 25 m and 50 m (9.1 – 9.3 nmol L⁻¹ h⁻¹). On average, dissimilation rates were higher during winter (6.5 ± 3.0 nmol L⁻¹ h⁻¹) and lowest in summer months (0.8 ± 1.0 nmol L⁻¹ h⁻¹). At the surface of station L4, rates averaged 2.2 ± 1.9 nmol L⁻¹ h⁻¹, the average methanol dissimilation rates then increased to 2.4 ± 2.0 nmol L⁻¹ h⁻¹ at 10 m and 2.7 ± 2.6 nmol L⁻¹ h⁻¹ at 25 m, before decreasing slightly at 50 m (2.2 ± 2.5 nmol L⁻¹ h⁻¹). Methanol assimilation (uptake in to biomass) rates at station L4 (Figure 4.1b) varied throughout the water column between 0.08 – 2.9 × 10⁻² nmol L⁻¹ h⁻¹ from August 2017 to September 2018. On average, methanol assimilation rates were higher during winter and autumn (2.2 ± 0.3 and 1.8 ± 0.6 × 10⁻² nmol L⁻¹ h⁻¹, respectively) and lower during spring and summer (0.7 ± 0.4 and 0.8 ± 0.3 × 10⁻² nmol L⁻¹ h⁻¹, respectively). The highest assimilation rates were determined during November, and by mid-spring positive assimilation rates could not be determined. Methanol assimilation rates at the surface and bottom of the L4 water column appeared to follow very similar annual cycles, and statistically correlated with each other ($r = 0.767$, $n = 16$, $P < 0.001$). However, as assimilation rates were only determined at the surface and 50 m, it cannot be affirmed that intermediate depths follow the same annual cycle.

Methanol dissimilation and assimilation rates at station L4 indicated that they shared similar seasonal cycles and were statistically correlated at the surface and bottom of the L4 water column ($r = 0.823$, $n = 18$, $P < 0.001$), and both uptake rates were

on average highest during winter months. However, the highest assimilation rates were found in late-autumn at the bottom of the water column (50 m), in addition the dissimilation rates were also determined to be higher, deeper in the water column (25 m and 50 m). Dissimilation rates were lowest during summer months and assimilation rates were lowest during spring, perhaps indicating a more staggered seasonal cycle when comparing the two. Using the methanol assimilation and dissimilation rates, the bacterial growth efficiency for methanol (BGE_M) was determined (Figure 4.1c), yielding an average of $0.8 \pm 0.7\%$ at the surface and $2.5 \pm 4.4\%$ at 50 m, throughout the time series. Seasonal averages indicated that BGE_M was highest during summer months ($2.8 \pm 4.7\%$) and lowest during spring ($0.5 \pm 0.7\%$). This is slightly higher than that reported by Sargeant (2013) for the same sampling location, but is still consistent with the reported by Dixon, Beale and Nightingale (2011) in coastal waters.

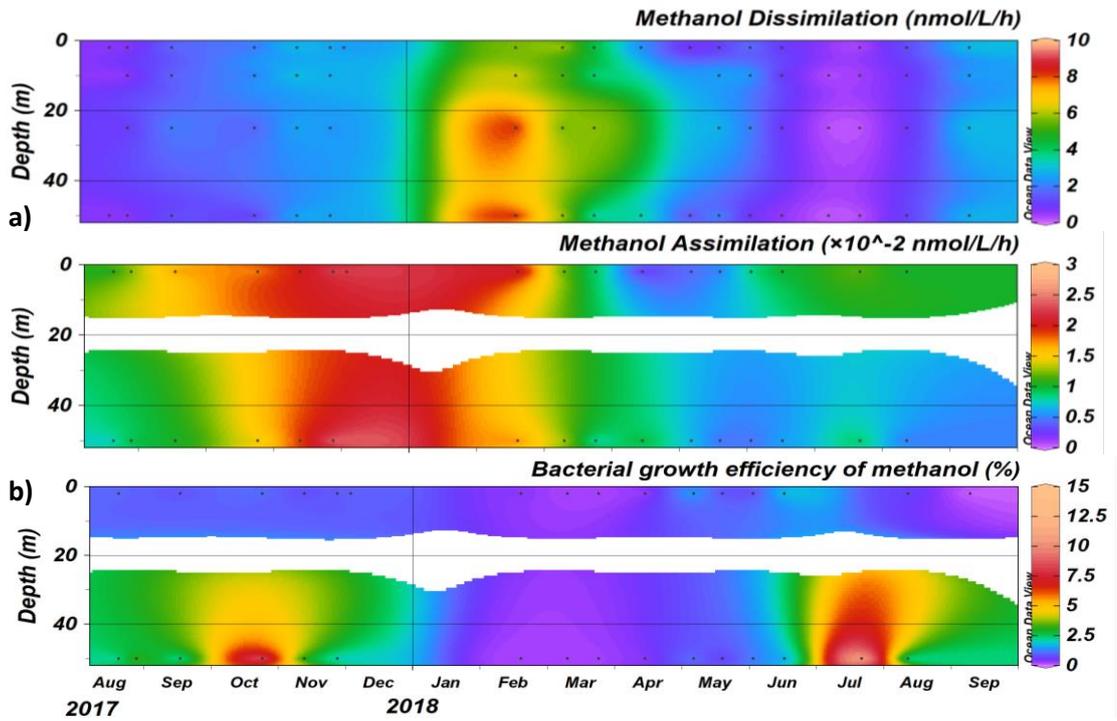


Figure 4.1 Temporal variation of (a) methanol dissimilation rates throughout the water column and (b) methanol assimilation rates and (c) bacterial growth efficiency for methanol (BGE_M) at station **L4** (August 2017 – September 2018).

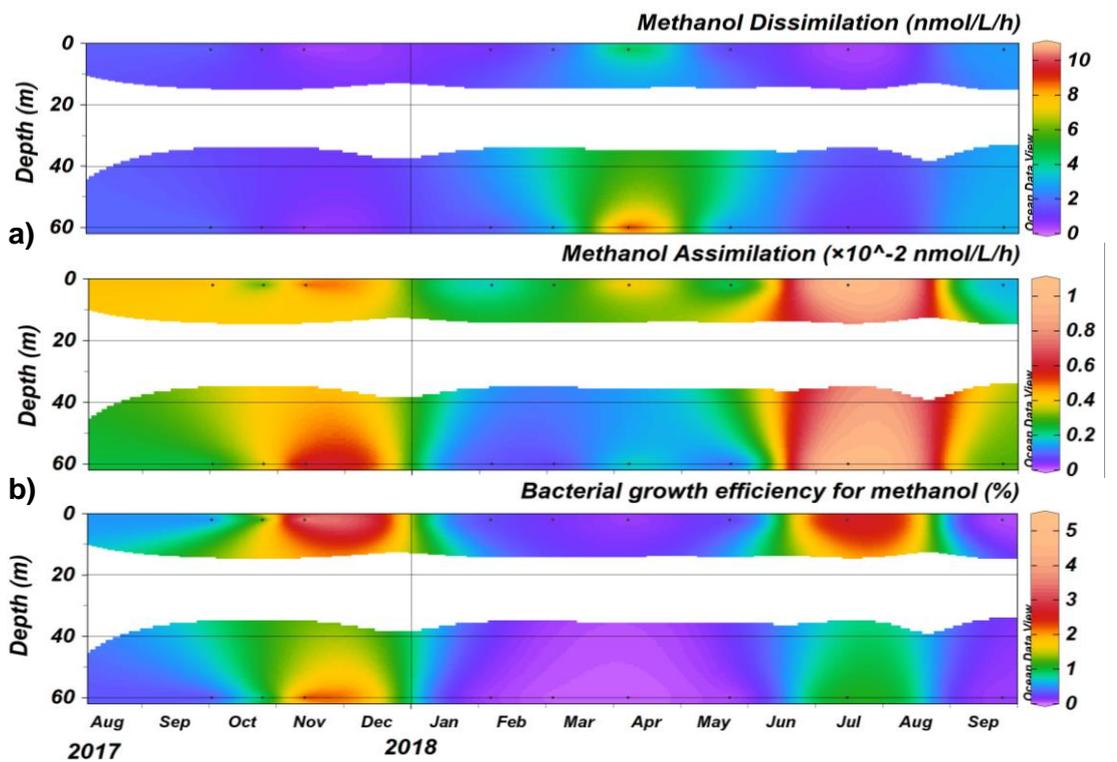


Figure 4.2 Temporal variation of (a) methanol dissimilation rates, (b) methanol assimilation rates and (c) bacterial growth efficiency for methanol (BGE_M) throughout the water column at station **E1** (August 2017 – September 2018).

Methanol dissimilation rates at station E1 ranged from 0.1 – 5.3 nmol L⁻¹ h⁻¹ at the surface and 0.2 – 10.6 nmol L⁻¹ h⁻¹ at 60 m between August 2017 and September 2018 (Figure 4.1a). On average, the highest dissimilation rates occurred in spring (3.8 ± 3.6 nmol L⁻¹ h⁻¹) and were lowest during summer months (0.6 ± 0.4 nmol L⁻¹ h⁻¹). Methanol assimilation rates at station E1 varied between 0.16 – 1.05×10^{-2} nmol L⁻¹ h⁻¹ at the surface and 0.04 – 1.1×10^{-2} nmol L⁻¹ h⁻¹ at 60 m (Figure 4.1b). On average methanol assimilation rates were highest during summer ($1.07 \pm 0.03 \times 10^{-2}$ nmol L⁻¹ h⁻¹) and lowest during winter months ($0.12 \pm 0.05 \times 10^{-2}$ nmol L⁻¹ h⁻¹). Bacterial growth efficiency (BGE_M) determined for E1 samples (Figure 4.1c), indicated averages of $0.54 \pm 1.0\%$ at the surface and $0.99 \pm 1.7\%$ at 60 m. The BGE_M was on average lowest in spring ($0.08 \pm 0.1\%$) and highest during summer ($1.9 \pm 1.1\%$). Total average methanol uptake rates were very similar between station L4 (2.1 ± 2.2 nmol L⁻¹ h⁻¹) and E1 (2.2 ± 2.4 nmol L⁻¹ h⁻¹) within the WEC throughout August 2017 to September 2018. The seasonal cycle of dissimilation rates at station E1 was slightly different to L4, as maximum rates were determined two months apart (February and April) as that at E1; however, both stations do exhibit seasonal minima during summer months.

The rates of methanol dissimilation across all depths at L4 and E1 have similar averages (2.4 ± 2.2 and 2.16 ± 2.4 nmol L⁻¹ h⁻¹, respectively), and both plots indicate similar annual peaks between February and April 2018. The methanol dissimilation rate range at station L4 (0.0 – 9.3 nmol L⁻¹ h⁻¹; Figure 4.1a) was within the same range as that previously reported by Sargeant *et al.* (2016; 0.7 - 11.2 nmol L⁻¹ h⁻¹). Station L4 appears to indicate higher uptake rates at deeper depths, and the difference between the surface and bottom sampled depths at E1 would indicate the same during early spring. However, averages at the surface and bottom of E1 would not indicate this (0.9 ± 1.7 and 0.5 ± 1.0 nmol L⁻¹ h⁻¹, respectively). Methanol assimilation rate averages at the surface and bottom of the water columns of both stations indicate that they are higher at L4, compared to E1 (1.2 ± 0.7 and 0.4 ± 0.3 nmol L⁻¹ h⁻¹, respectively). E1 values indicated opposing seasonal uptake cycles between methanol dissimilation and assimilation rates, although not significant ($r = -0.244$, $n = 9$, $P < 0.5$). There were similarly larger assimilation rates in 2017 at L4; however, the assimilation rates did not peak as high in 2018 when dissimilation rates

were lower but were increasing from April to August. BGE_M indicated very similar seasonal patterns at both stations, with peaks in autumn 2017 and mid-summer 2018 (Figure 4.1c and Figure 4.2c). These peaks appeared prevalent at the surface and bottom of the E1 water columns, whereas they were less established at the surface of L4.

4.2.2. *Virioplankton abundances*

Total virus-like particle (VLP) abundances at L4 ranged between $5.5 - 37.4 \times 10^6$ VLPs mL⁻¹ (Figure 4.3a), and the largest total VLP abundance was detected in surface waters in mid-August 2018, and was lowest in mid-February at 10 m depth. However, total VLP abundance was a combination of three virus subgroups (VLP 1, VLP 2 and VLP 3, Figure 2.3), which ranged in abundance between $1.2 - 30.3 \times 10^6$, $2.3 - 7.9 \times 10^6$ and $0.1 - 1.6 \times 10^6$ VLPs mL⁻¹, respectively.

Total VLP abundance was dominated throughout the sampling period by VLP 1 (16 – 92%; Figure 4.3b) and VLP 2 (7 – 73%; Figure 4.3c), whereas VLP 3 only made up between 1 – 14% (Figure 4.3d). VLP 1 abundance was lower during autumn ($1.2 - 6.1 \times 10^6$ VLPs mL⁻¹) and winter ($5.0 - 1.2 \times 10^6$ VLPs mL⁻¹), before increasing in May, with a prolonged peak during summer. In contrast, VLP 2 and VLP 3 had a higher abundance during autumn, and both peaked in November at 25 m (7.9×10^6 VLPs mL⁻¹ and 1.6×10^6 VLPs mL⁻¹, respectively). Furthermore, VLP 1 was inversely correlated against the VLP 2 and VLP 3 abundance values ($r = -0.577$, $n = 19$, $P < 0.01$). VLP 2 was the more dominant virus subgroup at all depths during autumn and winter (42 - 73% of total VLPs), compared to VLP 1 (16 – 49%), and then during spring there was more variability in dominance by VLP 1 (17 – 92%) and VLP 2 (7 – 73%). However, VLP 1 became more established during the summer months (36 – 90%). VLP 3 made up 5 – 14% in autumn and winter, then varied slightly more in spring (1 – 12%), but then only constituted 1 – 8% of the total VLP abundance in summer.

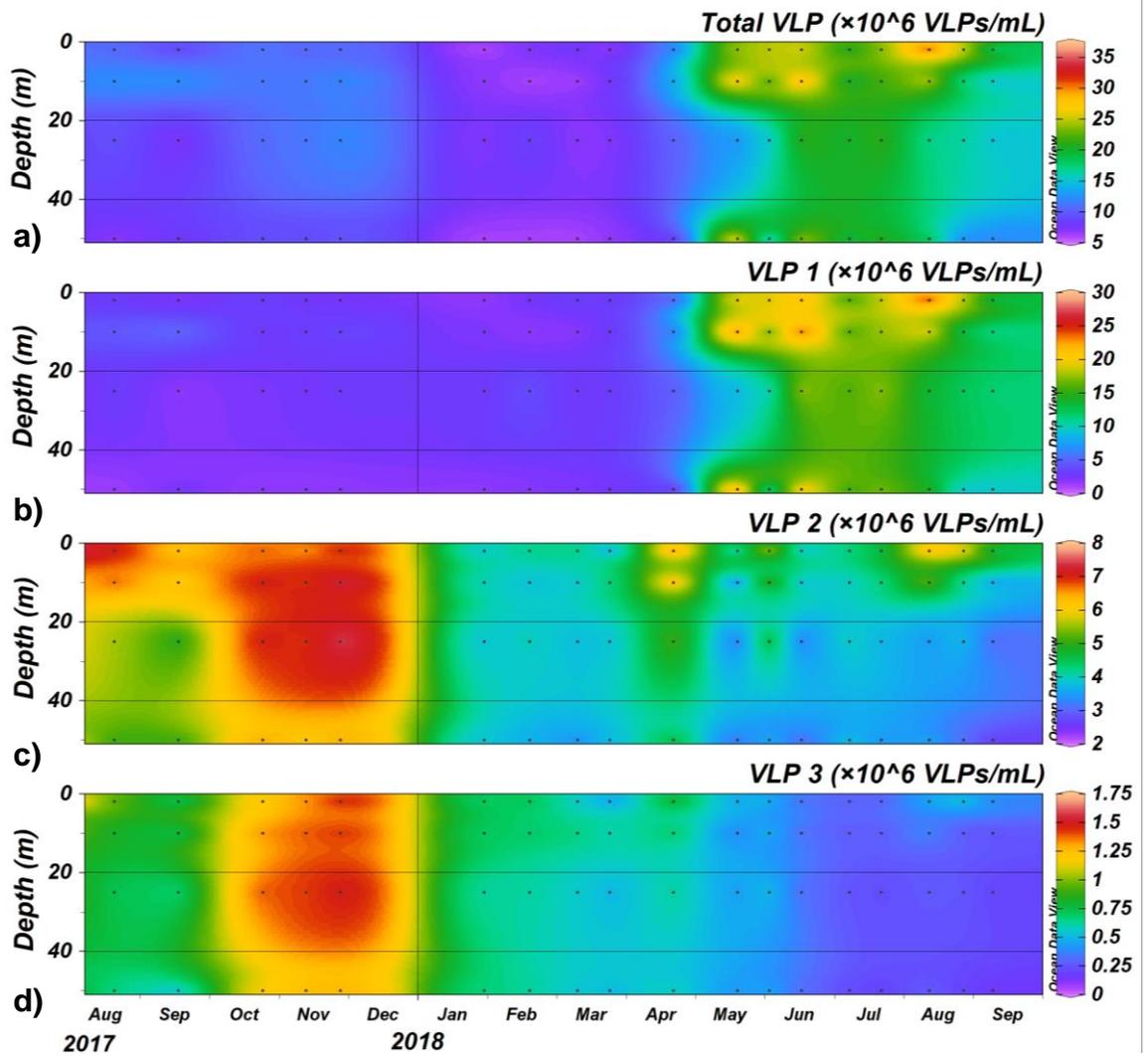


Figure 4.3 Abundances of (a) total virus-like particles (VLPs), and virus subgroups (b) VLP 1, (c) VLP 2 and (d) VLP 3 as VLPs mL⁻¹ through the water column at station L4 (August 2017 – September 2018).

The total VLP (Figure 4.4a) abundance at station E1 ranged between $4.3 - 12.7 \times 10^6$ VLPs mL⁻¹ at the surface, and between $3.5 - 9.1 \times 10^6$ VLPs mL⁻¹ at 60 m. The lowest and highest total VLP abundances were both detected during spring, where the lowest occurred at the start of spring, whereas the highest abundance happened at the end of spring. On average autumn and spring were similar in terms of total VLP abundance between the surface waters and 60m, ranging between 6.0 ± 2.2 (n = 12) and $6.3 \pm 3.4 \times 10^6$ VLPs mL⁻¹ (n = 18), respectively. The average total VLP abundance was lower in winter ($5.4 \pm 0.2 \times 10^6$ VLPs mL⁻¹, n = 6) and then highest in summer ($9.2 \pm 4.5 \times 10^6$ VLPs mL⁻¹, n = 6). The different virus subpopulation (VLP 1, VLP 2 and VLP 3) abundances varied throughout the sampling period and ranged between $1.4 - 9.6 \times 10^6$, $1.5 - 4.7 \times 10^6$ and $0.3 - 0.7 \times 10^6$ VLPs mL⁻¹, respectively. These subpopulations made up varying proportions of the total VLP abundance, and as with station L4 were typically dominated by VLP 1 (28 – 76%; Figure 4.4b) and VLP 2 (20 – 59%; Figure 4.4c), and again VLP 3 made up the smallest proportion (4 – 12%; Figure 4.4d).

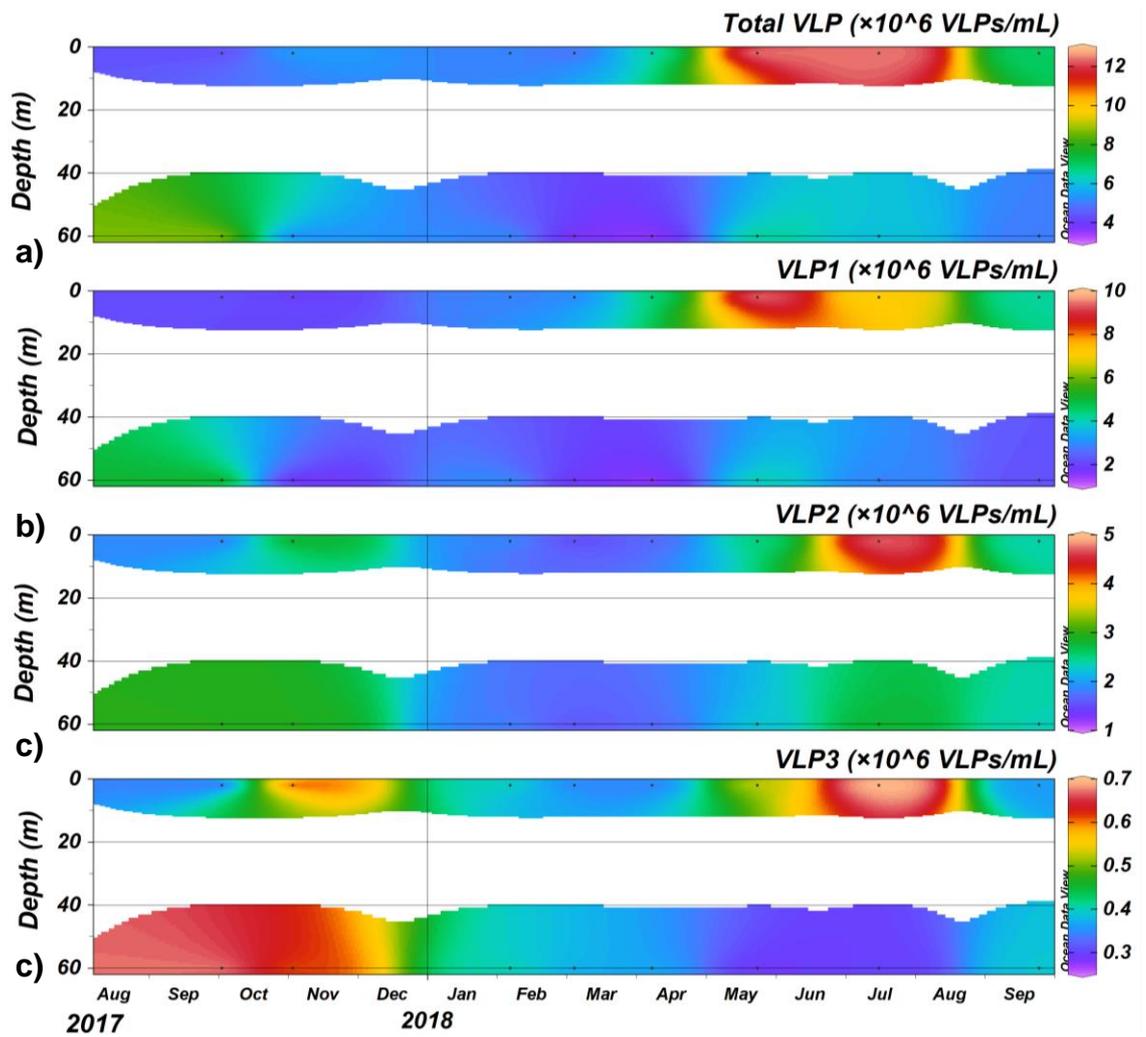


Figure 4.4 Abundances of (a) total virus-like particles (VLPs), and virus subgroups (b) VLP 1, (c) VLP 2 and (d) VLP 3 as VLPs mL^{-1} at the surface and 60 m at station E1 (August 2017 – September 2018).

4.3. Biogeochemical and abundance dynamics

4.3.1. *Biogeochemical properties*

The seawater temperatures at L4 throughout the water column ranged between 7.8 - 19.2°C (Figure 4.5a), with the coldest seawater temperature occurring in March at all depths, and the warmest seawater temperature recorded in July at the surface, when thermal stratification of the water column occurred between the surface and 10 m. This was consistent with the seasonal water column averages, indicating that the summer months were warmest ($14.9 \pm 1.8^\circ\text{C}$, $n = 52$), and spring was colder ($9.4 \pm 1.6^\circ\text{C}$, $n = 48$). Salinity measurements revealed only slight variations in salinity at L4 (33.4 – 35.4 PSU; Figure 4.5b), and the average seasonal salinity changed very little throughout the year: autumn (35.2 ± 0.1 PSU, $n = 40$); winter (35.0 ± 0.2 PSU, $n = 32$); spring (35.0 ± 0.3 PSU, $n = 48$) and summer (35.1 ± 0.1 PSU, $n = 52$). The lowest readings were indicated in surface water during March (33.4 PSU), and the highest salinity was in August 2017 at 50 m (35.4 PSU). Surface waters also had the greatest salinity variation (2 PSU), compared to bottom depths, where it only varied by 0.4 PSU throughout all seasons. Fluorescence readings at L4 ranged between 0.0 – 9.3 and indicated that during autumn and winter (Figure 4.5c; 0.9 ± 0.3 and 0.5 ± 0.6 , respectively), measurements varied little between the surface and the bottom. However, at the end of spring an increase occurred at 25 and 50 m, peaking in late-May at 50 m (9.3). This peak then decreased during June and July, but the fluorescence appeared to increase through the water column over the following months, with peaks in late-July (3.8) at 25 m and in mid-August (5.9) at 10 m which was maintained during September (2.8 ± 1.2).

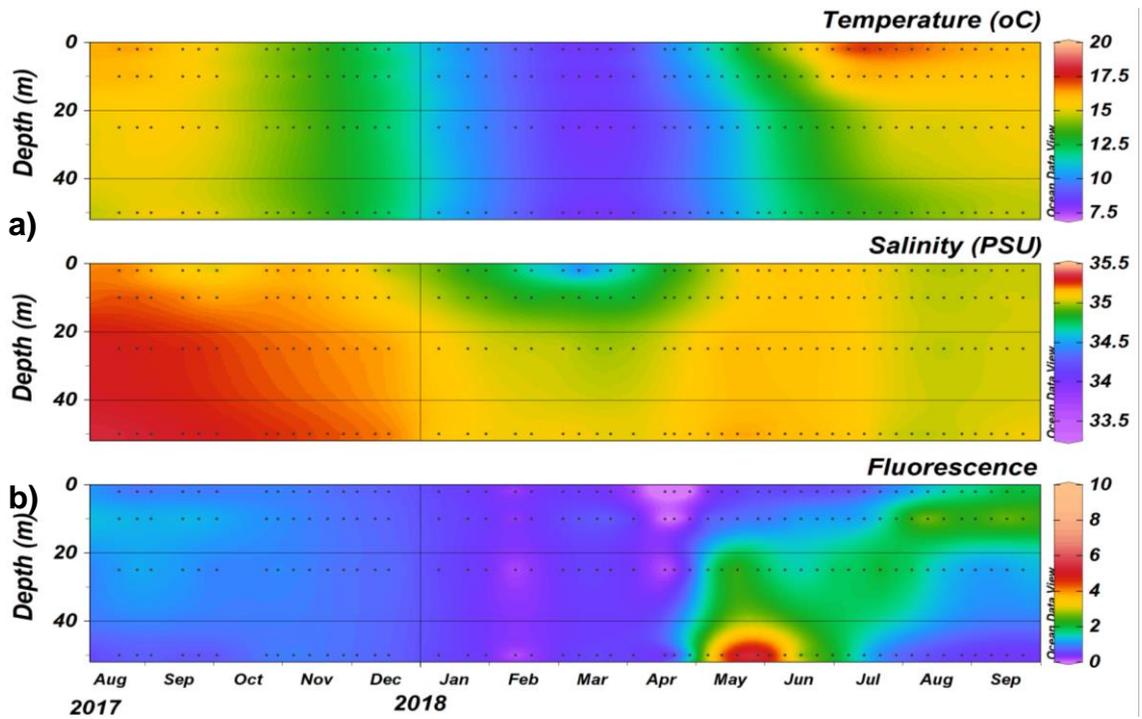


Figure 4.5 Depth profile of (a) seawater temperature, (b) salinity and (c) fluorescence values (in volts) through the water column at station **L4** (August 2017 – September 2018). Data provided by the Western Channel Observatory (<https://westernchannelobservatory.org.uk/>).

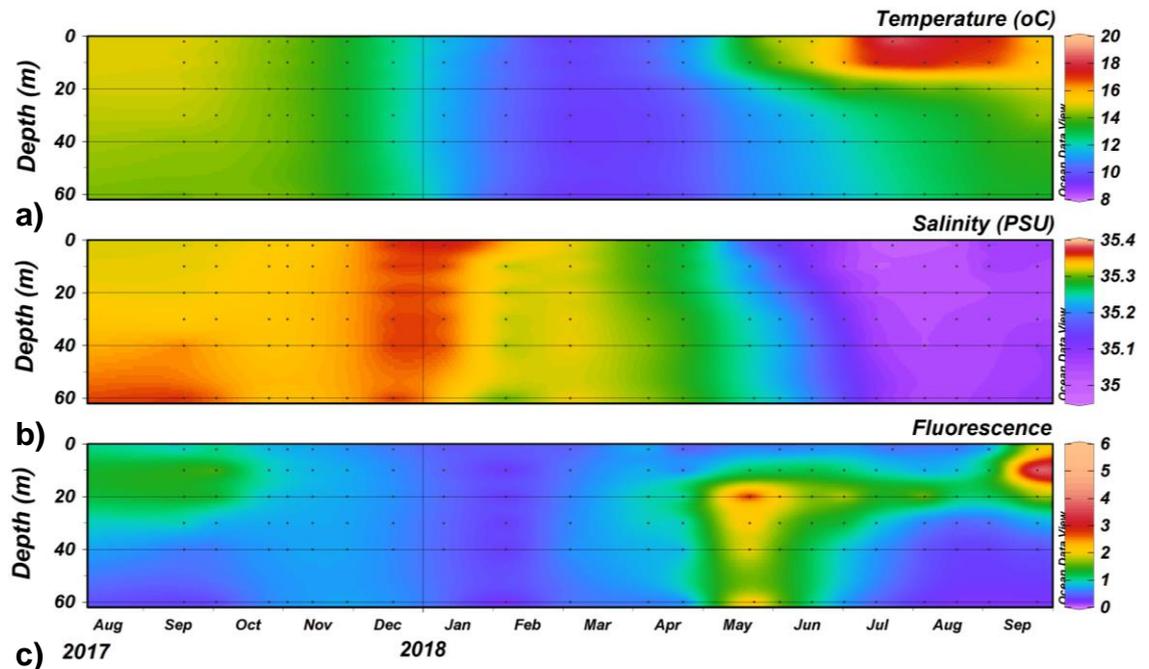


Figure 4.6 Depth profile of (a) seawater temperature, (b) salinity and (c) fluorescence values (in volts) through the water column at station **E1** (August 2017 – September 2018). Data provided by the Western Channel Observatory.

Similarly to L4, the seawater temperatures at E1 (Figure 4.6a) peaked in July at the surface (19.8°C), and the lowest temperature was recorded in March at 20 m (8.9°C). Winter showed the smallest temperature variation (10.1 – 12.0°C), and the largest variation occurred during summer months (11.1 – 19.8°C), and indicated a more pronounced thermal stratification between 10 and 20 m than at L4 during the same period. The salinity measurements at E1 varied between 35.0 – 35.4 PSU (Figure 4.6b). Salinity trends were very similar at all depths throughout the timeseries, starting off higher in autumn 2017, slowly reducing during winter and then quickly dropping during spring. The fluorescence range during autumn and summer was similar to L4 (Figure 4.6c; 0.2 – 2.4 and 0.2 – 3.0, respectively), and measurements were lowest during winter (0.0 - 0.7). The highest fluorescence occurred in May and September (5.8 at 25 m and again at 10 m), with a clear increase in fluorescence between April and May, and again in September in the sub-surface depths.

4.3.2. *Nutrients*

The concentration of nitrate at L4 (Figure 4.7a; 0 – 15.4 μM) increased during autumn and winter months, and reached a maximum at the surface during early-April, then dropped sharply by the end of that month (5.71 μM). This trend was similar throughout the water column, but most pronounced at the surface, where it then reached a summer minimum beneath the detectable range. This trend was consistent with two other major inorganic nutrients: silicate (Figure 4.7c; 0.05 – 6.4 μM) and phosphate (Figure 4.7d; 0.0 – 0.5 μM); which peaked at the surface in spring (6.4 and 0.5 μM , respectively), and were all at a minimum by the end of June, when nitrate and phosphate were undetected at the surface. Silicate concentrations (Figure 4.7c) began to increase again in late-September 2018 (4.09 μM) at 50 m and were beginning to increase at the surface as well after the minimum in June. Nitrite ranged between 0.0 – 2.44 μM (Figure 4.7b), and initially peaked during mid-September 2017 at all depths (increasing in concentration with depth and was highest at 50 m - 1.86 μM) and again in 2018 (reaching 2.4 μM at 50 m). However, nitrite was not detected in the summer months at the surface and 10 m only ranging between 0.0 – 0.17 μM . Ammonia ranged between 0.0 – 4.2 μM (Figure 4.7f), but

only showed slight variation in the upper 10 m (0.0 – 1.04 μM) and less variation at all depths in autumn and winter (0.0 – 0.3 and 0.1 – 0.5 μM). Ammonia concentrations then peaked at 50 and 25 m in late-April (4.2 and 4.1 μM , respectively) and there was a sustained maximum during July and August 2018 at 50 m (peaking at 2.7 μM). Ratios of nitrogen-to-phosphorous (N:P ratio; Figure 4.7e) were determined for station L4 and averaged 7.5 ± 6.7 ($n = 203$) between August 2017 and September 2018. Across all sampled dates and depths, the N:P ratio ranged between 0.0 – 42.75, and seasonal averages varied from a maximum in winter months (14.4 ± 2.3 , $n = 39$) to a minima during summer months (1.3 ± 1.5 , $n = 50$).

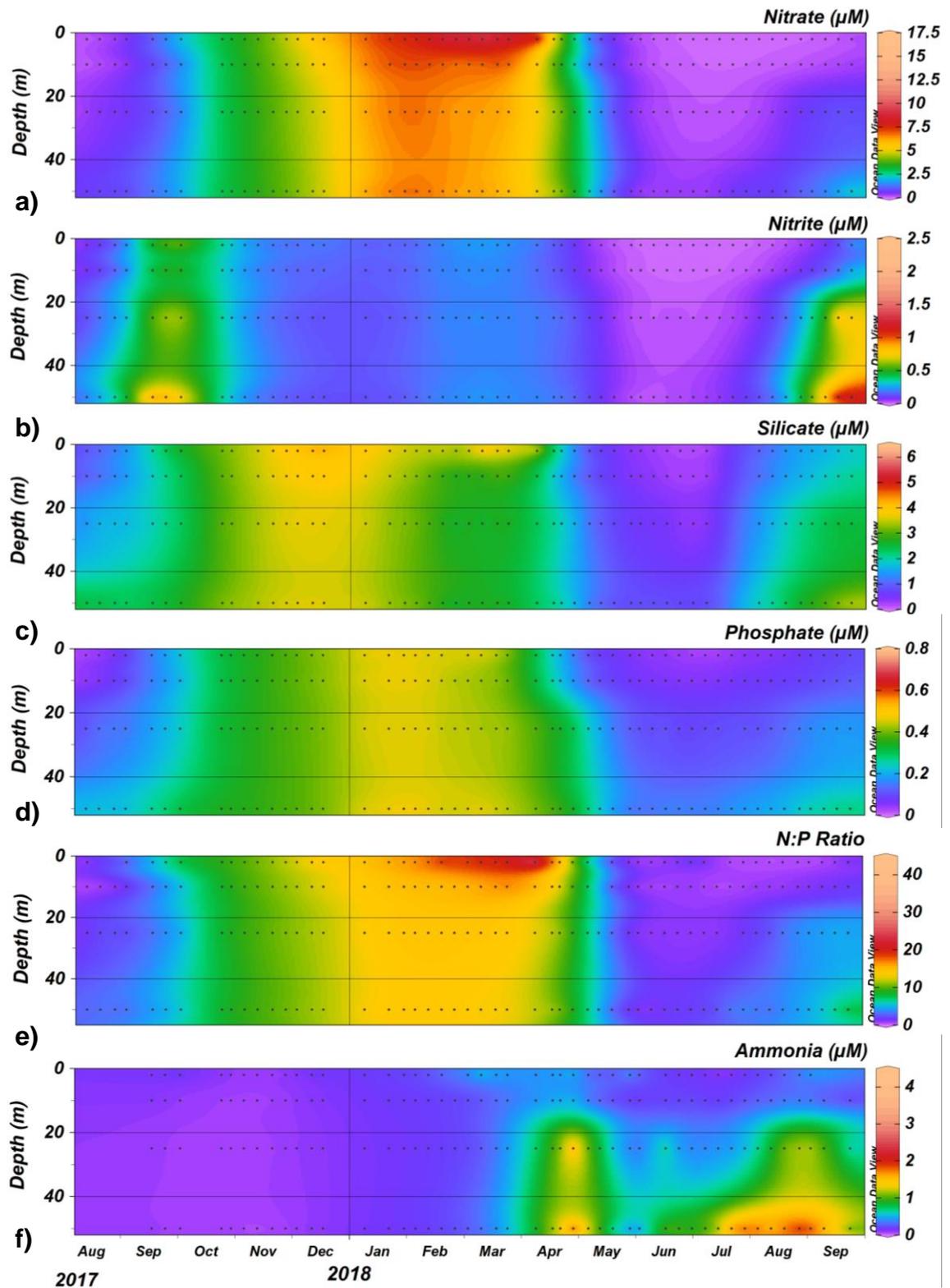


Figure 4.7 Concentrations of (a) nitrate, (b) nitrite, (c) silicate, (d) phosphate, (e) N:P ratio and (f) ammonia through the water column at station L4 (August 2017 – September 2018). Data provided by the Western Channel Observatory.

At E1, the nitrate, phosphate and silicate (Figure 4.8a, d and c, respectively) concentrations reflected the trends at L4, with some differences in the concentration ranges (0.0 – 5.4 μM , 0.0 – 0.9 μM and 0.1 – 4.6 μM , respectively). The highest nitrate and silicate concentrations occurred in late-September 2018 (5.4 μM and 4.6 μM , respectively), and as with L4, indicated winter as the season with the highest nitrate concentrations throughout the water column (4.6 – 5.3 μM). Although silicate and phosphate concentrations were higher during the same period as nitrate, the nutrient concentration peaks appear to be staggered by 2 months, beginning with silicate in November (2.9 - 3.6 μM , peaking at 30 m) and then phosphate in February (0.4 - 0.5 μM , peaking at 40 m). Staggered concentration observations were also observed at L4, however the highest concentrations of both silicate and phosphate occurred in early-September (4.6 μM and 0.9 μM , respectively) at 20 m. Additionally, the upper 10 m between May and late-September 2018 indicated much lower nitrite (Figure 4.8b), nitrate (0.0 – 0.05 μM) and phosphate concentrations (0.0 – 0.1 μM). Although silicate was detectable in all samples, its lowest concentration occurred in June (0.13 μM), aligning with the summer minima of nitrite, nitrate and phosphate. Nitrite concentrations indicated two peaks occurring in September 2017 and 2018 between 20 - 60 m, with the larger peak occurring in 2018, reaching 1.5 μM at 60 m. Nitrite was undetected in the upper 10 m during summer months, and only ranged between 0.0 – 0.3 μM for the rest of the year at these depths. Nitrite levels during spring and summer ranged between 0.0 and 0.4 μM and began to increase during the end of summer 2017 and 2018, with the maximum occurring in 2018. Ammonia ranged between 0.0 – 2.8 μM (Figure 4.8f), but was limited in the upper 10 m throughout the timeseries (0.0 – 0.4 μM) and throughout the water column during autumn and winter (0.0 – 0.2 μM and 0.2 μM , respectively). However, in spring ammonia concentrations began to increase between 30 and 60 m, and during summer months reached the highest concentration (2.8 μM) at 30 m in late-August. The concentration of ammonia then began to decrease during September. The thermocline present at station E1 indicated a greater influence on nutrient concentrations in the upper 10 m, compared to station L4, for nitrate, nitrite, phosphate and ammonia.

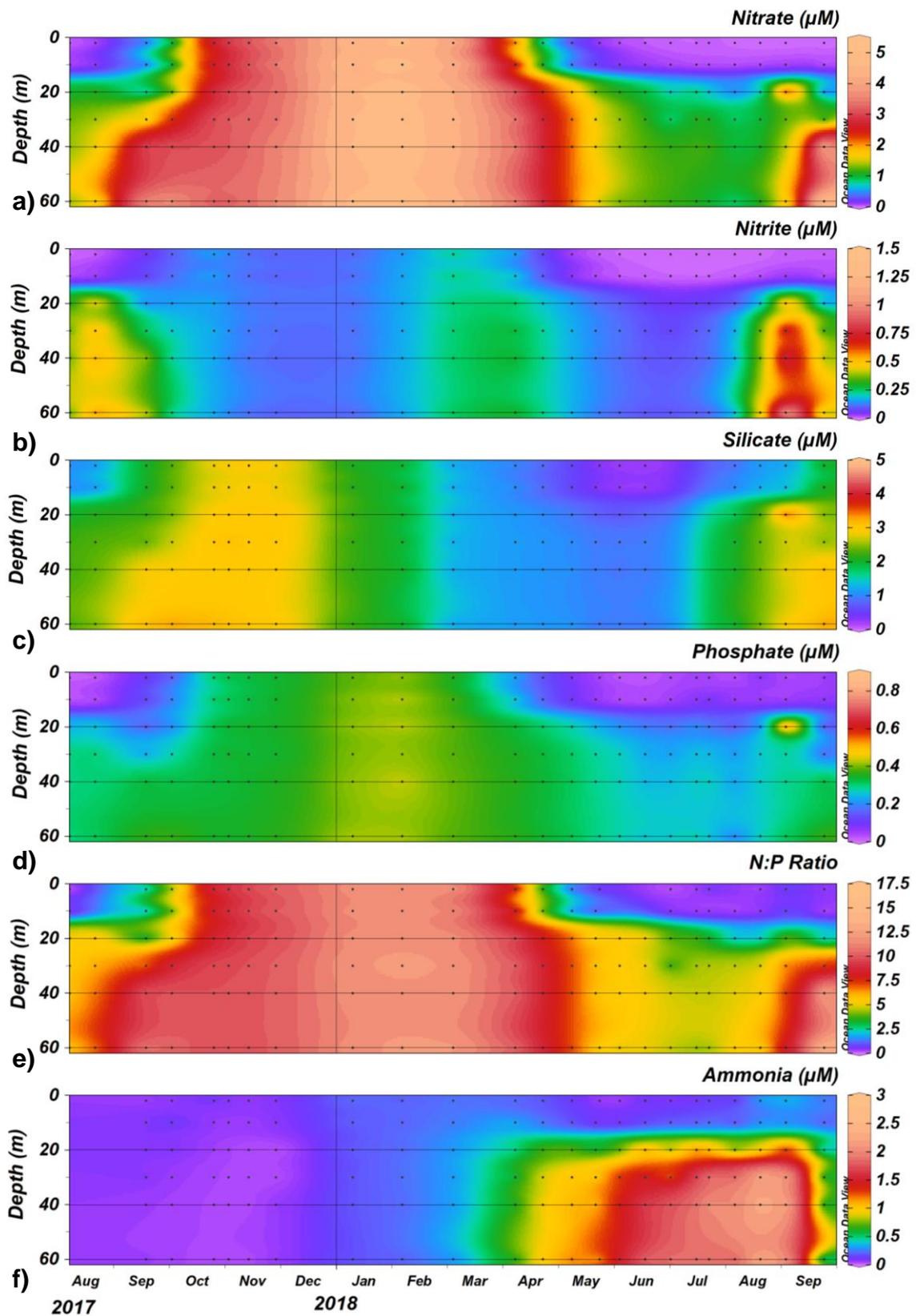


Figure 4.8 Concentrations of (a) nitrate, (b) nitrite, (c) silicate, (d) phosphate, (e) N:P ratio and (f) ammonia through the water column at station **E1** (August 2017 – September 2018). Data provided by the Western Channel Observatory.

4.3.3. Nano- and picoplankton populations

Synechococcus sp., pico-eukaryote and nano-eukaryote cell abundances at L4 (Figure 4.9a, b and c) varied between 85 - 328,160 cells mL⁻¹, 308 – 39,770 cells mL⁻¹ and 86 – 7,423 cells mL⁻¹, respectively, throughout the year. *Synechococcus* sp. began to bloom at the end of summer in the upper 10 m and then declined by mid-September. The lowest *Synechococcus* sp. abundance was determined in late August 2017 in surface waters (85 ± 12 cells mL⁻¹), indicating a stark contrast compared to the same time point in 2018 (202,706 ± 3,021 cells mL⁻¹). The lowest *Synechococcus* sp. abundance was observed in spring (791 – 7,275 cells mL⁻¹), and the highest abundance in summer (195 – 202,706 cells mL⁻¹). Pico-eukaryotic cells were present throughout the year and also reached a maximum at 10m during the end of summer (39,770 ± 616 cells mL⁻¹), which is also the season when the lowest abundance (308 ± 60 cells mL⁻¹) occurred, at 50 m. Nano-eukaryotes also indicated similar trends in the upper 10 m of L4, peaking at the end of summer 2017 (7,423 ± 108 cells mL⁻¹), however maximum abundance was not as high in summer 2018 (2,224 ± 135 cells mL⁻¹) and numbers were lowest in winter months (231 - 763 cells mL⁻¹). Coccolithophores (Figure 4.9d) were most abundant in summer in the upper 10 m at the end of June (4,568 cells mL⁻¹), and abundance was much lower during autumn (2 – 230 cells mL⁻¹), winter (1 – 186 cells mL⁻¹) and spring (5 – 130 cells mL⁻¹). *Phaeocystis* sp. (Figure 4.9e) was only detected in spring (73 – 1530 cells mL⁻¹) and summer (138 – 4,072 cells mL⁻¹), and the highest concentration was detected at 50 m in mid-June. During summer, *Phaeocystis* sp. was present at all depths and appeared to peak twice at the surface in early July (4,072 ± 181 cells mL⁻¹) and August (2,684 ± 77 cells mL⁻¹). Although present throughout much of the year, Cryptophyte abundance at L4 (Figure 4.9f; 12 – 1,223 cells mL⁻¹) appeared to be most prevalent in autumn (114 – 1,223 cells mL⁻¹). Peaking in early-October at 10 m and abundant at all depths during November (555 – 906 cells mL⁻¹), the prolonged cryptophyte autumn maximum then began to decrease during winter.

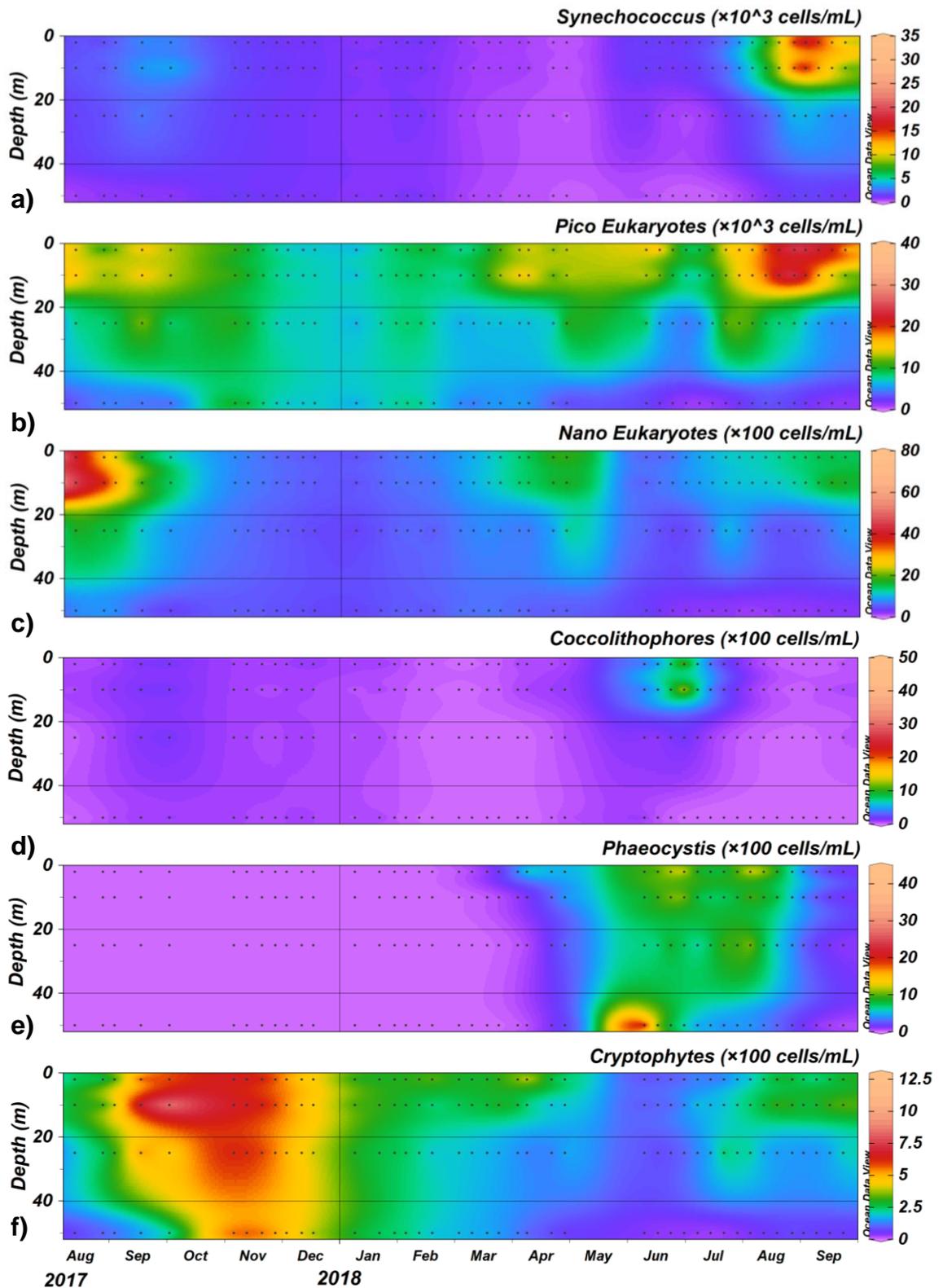


Figure 4.9 Abundances of (a) *Synechococcus*, (b) picoeukaryotes, (c) nanoeukaryotes, (d) coccolithophores, (e) *Phaeocystis* sp. and (f) cryptophytes as cells mL⁻¹ through the water column at station L4 (August 2017 – September 2018). Data provided by the Western Channel Observatory.

Numbers of nano- and picoplankton at E1 for were generally lower than at L4 (Figure 4.10a, b and d), and abundance varied between 430 – 119,514, 91 – 29,111 and 16 – 2,999 cells mL⁻¹, for *Synechococcus*, pico-eukaryotes and nano-eukaryotes, respectively. All the phytoplankton groups investigated at E1 increased in early-autumn, except for *Phaeocystis* sp. and the cryptophytes in 2018, and the autumn bloom lasted for 1-2 months. The abundance of *Synechococcus* sp. and nano-eukaryote was highest during the initial autumn maxima; however pico-eukaryotes were most abundant during the summer maxima, coinciding with a second *Synechococcus* bloom. *Synechococcus* sp. abundance was lowest during spring at 10 m, which was also the season with the lowest overall abundance (430 – 11,306 cells mL⁻¹). This was different from the pico- and nano-eukaryotes, which both reached minima during summer at 60 m, and had lower abundance ranges in winter (3,923 - 8,757 and 277 - 743 cells mL⁻¹, respectively). As with L4, *Phaeocystis* sp. cells (Figure 4.10c) were only detected in spring (191 – 9,019 cells mL⁻¹) and summer (47 – 2,228 cells mL⁻¹), however, they were more abundant at E1 during the spring bloom than at L4. Additionally, the *Phaeocystis* sp. bloom at E1 occurred throughout the water column, but the largest cell concentrations occurred in late-April in the upper 10 m (8,640 - 9,019 cells mL⁻¹). Cryptophytes ranged between 6 – 1,797 cells mL⁻¹ (Figure 4.10e), with maximum abundance occurring in the upper 20 m during the first half of autumn. The remaining seasons indicated much lower cryptophyte abundances (6 – 542 cells mL⁻¹) which were lowest in summer (6 – 222 cells mL⁻¹). Coccolithophore abundance (Figure 4.10f) changes throughout the year reflected the majority of the E1 plankton populations, with a bloom in the upper 20 m in early-autumn (330 ± 205 cells mL⁻¹, n = 9), followed by comparatively low numbers in the following spring (13 – 90 cells mL⁻¹). In summer, a second coccolithophore bloom occurred at the same time as at L4, reaching 734 cells mL⁻¹ (no replicates available) at 30 m.

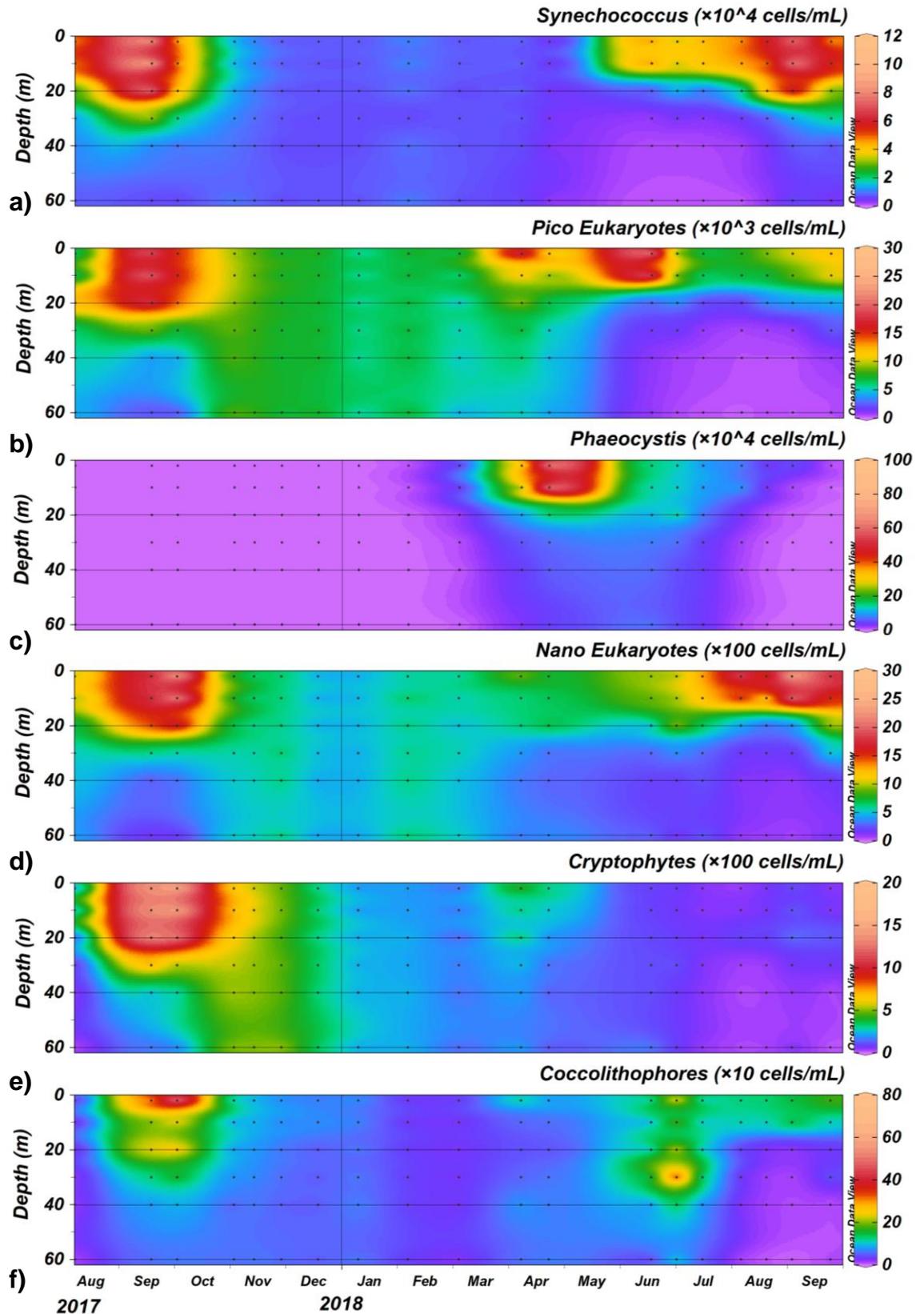


Figure 4.10 Abundances of (a) *Synechococcus*, (b) picoeukaryotes, (c) *Phaeocystis* sp., (d) nanoeukaryotes, (e) cryptophytes and (f) coccolithophores as cells mL⁻¹ through the water column at station E1 (August 2017 – September 2018). Data provided by the Western Channel Observatory.

4.3.4. Bacterial abundance and production

Total heterotrophic bacterial abundance ranged between $5.4 - 36.3 \times 10^5$ cells mL⁻¹ (Figure 4.11a) and peaked in early-June and in late-August and September. The peaks in June and September 2018 were highest in surface samples, and the lowest total bacterial abundance occurred in January at 10 m (5.4×10^5 cells mL⁻¹). Based on their flow cytometric profile, total heterotrophic bacteria were separated into two groups high (HNA) and low (LNA) nucleic acid-containing bacteria, and abundance ranged between $2.7 - 26.4 \times 10^5$ cells mL⁻¹ (Figure 4.11b) and $0.8 - 12.0 \times 10^5$ cells mL⁻¹ (Figure 4.11c), respectively. These subgroups followed very similar annual trends; however, the HNA bacteria tended to be more dominant throughout the year (48 – 90%), especially at 50 m where HNA bacteria made up 90% of the summer maximum. LNA bacteria indicated a prolonged decrease in abundance at all depths following the autumn maximum in September, compared to HNA bacteria which dropped much quicker.

Measured ³H-leucine incorporation rates at L4 ranged between $4.6 - 149$ pmol Leu L⁻¹ h⁻¹ and were highest during summer months and lowest during autumn. These leucine incorporation rates were then used to derive the bacterial production rates, which were calculated using the theoretical conversion factor (TCF) described by Simon and Azam (1989) of 1.55 kg C mol Leu⁻¹, and ranged between $7.2 - 231$ ng C L⁻¹ h⁻¹ (Figure 4.11d). The bacterial production maximum occurred in the upper 10 m of the water column; however, a peak at 25 m (215 ng C L⁻¹ h⁻¹) occurred one month previously to the upper 10 m peak. On average, bacterial production was highest in the upper 10 m (Surface = 61 ± 64 ng C L⁻¹ h⁻¹ and 10 m = 68 ± 69 ng C L⁻¹ h⁻¹) and then decreased with depth (25 m = 54 ± 63 ng C L⁻¹ h⁻¹ and 50 m = 34 ± 39 ng C L⁻¹ h⁻¹).

Seasonally, bacterial production was on average highest during summer (104 ± 69 ng C L⁻¹ h⁻¹) and lowest during autumn (12.7 ± 3.5 ng C L⁻¹ h⁻¹). Bacterial production was also used to determine the bacterial production per cell at station L4, which varied between $0.05 - 1.79 \times 10^{-7}$ ng C cell⁻¹ h⁻¹ (Figure 4.11e), and on average was highest during summer ($1.8 \pm 0.1 \times 10^{-7}$ ng C cell⁻¹ h⁻¹) and lowest during autumn ($0.17 \pm 0.05 \times 10^{-7}$ ng C cell⁻¹ h⁻¹). Normalising to a per cell bacterial

production rate indicated highest measurements at 25 m ($1.79 \pm 0.05 \times 10^{-7}$ ng C cell⁻¹ h⁻¹), instead of in the upper 10 m as with the total bacterial production rate. However, the average rate varied very little in the upper 25 m ($0.38 - 0.41 \times 10^{-7}$ ng C cell⁻¹ h⁻¹), and was lowest at 50 m ($0.26 \pm 0.20 \times 10^{-7}$ ng C cell⁻¹ h⁻¹).

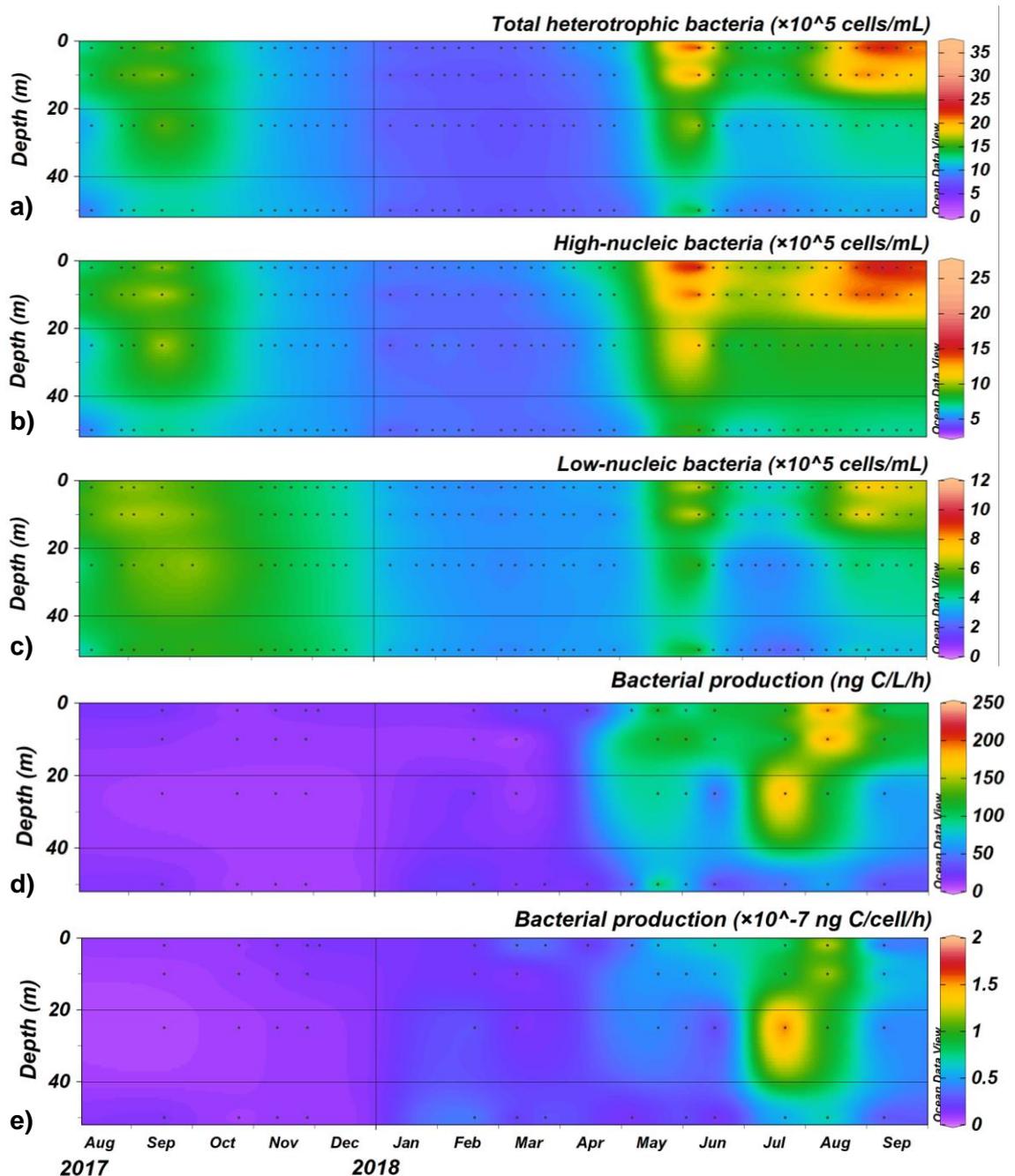


Figure 4.11 Abundances of (a) total heterotrophic bacteria, (b) high nucleic acid-containing (HNA) bacteria, (c) low nucleic acid-containing (LNA) bacteria, (d) bacterial production and (e) bacterial production per cell through the water column at station **L4** (August 2017 – September 2018). Abundance data in cells mL⁻¹. Bacterial production values determined using theoretical conversion factor of 1.55 kg C mol Leu⁻¹ (Simon and Azam, 1989b). Bacteria abundance data provided by the WCO.

The total heterotrophic bacterial abundance at E1 ranged between $3.6 - 27.6 \times 10^5$ cells mL⁻¹ (Figure 4.12a), and the HNA (Figure 4.12b) and LNA (Figure 4.12c) bacteria varied between $1.8 - 19.8 \times 10^5$ and $0.6 - 10.4 \times 10^5$ cells mL⁻¹, respectively. The same annual trends that occurred at L4 in September and June also occur at E1; however, in contrast to L4 the autumn maximum (27.1×10^5 cells mL⁻¹; no replicates available) was similar to the summer maximum (27.6×10^5 cells mL⁻¹; no replicates available).

Measurements of ³H-leucine incorporation at station E1 (Figure 4.12d) were also measured and ranged between $0.0 - 104$ pmol L⁻¹ h⁻¹ and were used to derive the bacterial production for E1 samples (Figure 4.12e) using the same conversion factor as at station L4. Bacterial production rate ranged between $0.0 - 160$ ng C L⁻¹ h⁻¹, and was highest on average during spring (40.7 ± 59.5 ng C L⁻¹ h⁻¹, n = 12) and summer (30.2 ± 13.3 ng C L⁻¹ h⁻¹, n = 6), compared to winter (11.8 ± 2.7 ng C L⁻¹ h⁻¹, n = 6) and autumn (5.5 ± 14.5 ng C L⁻¹ h⁻¹, n = 16) when it was lowest. Similarly, to station L4, the depth made a large difference to the maximum determined bacterial production and ranged between $0.0 - 34.9$ ng C L⁻¹ h⁻¹ at 60 m and $0.0 - 160$ ng C L⁻¹ h⁻¹ at the surface between August 2017 and September 2018. Bacterial production per cell was also determined for E1 measurements and ranged between $0.0 - 0.98 \times 10^{-7}$ ng C cell⁻¹ h⁻¹. Average bacterial production per cell at E1 reflected station L4, and was highest in summer ($0.4 \pm 0.2 \times 10^{-7}$ ng C cell⁻¹ h⁻¹, n = 6) and lowest in autumn ($0.0 \pm 0.1 \times 10^{-7}$ ng C cell⁻¹ h⁻¹, n = 18). As with station L4, bacterial production per cell was also determined to be lowest on average at the bottom of the water column at E1 ($0.12 \pm 0.14 \times 10^{-7}$ ng C cell⁻¹ h⁻¹, n = 27) and was highest at the surface ($0.27 \pm 0.33 \times 10^{-7}$ ng C cell⁻¹ h⁻¹, n = 27).

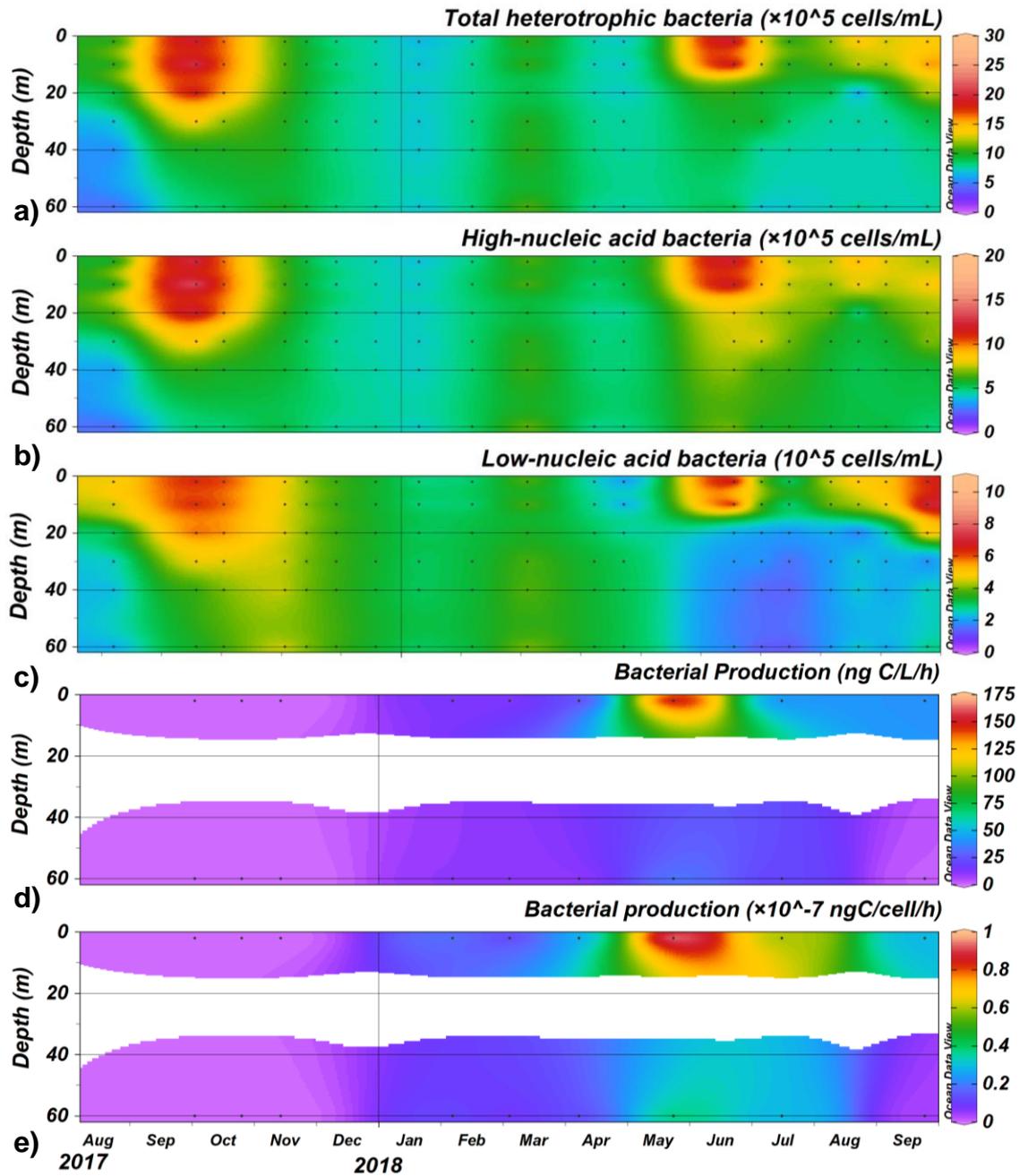


Figure 4.12 Abundances of (a) total heterotrophic bacteria, (b) high nucleic acid-containing (HNA) bacteria, (c) low nucleic acid-containing (LNA) bacteria, (d) bacterial production and (e) bacterial production per cell at station **E1** (August 2017 – September 2018). Abundance data in cells mL⁻¹ and throughout the water column. Bacterial production only at the surface and 60 m, and values determined using theoretical conversion factor of 1.55 kg C mol Leu⁻¹ (Simon and Azam, 1989b). Bacteria abundance data provided by the WCO.

4.3.5. *Virus-to-bacteria ratios*

As variability existed in the virus subgroups (Figure 4.3 and Figure 4.4) which make up the total virus-like particle abundances, when the virus-to-bacteria ratios were investigated the different VLP groups were also considered.

The ratio of total virus-to-bacteria (total VBR; Figure 4.13a) ranged from 3 - 28 across all depths and samples at station L4, and on average was highest in summer (16.9 ± 6.7 ; $n = 60$), and lowest during autumn (8.8 ± 3.7 ; $n = 36$). On average, the total VBR increased slightly with depth, being 11.0 ± 5.1 ($n = 45$) at the surface and was 12.9 ± 7.4 ($n = 45$) at 50 m depth. The VLP1-to-bacteria ratio (VBR 1; Figure 4.13b) ranged from 0.5 – 25, and reflected the total VBR, with higher values in summer (25 ± 3.5 ; $n = 60$) and lowest in autumn (4.4 ± 0.5 ; $n = 60$), and also displayed an increase with depth. The VLP2-to-bacteria ratio (VBR 2; Figure 4.13c) ranged from 1.1 - 9.6 and indicated little seasonal average variation from autumn to spring (5.2 – 5.8; $n = 96$), but was lower in summer (2.9 ± 0.9 ; $n = 45$), there was also little depth variation in average values throughout the water column (4.5 – 4.6, $n = 180$). The VLP3-to-bacteria ratio (VBR 3; Figure 4.13d) ranged between 0.1 – 1.8 and again varied little seasonally (except in summer), and with depth.

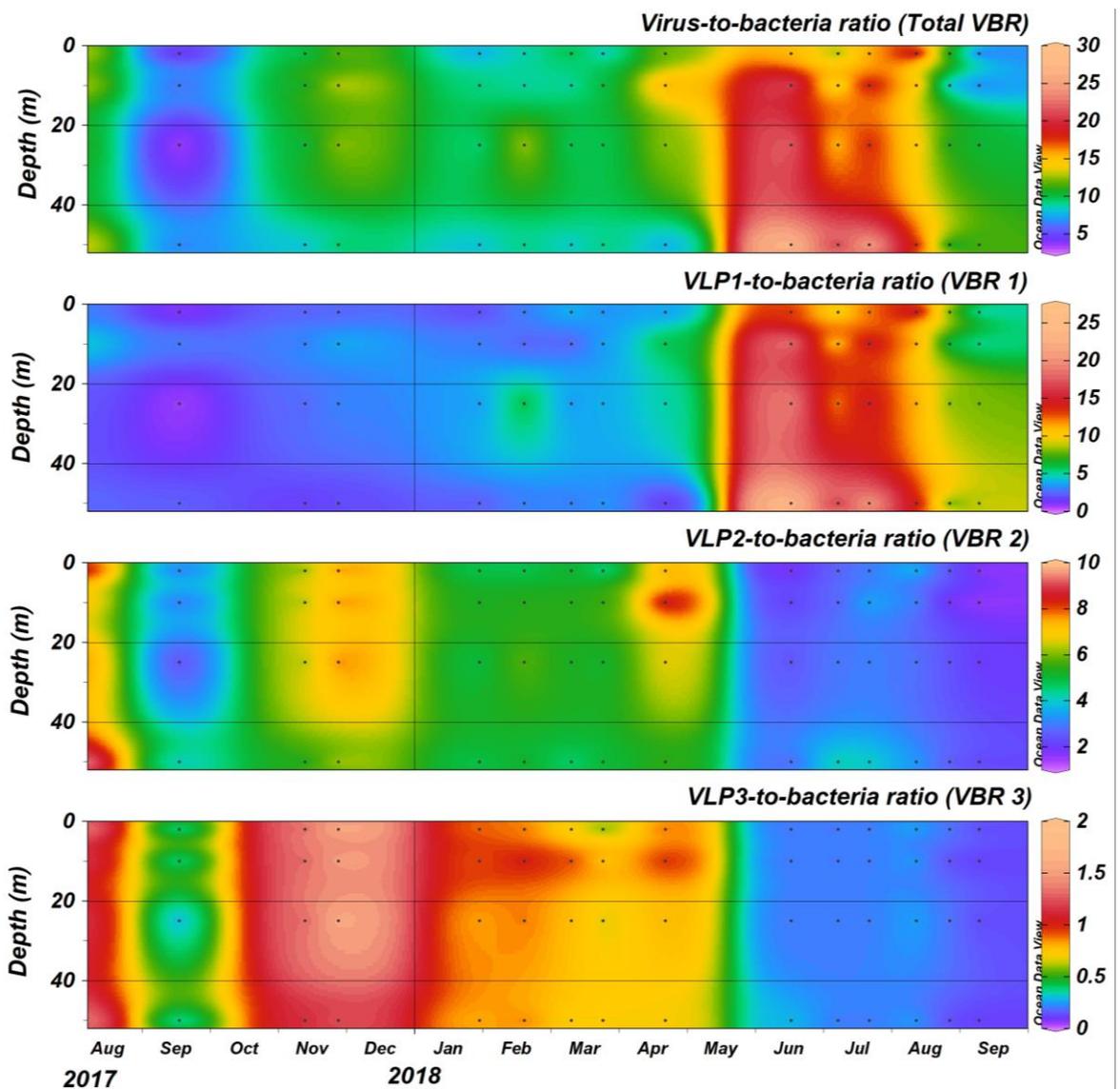


Figure 4.13 Abundances of (a) total virus-to-bacteria ratio, (b) VLP 1-to-bacteria ratio, (c) VLP 2-to-bacteria ratio, (d) VLP 3-to-bacteria ratio throughout the water column at station L4 (August 2017 – September 2018).

The ratio of total bacteria-to-virus-like particles (total VBR; Figure 4.14a) ranged from 2.8 - 19 across all depths and samples at station E1, and on average was highest in summer (13.6 ± 7.2 ; $n = 6$), and lowest during winter (7.5 ± 7.1 ; $n = 6$) and on average, was highest at the surface (7.4 ± 5.0 ; $n = 24$). The VLP1-to-bacteria ratio (VBR 1; Figure 4.14b) ranged from 1.1 – 10.5, and also had higher values in summer (7.3 ± 4.5 ; $n = 6$) and was lowest in autumn (2.7 ± 2.7 ; $n = 12$), with little variation in depth. The values of VBR 2 and VBR 3 ranged from 1.1 – 7.1 and 0.2 – 1.1, respectively (Figure 4.14c and Figure 4.14d). Both were highest in summer and also showed little variation in depth.

Higher total VBR values were indicated at station L4 (28.3 and 18.7, respectively) when compared to station E1, with the lowest values remaining similar (2.8 and 2.8, respectively) between the stations. This was the same with the other VBR values and is further demonstrated by the overall average at each station, which was higher at L4, compared to E1 (12.0 ± 5.7 and 7.0 ± 3.9 , respectively), which has also been demonstrated previously in summer during a bloom in *Emiliana huxleyi* by Wilson *et al.* (2002) and for other coastal regions (Wommack and Colwell, 2000). Both stations indicated that the highest total VBR occurred in summer, and the lowest in winter. As with the previously mentioned VLP abundances (Section 4.2.2), there was a clear difference in the seasonal cycles between the VLP1 subgroup and the other two virus subgroups (VLP2 and VLP3), which was also demonstrated in the VBR values. However, there were differences in the total VBR and VBR 1 values at the top and bottom of the water column between the two stations, where the highest was recorded at the bottom depths of L4, whereas they were determined to be at the surface at E1. The VBR 2 and VBR 3 values also indicated differences between the stations, where they were both higher at all depths (with the exception of September at L4) from August 2017 until April 2018 at L4, but at E1 there was no clear similar trend for the same period.

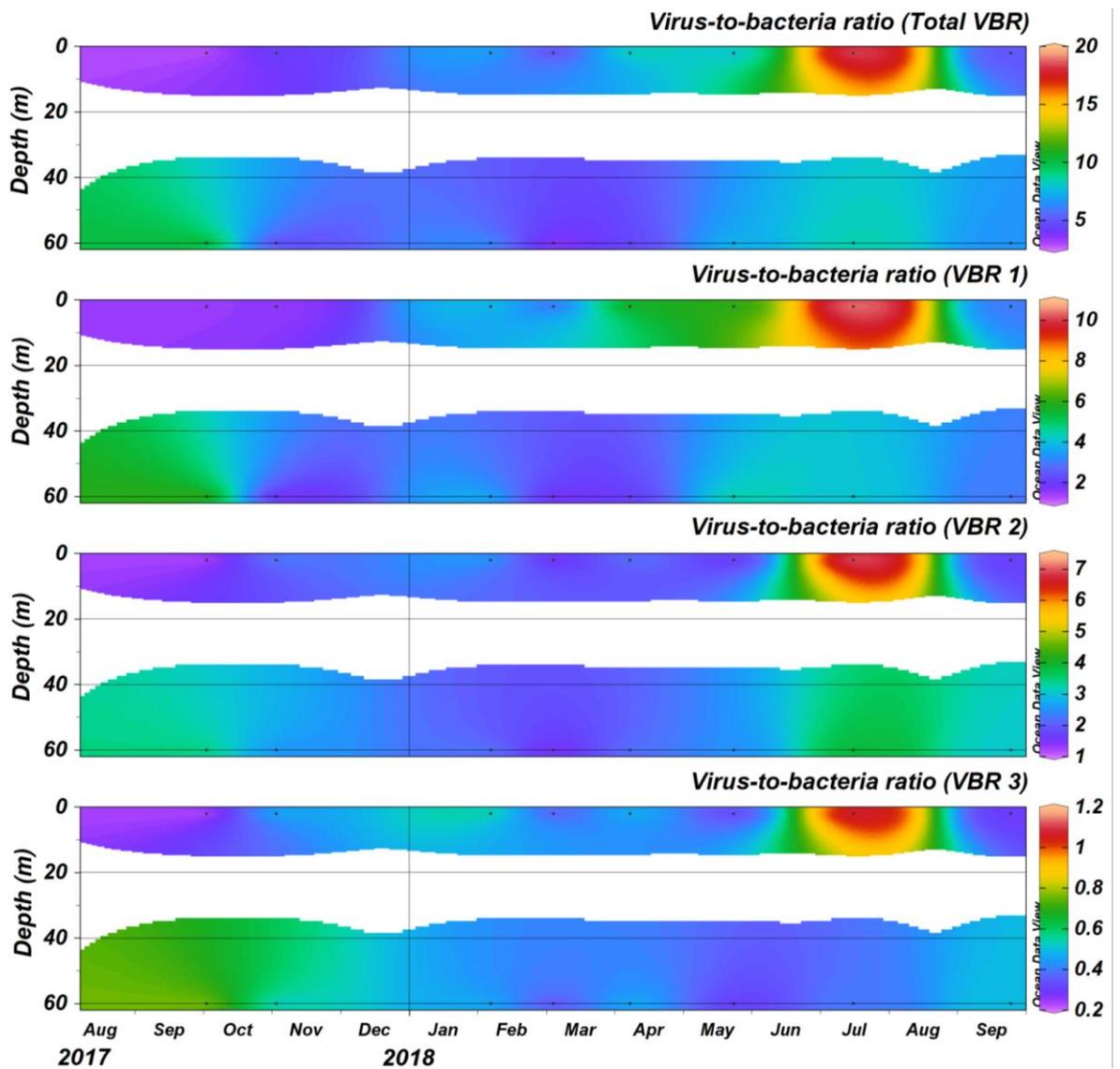


Figure 4.14 Abundances of (a) total virus-to-bacteria ratio, (b) VLP 1-to-bacteria ratio, (c) VLP 2-to-bacteria ratio, (d) VLP 3-to-bacteria ratio at the surface and 60 m at station **E1** (August 2017 – September 2018).

4.4. Microbial diversity

The microbial diversity was explored in surface and bottom waters at stations L4 and E1 in the English Channel between August 2017 and September 2018. Microbes of interest were investigated by sequencing targeted genes from DNA extracted from seawater samples to characterise the microbial community (16S V4 *rRNA*) and also to investigate seasonal variations in the diversity of bacteria containing the *xoxF5* gene clade (Section 2.10.7 and 2.10.8).

4.4.1. 16S rRNA *bacterial diversity*

The temporal changes of the bacterial assemblage at the surface and bottom depths of stations L4 and E1 was determined by examining the V4 region of the 16S rRNA gene. Sequencing of V4 amplicons yielded 1,410,967 reads in total from 50 samples after pre-processing steps using the DADA2 pipeline (filtering, merging and removal of chimeric and singleton reads; Section 2.10.8). After these steps were completed, 25% of the total reads had been removed (including 135,213 chimeric reads), with 323,067 reads derived from L4 samples and 730,088 reads derived from E1 samples. Therefore, although E1 samples only made up 14 of the 50 samples, they accounted for 69% of the total number of sequence reads. Sequences were then assigned into amplicon sequence variants (ASVs; described further in Section 2.10.8), assigning 2432 ASVs in total, where only 53 ASVs represented 50% of the reads. To mitigate bias resulting from uneven sampling depth (number of sequences per sample), samples with < 1000 reads were removed and samples with >100,000 reads were re-sampled to 50,000 reads. This removed three station L4 samples (977 total reads), and only one sample (from station E1) was sub-sampled. These ASVs were then annotated using the SILVA database for 16S rRNA sequences. As expected, the majority of sequences were aligned with bacterial taxonomic sequences (89.2%), however 10.1% were attributed to archaea and 0.73 % aligned with eukaryotes, these were both removed leaving 2289 bacteria-associated ASVs. Of these ASVs, 51 - 52% and 65 - 66% of ASVs could be classified to the genus and family taxonomic level in station L4 and E1 samples, respectively. Furthermore,

82 – 89% and 83 - 90% of ASVs could be classified to the order and class taxonomic levels in samples from stations L4 and E1, respectively.

In general, similar taxonomic groups dominate at both depths of station L4 (Figure 4.15), such as *Proteobacteria* which is the dominant phylum and made up 44 – 83% and 39 - 88% of the total bacterial community over the year at the surface and 50 m, respectively. *Bacteroidetes* made up 8 - 36% and 6 – 49% at the surface and 50 m, respectively and then the *Cyanobacteria* (2 – 24% at both depths) and *Actinobacteria* (3 – 16% and 2 – 11% at the surface and 50 m, respectively) phyla, which were less dominant throughout the year. Within the *Proteobacteria*, the *Gamma*- and *Alphaproteobacteria* were the predominant classes, and at the surface the *Gammaproteobacteria* ranged between 24 – 56% of the surface bacterial community, and was highest in late-July, but also lowest in early-July, showing variability between seasons. This contrasts with the bottom of the water column where there was a clear seasonal cycle, decreasing during autumn 2017 ($51 \pm 10\%$) to a minimum in spring 2018 ($25 \pm 7\%$), before increasing over the summer months ($52 \pm 9\%$).

The *Gammaproteobacteria* were primarily made up of the SAR86 clade, *Cellvibrionales*, *Oceanospirillales* and the HOC 36 clade. The SAR86 clade was present in all samples, comprising 7 – 39% of the total surface bacterial community at station L4 and had a consistent abundance throughout the year, averaging 18 – 24% at both depths. Additionally, the *Oceanospirillales* order was also detectable in all seasons, but at a lower abundance, comprising 2 – 11% of the surface community, and was highest in autumn and lowest in early-summer. The HOC 36 clade and *Cellvibrionales* exhibit an apparent switch in dominance between the two orders, where HOC 36 was only detected in autumn and winter months (2 – 16% of the surface community) and then *Cellvibrionales* was only detected from late-winter until late-summer, consisting of 4 – 16% of the surface community. At the bottom of the L4 water column, the SAR86 clade constituted a similar percentage of the total bacterial community as at the surface (4 – 48%) and had a more variable seasonal cycle averaging $7 \pm 4\%$ during spring and $26 \pm 10\%$ during autumn and summer. *Oceanospirillales* demonstrated a slightly higher abundance compared to surface

samples (4 – 18%), but the same seasonal maxima and minima. The HOC36 clade was not as prevalent at 50 m (2 – 8% of the community) but was again only detectable during autumn and winter. However, the *Cellvibrionales* order (3 – 15% of the community at 50 m) was detected in all seasons and was highest during early-summer (averaging 8 – 10%) and lower in autumn and winter (averaging 3 – 5%).

The *Alphaproteobacteria* primarily consisted of the *Rhodobacterales*, *Thalassobaculales* and *Parvibaculales* orders. The *Rhodobacterales* order made up 3 – 35% of the total bacterial community at both depths of station L4, and also showed the same seasonal variation where abundance was on average highest during spring and summer months at the surface ($22 \pm 7\%$) and 50 m ($19 \pm 8\%$) and the lowest during autumn and winter ($5 \pm 1\%$ and $6 \pm 2\%$) at the surface and 50 m depth, respectively. The *Thalassobaculales* and *Parvibaculales* orders were only detected in autumn and winter at the surface (and in late summer at 50 m), and only made up 2 – 8% of the microbial community at both depths. The *Bacteroidetes* were the other main phylum detected at station L4, which principally consisted of the *Flavobacteriales* order (within the *Bacteroidia* class), ranging between 8 – 36% and 6 – 49% of the total community at the surface and 50 m depths, respectively. The *Chitinophagales* were another order detected belonging to the *Bacteroidetes*, but only comprised 2 – 5% of the bacterial community, except at one sample point at 50 m depth in late-April, where it made up 29% of the community at 50 m. At the surface, the *Bacteroidetes* do not vary seasonally on average, ranging between $19 \pm 6\%$ and $20 \pm 20\%$, compared to 50 m where they only comprised $11 \pm 3\%$ of the total bacterial community in autumn, and then increase to an average of $39 \pm 14\%$ in spring. This average maximum in the *Bacteroidia* class occurs during the *Gammaproteobacteria* minima, perhaps indicating competition between the SAR86 and *Flavobacteriales* orders. Some of the less abundant orders at station L4 include the *Synechococcales* (from the *Cyanobacteria* phylum), *Actinomarinales* (from the *Actinobacteria* phylum) and *Pirellulales* (from the *Planctomycete* phylum). The *Actinomarinales* order ranged between 3 – 15% at the surface and 2 – 10 % at 50 m, and was variably detected in different samples throughout the timeseries, being highest on average during summer at the surface ($10 \pm 7\%$) and during winter at 50 m depth ($8 \pm 3\%$).

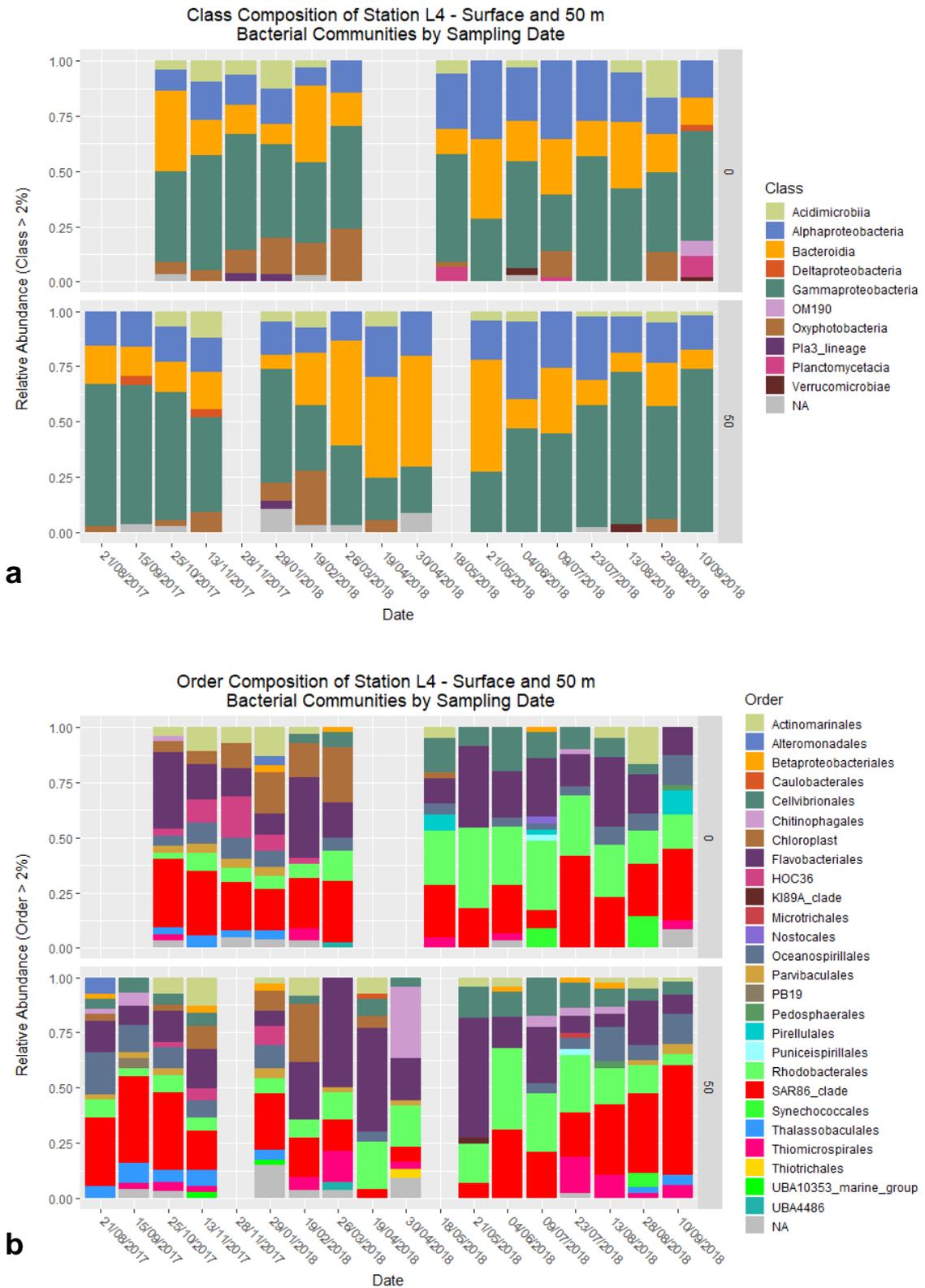


Figure 4.15 Relative abundance (> 2%) of the bacterial 16S *rRNA* community at the surface and 50m depths (depth indicated as a number value on the right hand side of the plots) of station L4 (August 2017 and September 2018), at the (a) Class and (b) Order taxonomic levels.

Surface and 60 m samples from station E1 also indicated that the microbial community mainly consisted of *Proteobacteria*, *Cyanobacteria* and *Bacteroidetes* phyla (Figure 4.16). The *Proteobacteria* accounted for 43 – 67% of the station E1 bacterial community at the surface and 70 – 80% at 60 m depth. Again, this phylum mainly consisted of *Alpha*- and *Gammaproteobacterial* classes, which were evenly abundant throughout the time series (20 – 35% and 19 – 34% at both depths, respectively). These classes appeared to have differing seasonal cycles at the surface, where *Alphaproteobacteria* was more abundant on average during autumn ($31 \pm 4\%$), whereas the *Gammaproteobacteria* were on average more abundant during winter and spring ($33 \pm 2\%$) and both were less abundant during summer months (25 ± 7 and $24 \pm 1\%$, respectively). Although less station E1 samples were available for analysis at 60 m, the five samples analysed indicated that the *Alphaproteobacteria* were on average most abundant during spring ($43 \pm 1\%$) and least abundant during summer (29%), and the *Gammaproteobacteria* were more abundant during summer (52%) and least abundant during spring ($30 \pm 5\%$). The *Alphaproteobacteria* mainly consisted of the SAR11 clade and *Rhodobacterales* at both depths, which constituted 3 – 18 % and 3 – 17% in surface samples, respectively, and displayed statistically significant opposing relative abundances in surface E1 samples ($r = -0.791$, $n = 9$, $P < 0.02$).

The *Gammaproteobacteria* mainly comprised of SAR86, *Thiomicrospirales* and *Cellvibrionales*. SAR86 ranged between 3 – 14% of the microbial community at the surface and 6 – 15% at 60 m, averaging $13 \pm 2\%$ in autumn, winter and spring at the surface and $8 \pm 7\%$ in summer in surface samples. However, at 60 m, SAR86 was at a minimal abundance in May (6%) and then increased to $15 \pm 0.5\%$ in autumn months. *Thiomicrospirales* were another constituent order within the *Gammaproteobacteria* class ranging between 3 – 12% and 4 – 25% of the surface and 60 m community, respectively. The *Cyanobacteria* phylum mainly consisted of *Synechococcales*, which is the same as station L4 samples. *Synechococcales* appeared to make up 44% of the station E1 microbial community at the surface in August, but then only averaged $3 \pm 1\%$ until June when it peaked again at 18%, and then drops again in autumn to 2% of the bacterial community. At station E1, the *Bacteroidetes* only consisted of the *Flavobacteriales* order of bacteria which

represented 3 – 18% of the surface bacterial community and 3 – 14% at 60 m depth, where it was more dominant at the surface during the spring and summer (13 – 14%) compared to the winter and autumn months (3 – 10%), however little variation occurred at 60 m (6 – 9%).

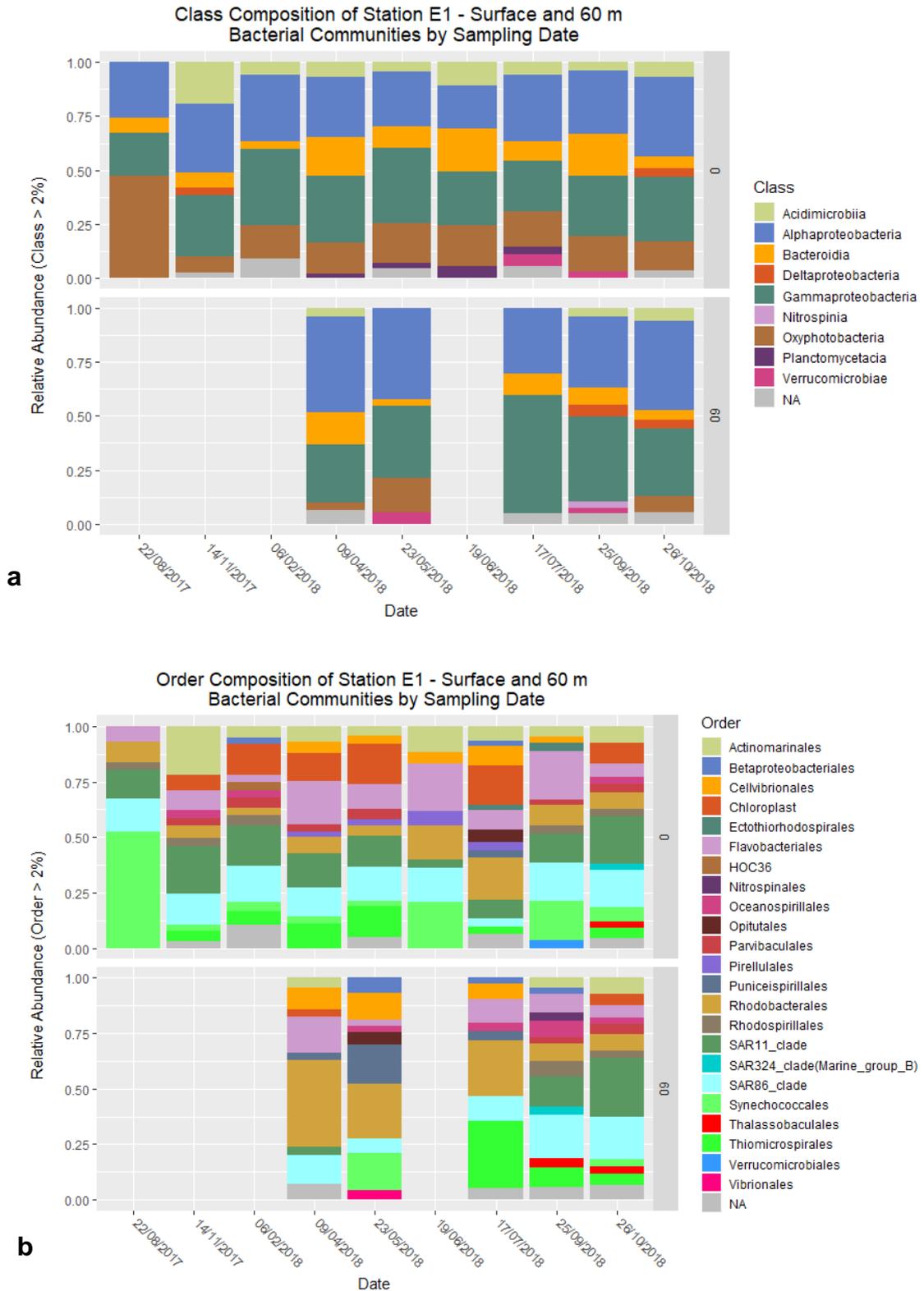


Figure 4.16 Relative abundance (> 2%) of the bacterial 16S *rRNA* community at the surface and 60m depths (depth indicated as a number value on the right hand side of the plots) of station **E1** (August 2017 and October 2018), at the (a) Class and (b) Order taxonomic levels.

4.4.2. *xoxF5* bacterial diversity

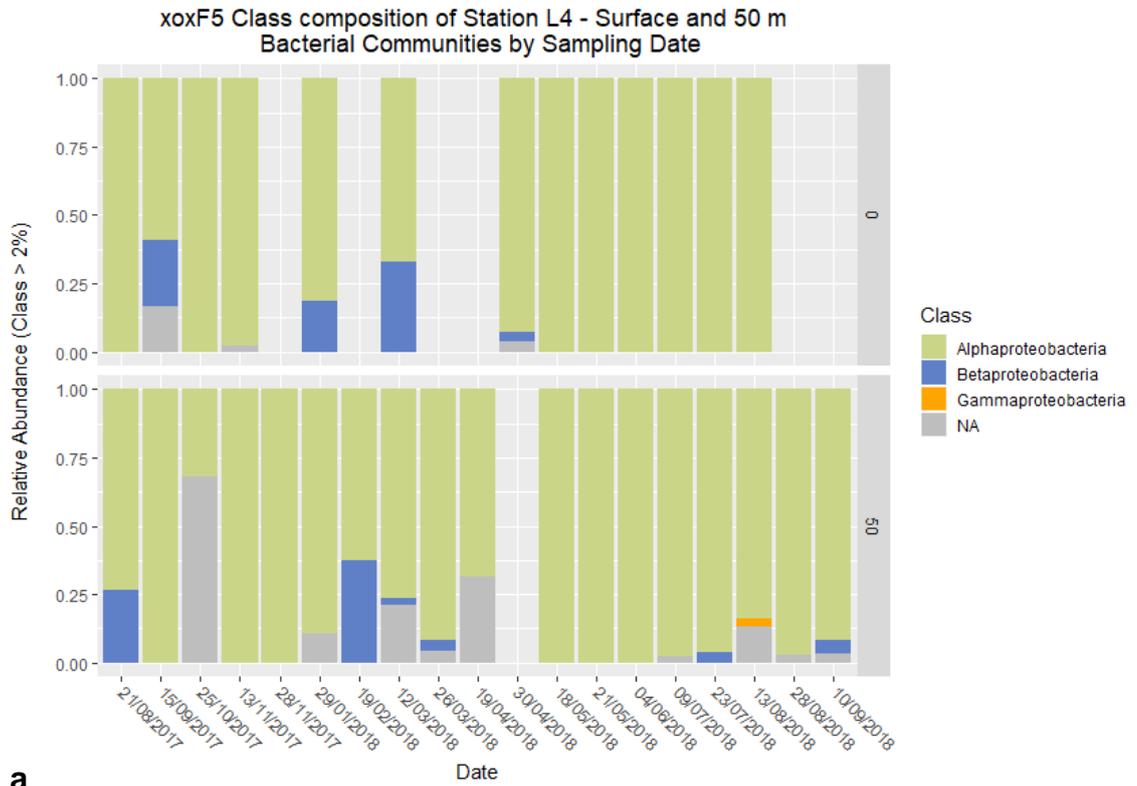
Temporal changes in the diversity of *xoxF5* at the surface and bottom depths of stations L4 and E1 was determined by examining the *xoxF5* gene. Sequencing of *xoxF5* gene amplicons yielded 1,780,753 reads in total from 48 samples after pre-processing steps using the DADA2 pipeline (filtering, merging and removal of chimeric and singleton reads). After these steps were completed, 30% of the total reads were removed (including 291,914 chimeric reads), with 831,792 reads derived from L4 samples and 414,052 reads derived from E1 samples. Reads were then assigned into ASVs, yielding 820 ASVs in total. To mitigate bias resulting from uneven sampling depth, the same approach was taken as with the 16S sequences, which also removed three station L4 samples (770 total reads), and only one sample was sub-sampled to 50,000 reads (from station L4). These ASVs were then annotated using a manually curated *xoxF* database (provided by Martin Taubert), containing 109 *xoxF5* sequences (described further in Section 2.10.8).

At station L4, the *xoxF5* sequences were primarily annotated as *Alphaproteobacteria*, but also included sequences from the *Beta-* and *Gammaproteobacteria*. *Alphaproteobacteria* accounted for 58 – 100% of the assigned *xoxF5* sequences at the surface and 32 – 100% at the bottom of the water column. The *Betaproteobacteria* consisted of 1 – 33% of the *xoxF5* diversity in surface samples and 1 – 37% of the *xoxF5* diversity at 50 m. The *Gammaproteobacteria* only accounted for 1 – 3% at both depths from station L4 samples. The *Alphaproteobacteria* were mainly composed of the *Rhodobacterales*, *Rhodobacteraceae*, which alone made up 15 – 99% of the total assigned *xoxF5* diversity at the surface and 32 – 98% at 50 m depth. *Rhodobacteraceae* exhibited seasonal variation, and on average made up the lowest proportion of the community during autumn ($55 \pm 10\%$) and was then highly dominant during the summer months ($86 \pm 16\%$) at the surface. At the bottom of the water column, a seasonal cycle was also clear, but was instead lowest during winter months ($39 \pm 8\%$) and was also dominant in summer ($87 \pm 12\%$). The *Rhizobiales* order was also a major component of the *xoxF5* diversity, comprising the *Methylobacteriaceae* and *Bradyrhizobiaceae* families. Together these they ranged between 1 – 85% of the

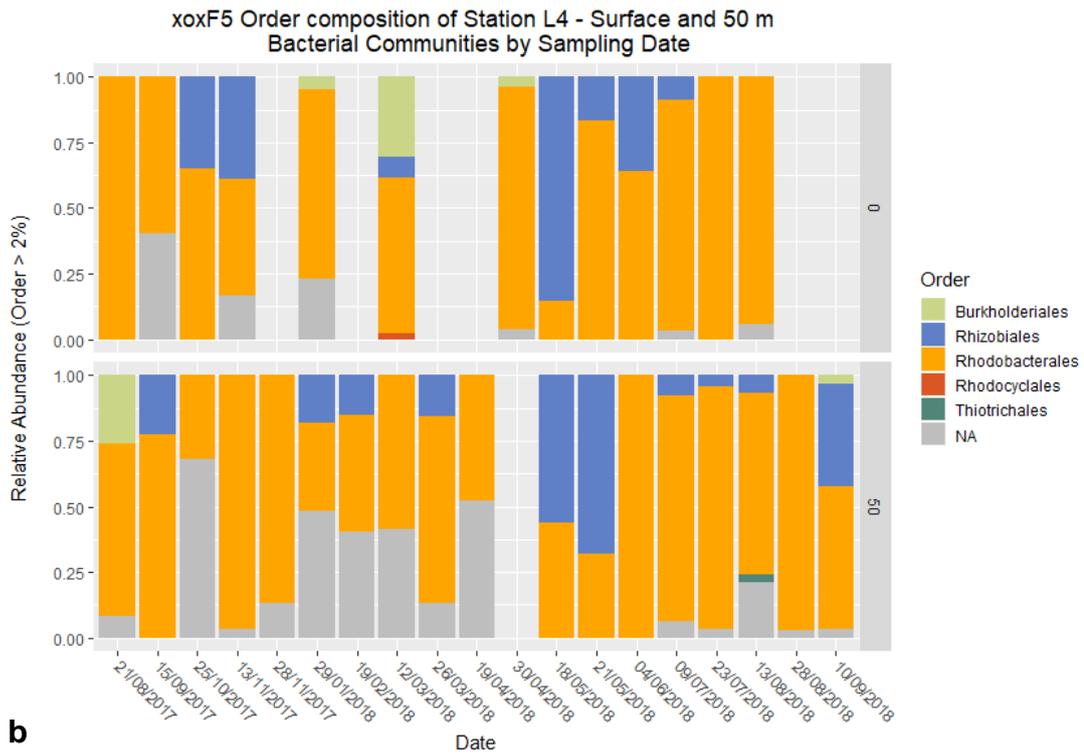
assigned *xoxF5* community at the surface of station L4, and 4 – 68% at 50 m depth. However, the *Rhizobiales* were much more inconsistent throughout the time series making any seasonal cycles hard to distinguish when compared to the *Rhodobacterales* which were identified in all samples. However, they do appear to increase in relative abundance when the *Rhodobacterales* decreases, as evident in autumn and spring. *Methylobacteriaceae* ranged between 1 – 66% and 4 – 68% of the *xoxF5* community at the surface and 50 m, respectively, and the *Bradyrhizobiaceae* ranged between 1 – 19% and 1 – 22% at the surface and 50 m depth. The *Betaproteobacteria* were less abundant and mainly comprised of the *Burkholderiales* order, which constituted between 5 – 29% of the annotated surface *xoxF5* community and 1 – 24% at 50 m. However, again this taxonomic group was variable throughout the time series, and was only identified in four samples, which ranged all seasons.

A similar reflection of the *xoxF5* diversity characterised at station L4 also existed at station E1, where the *Alphaproteobacteria* were again the dominant taxonomic class, ranging between 61 – 97% of the annotated *xoxF5* sequences at the surface and 77 – 99% at 60 m depth. Additionally, the *Betaproteobacteria* were still relatively abundant at the surface comprising between 1 – 10% of the *xoxF5* diversity, and only 1% at 60 m depth. The dominance of the *Alphaproteobacteria* was again driven by the *Rhodobacterales*, *Rhodobacteraceae*, and the *Rhizobiales* families, *Methylobacteriaceae* and *Bradyrhizobiaceae*. *Rhodobacteraceae* again constituted a large percentage of the assigned *xoxF5* diversity, ranging between 29 – 94% at the surface and 41 – 92% at 60 m. However, a seasonal cycle was less obvious in samples from station E1, which indicated that the lowest relative abundance occurred in surface winter samples (29%) and was highest during spring ($93 \pm 1\%$). As a low number of samples from 60 m at station E1 were successfully sequenced, seasonal cycles were difficult to infer. The *Rhizobiales* again constituted an average of $14 \pm 22\%$ to the *xoxF5* surface community and $10 \pm 14\%$ at 60 m, although again any seasonal variation could not be determined due to the variability in relative abundance.

Within some *xoxF5* sample sequences, the unassigned sequences at the class level ranged between 1 – 18% at the surface and 1 – 68% at 50 m depth and averaging $3 \pm 5\%$ for surface samples and $11 \pm 18\%$ for 50 m samples. As would be expected, the number of unassigned sequences increased at lower taxonomic levels, ranging between 1 – 42% for surface samples and 2 – 68% in 50 m samples at the order level, averaging $10 \pm 13\%$ and $20 \pm 21\%$ for surface and 50 m samples. On average, this was the same across *xoxF5* sequences of E1 samples.



a



b

Figure 4.17 Relative abundance (>2%) of the bacterial *xoxF5* community at the surface and 50m depths (depth indicated as a number value on the right hand side of the plots) of station **L4** (August 2017 and September 2018), at the (a) Class and (b) Order taxonomic levels.

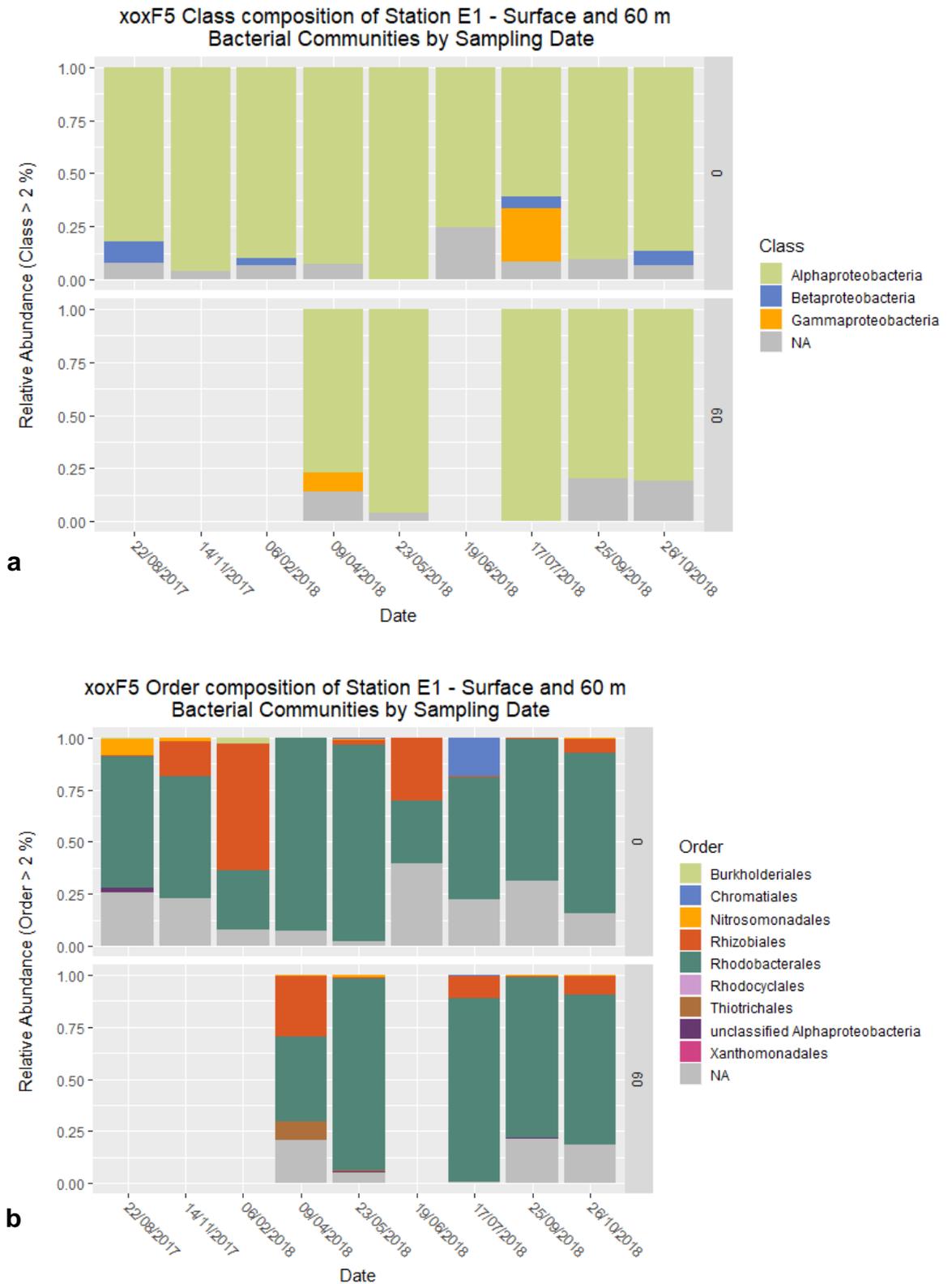


Figure 4.18 Relative abundance (>2%) of the bacterial *xoxF5* community at the surface and 60m depths (depth indicated as a number value on the right hand side of the plots) of station **E1** (August 2017 and September 2018), at the (a) Class and (b) Order taxonomic levels.

4.5. Discussion

4.5.1. *Environmental drivers of methanol utilisation in the Western Channel Observatory*

Between August 2017 and September 2018 methanol uptake rates were determined in seawater samples collected from throughout the water column of station L4, and at the surface and bottom depths of station E1. The methanol dissimilation rates at all sampled depths of station L4 indicated a similar range ($0.0 - 9.3 \text{ nmol L}^{-1} \text{ h}^{-1}$) to the L4 surface range previously reported by Sargeant *et al.* (2016; $0.7 - 11.2 \text{ nmol L}^{-1} \text{ h}^{-1}$), who also indicated seasonal peaks in methanol dissimilation during winter. Methanol dissimilation rates were also determined at station E1 for the first time, which is a more open shelf station within the Western Channel Observatory (WCO; Smyth *et al.*, 2015), where the range in methanol dissimilation rates ($0.1 - 10.6 \text{ nmol L}^{-1} \text{ h}^{-1}$) were also similar to that of Sargeant *et al.* (2016) for L4, and also showed a maxima in winter. Interestingly, higher methanol dissimilation rates in the WCO were detected in the subsurface at both stations when rates were highest in winter, however this variation between the surface and bottom uptake rates was not statistically significant, which follows the findings of Dixon and Nightingale (2012) who found no significant difference with depth in Atlantic samples. Rates of microbial methanol assimilation into biomass were also determined throughout the water column at L4 and for the first time at E1 ($0.1 - 2.9$ and $0.2 - 1.1 \times 10^{-2} \text{ nmol L}^{-1} \text{ h}^{-1}$, respectively), where the uptake range was indicated to be in the same range as surface measurements previously reported for L4 (Sargeant *et al.*, 2016; $0.2 - 2.8 \times 10^{-2} \text{ nmol L}^{-1} \text{ h}^{-1}$). Low methanol assimilation rates have previously been suggested as a result of facultative methylotrophs preferentially utilising other substrates for growth instead of solely relying on methanol (Chistoserdova, 2011; Sargeant *et al.*, 2016). The low methanol dissimilation rates during summer months at both stations was contrasted by higher bacterial production, a trend also noted by Sargeant (2013) in surface samples. This was suggested as a result of methylotrophic bacteria fully exploiting methanol during periods of low bacterial productivity in winter months when other DOC compounds from phytoplankton are

scarce. This is further demonstrated by the higher methanol dissimilation rates at deeper depths where bacterial production is generally lower. Facultative methylotrophs, such as *Methylorubrum extorquens* AM1, have previously been indicated to utilise methanol and other carbon sources under mixed substrate conditions, and to also switch between C1 compounds (such as methanol) and multi-carbon compounds (Skovran *et al.*, 2010; Peyraud *et al.*, 2012). Peyraud *et al.* (2012) indicated that dissimilation and assimilation pathways could be decoupled, where methanol was utilised for dissimilation, and succinate (a multi carbon compound) was primarily used for biomass synthesis. This indicates the metabolic capacity that exists in methylotrophs and could be used by other facultative methylotrophs in the WCO. This indicates a preference for methanol dissimilation and assimilation at different times of the year and could also drive methylotrophic diversity at different depths where this capacity could yield a competitive advantage during periods of lower productivity of the microbial community. This is further represented by the BGE_M, which was on average higher at lower depths ($2.4 \pm 4.4\%$) and during summer months ($2.8 \pm 4.7\%$), although these low percentages still indicate a preference for dissimilation over assimilation of methanol, as has previously been indicated (Sargeant, 2013).

On average, the total heterotrophic bacterial abundance was higher at station L4 compared to E1 (11.6 ± 5.4 and $9.7 \pm 4.6 \times 10^5$ cells mL⁻¹), between August 2017 and September 2018. Seasonal trends were indicated at both stations, with peaks occurring in August and September in 2017 and again in 2018 for the total heterotrophic bacterial abundance (and the two subgroups of nucleic acid-containing bacteria), which is 1-2 months later than that previously reported by Sargeant *et al.* (2016), but similar to a longer time series reported by Tarran and Bruun (2015). The bacterial production values at L4 coincide with the bacterial abundance peak in June in the upper 10 m and subsequent peak in August 2018, and at the surface indicate a similar peak in bacterial production in summer months at L4 as previously found (Sargeant *et al.*, 2016). However, although a peak in bacterial production was also detected at E1 in summer, an additional peak in August was not determined, although this is likely down to the lower sampling

resolution at E1. Interestingly, a peak in bacterial production was observed at 25 m at L4, which does not coincide with an increase in bacterial abundance.

Within the WCO, the nano- and picoplankton population abundances and seasonal trends were similar to previous findings and within their expected ranges, the *Synechococcus* sp. maximum peak in 2018 was more than double that of the maximum range determined over a 7 year time series (Tarran and Bruun, 2015). This difference in *Synechococcus* sp. abundance between the autumn blooms in 2017 and 2018 was statistically significant ($P < 0.05$), and may be a result of the lack of nutrient limitation typically caused by a stronger thermocline at L4 (Figure 4.5). The *Phaeocystis* sp. bloom was driven by nutrient enrichment and more favourable temperatures, and likely consisted mainly of the *P. globosa* species at this latitude (Schoemann *et al.*, 2005). These plankton blooms consist of large gelatinous colonies, which have significant sinking rates, and would explain the increase in abundance at the bottom of the L4 water column at the end of spring as the clumped colonies would sink resulting in higher fluorescence in bottom waters at the end of spring (Figure 4.5c and Figure 4.6c) (Schoemann *et al.*, 2005). Interestingly, the coccolithophore abundances seemed not to indicate an annual September bloom at either station, with peaks in September and October 2017, but not in September 2018. This may be driven by the pronounced thermocline at E1 (Figure 4.6c), which appears to also restrict growth of the plankton groups (except *Phaeocystis* sp.) beneath 20 m in August and September 2018. The balanced nutrient conditions described as the Redfield ratio (16:1) were maintained on average at the surface of L4 during winter (16.2 ± 3.5 , $n = 9$), and decreased with depth (Redfield, Ketchum and Richards, 1963; Obernosterer and Herndl, 1995). Whereas the N:P ratio (Figure 4.8e) at station E1 never reached the Redfield ratio (N:P = 16:1), but was closest during autumn. This ratio is used as an indicator of the average elemental composition of marine organisms, and has been used to assess the physiological status of phytoplankton - where values below this ratio is indicative of growth limitation due to lower nutrient availability (Grob *et al.*, 2013). This ratio is worth consideration if methanol uptake rates are lowest during periods of nutrient limitation. Further highlighting the reduction in methanol utilisation by facultative methylotrophs during summer months when alternate carbon sources are more

readily available due to increased plankton biomass which in turn lower the N:P ratio by utilising and limiting available nutrients.

Within the same sampling period, samples were also used to investigate the virus community within the WCO by determining the abundance of virus-like particles, which have been proven to have vast biogeochemical influences in the ocean via direct and indirect interactions with organisms (See Section 1.3). The total virus-like particle (VLP) abundance from samples in the WCO ranged between 5.5 and 37×10^6 VLPs mL⁻¹, where the total VLP abundance at station L4 was on average twice as abundant as at station E1 (15.6 ± 9 and $7.3 \pm 3.4 \times 10^6$ VLPs mL⁻¹, respectively), which is consistent with previous findings comparing inshore and further offshore waters (Breitbart, 2012). Peaks in total VLP abundance were similar at both stations, occurring between May and September 2018 (Figure 4.3a and Figure 4.4a), which is similar to the findings of Rodriguez *et al.* (2000) in surface waters at L4, who used a transmission electron microscope for direct counts. This range in VLP abundance is similar to a variety of other coastal environments, such as the coasts of Norway, Japan, and in the Bering and Chukchi Seas (Wommack and Colwell, 2000). As the microbial abundance was higher at station L4 than E1, this is likely the best explanation as to why VLP abundance was also higher at L4, as host availability has been shown to typically be the most influential factor on virus abundance (Weinbauer, 2004; Evans *et al.*, 2017). Virus abundance can also be linked with environmental conditions (Mojica and Brussaard, 2014), although this could be an indirect effect because bacterial abundance is driven more by environmental factors, such as temperature, phosphate and silicate concentrations than by trophic interactions (Gilbert *et al.*, 2009). Seasonal trends in total heterotrophic bacterial abundance (and the two subgroups of nucleic acid-containing bacteria) were indicated at both stations, with peaks occurring in August and September in 2017 and again in 2018, which is 1-2 months later than that previously reported by Sargeant *et al.* (2016), but more similar to a longer time series reported by Tarran and Bruun (2015).

Total VLP was further divided into three virus subgroups (VLP1, VLP2 and VLP3), which indicated contrasting seasonal variation throughout the year where VLP1 acted independently of VLP3, and VLP2 shared seasonal aspects with VLP1 and

VLP3. The VLP1 and VLP2 subgroups dominated the total virus abundance at both stations and throughout the water columns (averaging 51 – 61% and 33 – 35%, respectively), and were inversely correlated ($r = -0.539$, $n = 19$, $P < 0.02$), indicating dominance at different times of the year. However, VLP1 was the most abundant virus subgroup in the WCO and closely reflected the seasonal dynamics of the total bacterial abundances ($r = 0.613$, $n = 15$, $P < 0.02$), whereas VLP3 indicated seasonal trends which more closely followed that of silicate concentrations and cryptophyte abundances. These results align with previous findings, where VLP1 is a lower fluorescing group and likely consists of bacteriophage (viruses which infect bacteria), whereas VLP3 is more fluorescent and more associated with phytoplankton (Brussaard, Marie and Bratbak, 2000; Payet and Suttle, 2008b; Evans *et al.*, 2017). Viruses have been isolated which infect cryptophyte cells, and have been indicated to be larger viruses (approx. 203 nm), which would fluoresce more and therefore supports the previous statement regarding the association of VLP3 with phytoplankton (Barone and Naselli-Flores, 2003; Nagasaki *et al.*, 2009). Interestingly, cryptophyte cells (*Rhodomonas salina* CCMP1319) have been shown to produce methanol in culture, and also contain methanol within their cell wall which could be made available to methylotrophs upon cell death, virally mediated or otherwise (Mincer and Aicher, 2016). VLP1 consisting of bacteriophages is further supported by the total virus-to-bacteria ratio (VBR), which is largest during summer months when the VLP1 and bacterial abundances are highest (Evans *et al.*, 2017). Additionally, the VLP1 subgroup indicated a strong significantly positive correlation with bacterial production across all depths and both stations ($r = 0.783$, $n = 63$, $P < 0.001$). This is a known strong relationship that has previously been indicated in multiple studies, and is a clear demonstration of the link between abundant VLP numbers and a productive bacterial community (Steward, Smith and Azam, 1996; Wommack and Colwell, 2000; Thomas *et al.*, 2011). This variability and availability of different hosts likely drives the changes in the virus community composition in the WCO. Collectively, the environmental variables between station L4 and E1 were very similar, indicated by Figure 4.19a which highlights the high degree of similarity between these two stations across the length of the time series period, which was also reflected at the bottom of the water column.

Pearson correlation analysis was carried out to determine any correlations of the methanol uptake rates with any of the environmental variables and abundance values (Table 4-1). Using this approach, methanol dissimilation at station L4 was found to have significant positive correlations with nitrate, silicate and phosphate concentrations, and negative correlations with seawater temperature and bacterial production rates. These findings broadly agree with Sargeant *et al.* (2016), who focussed on surface samples from L4, and go further by investigating the whole water column and indicate that when these nutrients (nitrate, silicate and phosphate) are depleted at all depths during summer months, the methanol dissimilation rates were also at minima throughout the water column. These nutrients are rapidly depleted by the spring bloom and are then a limiting factor leading to its eventual decline (Smyth *et al.*, 2010). As the bacterial production is higher during these warmer periods of nutrient depletion, the marine bacteria are likely utilising another primary substrate for energy other than methanol (Dixon, Beale and Nightingale, 2011). This primary source of carbon is likely dissolved organic matter derived from phytoplankton, still present in the environment having accumulated during the plankton bloom (Giovannoni and Vergin, 2012).

Using the virus abundance data, a significant negative relationship was determined between the VLP1 virus subgroup and microbial methanol dissimilation rates at all depths of station L4, this is important due to the association of the VLP1 virus subgroup with bacteriophage (Table 4-1) (Brussaard, Marie and Bratbak, 2000). However, total VLP abundance indicated a more significant negative correlation with dissimilation rates, and although VLP1 accounts for a significant proportion of the total VLP abundance, when viruses are included out with the VLP1 grouping there is a higher correlation. Collectively these findings indicate that during warmer months, when nutrient concentrations are depleted and bacterial and virus abundance (and bacterial productivity) is higher at station L4, the methanol dissimilation rates were lowest at all depths in the water column

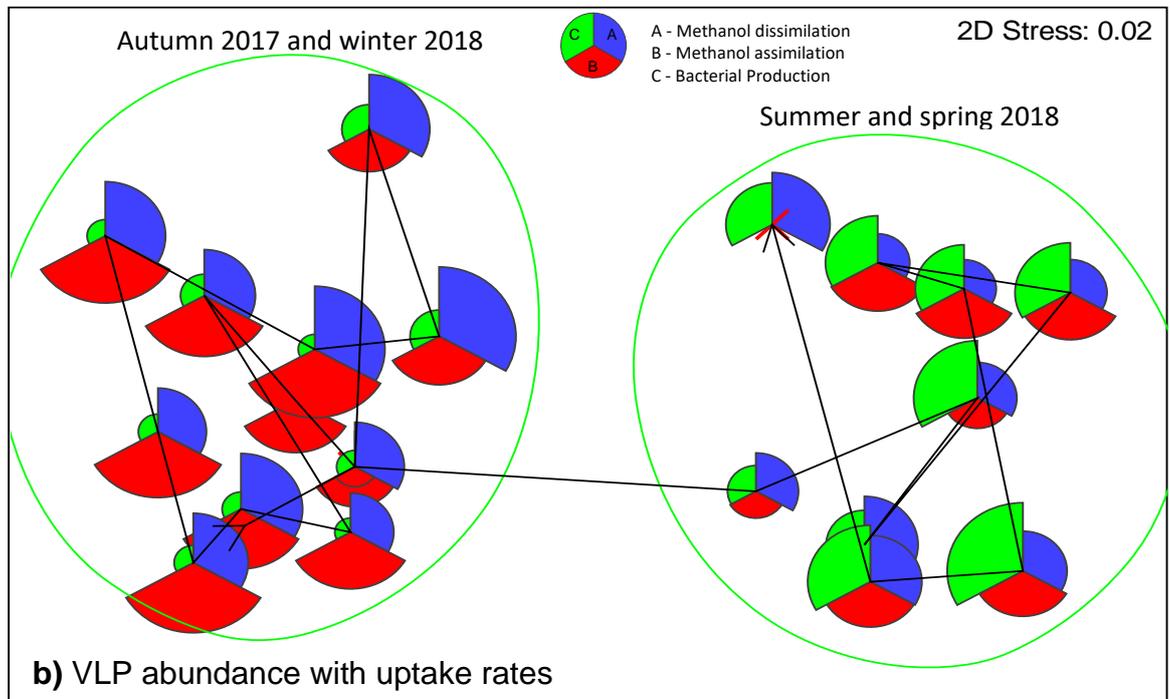
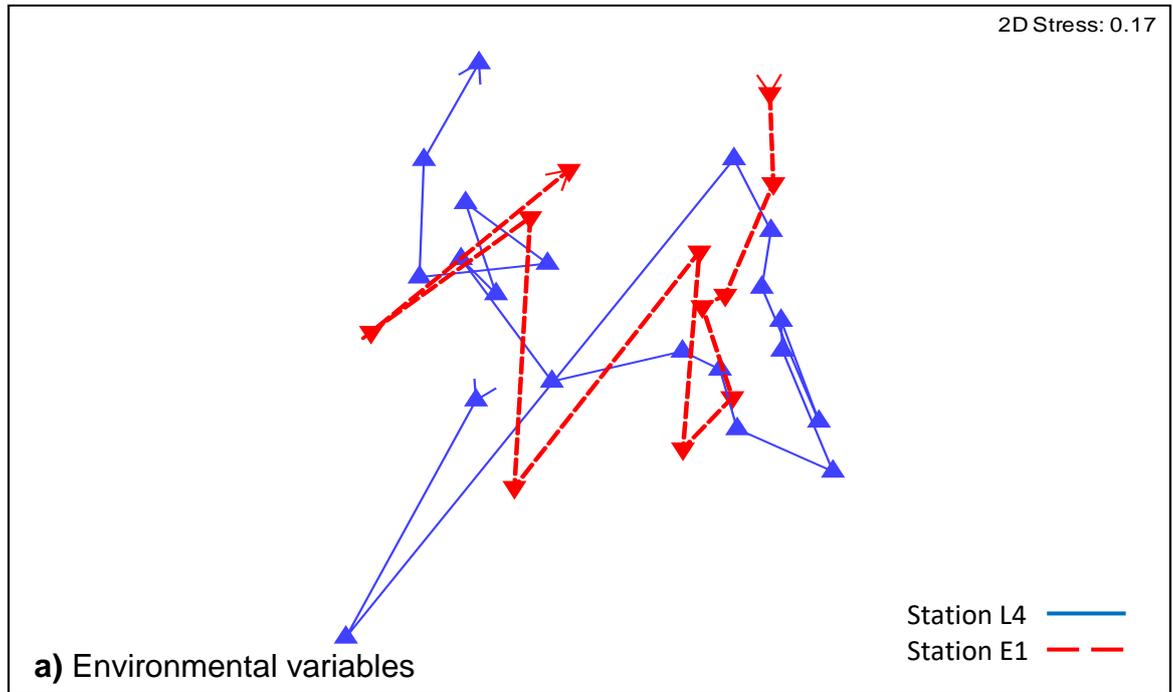


Figure 4.19 Non-metric multidimensional scaling (NMDS) ordination of a) a Euclidean distance resemblance matrix of log transformed surface environmental variables (including VLP subgroups, bacterial abundances, plankton abundances, nutrient concentrations, seawater temperature, salinity and bacterial productivity) from station **L4** (blue) and **E1** (red), and b) surface **L4** abundance data for the total VLP abundances, overlaid with bubble plots of the normalised methanol dissimilation (blue), assimilation (red) and bacterial production (green) rates for each sample.

Table 4-1 Pearson correlation coefficient values between microbial methanol dissimilation and assimilation at station **L4** and **E1** in the WCO with environmental variables. Significant correlations are shown in bold: *P≤0.05, **P≤0.02, ***P≤0.01.

Environmental Variables	Station L4		Station E1	
	Dissimilation	Assimilation	Dissimilation	Assimilation
n	77	41	19	19
VLP 1 (VLPs mL ⁻¹)	-0.519*	-0.420	-0.030	0.085
VLP 2 (VLPs mL ⁻¹)	0.101	-0.062	-0.581*	0.587*
VLP 3 (VLPs mL ⁻¹)	0.430	0.315	-0.535*	0.191
Total VLP (VLPs mL ⁻¹)	-0.609**	-0.433	-0.264	0.266
HNA Bacteria (cells mL ⁻¹)	-0.436	-0.403	-0.171	-0.128
LNA Bacteria (cells mL ⁻¹)	-0.152	0.063	-0.003	-0.539*
HNA + LNA (cells mL ⁻¹)	-0.378	-0.279	-0.164	-0.332
<i>Synechococcus sp.</i> (cells mL ⁻¹)	0.236	0.223	-0.273	0.017
Picoeukaryotes (cells mL ⁻¹)	0.140	-0.017	0.198	-0.324
Nanoeukaryotes (cells mL ⁻¹)	-0.147	-0.264	0.005	0.091
Coccolithophores (cells mL ⁻¹)	-0.274	-0.248	0.212	0.018
Cryptophytes (cells mL ⁻¹)	0.449	0.374	-0.092	-0.169
Total eukaryotes (cells mL ⁻¹)	0.073	-0.063	0.231	-0.231
Nitrite (μM)	0.292	0.246	0.472	-0.390
Nitrate (μM)	0.793***	0.554*	-0.025	-0.332
Ammonia (μM)	-0.039	-0.246	0.464	-0.079
Silicate (μM)	0.629**	0.472	-0.491	0.437
Phosphate (μM)	0.683**	0.605**	-0.030	-0.306
Temperature (°C)	-0.520*	-0.274	-0.475	0.316
Fluorescence (volts)	-0.220	0.094	0.052	0.032
Salinity (PSU)	-0.310	-0.180	-0.163	-0.049
Bacterial Production (ng C L ⁻¹ h ⁻¹)	-0.508*	-0.315	-0.054	0.066

As indicated, dissimilation rates are higher at times of the year when bacterial productivity is lower, potentially due to facultative methylotrophs fully utilising methanol as an energy source, as opposed to other carbon sources. This increased dissimilation of methanol is not a metabolic capability available to the entire microbial community, hence the lower bacterial abundance during winter (Figure 4.11), which would explain the lower total VLP abundance due to lower host availability (Weinbauer, 2004; Evans *et al.*, 2017). This is demonstrated with the

surface data in Figure 4.19b, where the total VLP abundance data clearly separates into two clusters during the time series, where the cluster for autumn 2017 and winter 2018 highlights the higher total methanol uptake rates and lower bacterial production when total VLP abundance is lower. Compared to the summer and spring 2018 cluster where the proportions of total methanol uptake and bacterial production are more proportional, when the total VLP abundance is higher along with bacterial abundance and productivity. This likely indicates that when increased virus pressure is not providing additional organic matter via cell lysis, as would occur in summer and spring months when bacterial productivity is higher, methylotrophic bacteria utilise methanol more efficiently than other carbon sources for energy and growth.

When the same correlations were determined using E1 surface and bottom data, different correlations are evident. Methanol dissimilation had a significant negative correlation with the VLP2 and VLP3 abundances, and assimilation rates positively correlated with VLP2 abundances and negatively with the LNA bacterial abundances. As VLP2 and VLP3 share similar seasonal cycles, then it follows that they would have similar correlation coefficients, although a higher coefficient was determined for VLP2. This appears to suggest that when viruses associated with the VLP2 subgroup have higher abundances during autumn and summer months, the methanol dissimilation rates were lower and assimilation rates were higher at the surface and bottom depths of station E1. Unlike VLP1, the VLP2 virus subgroup is typically more associated with pico-cyanobacteria and small eukaryotes (Payet and Suttle, 2008b; Personnic *et al.*, 2009), which is interesting as cyanobacteria and eukaryotes have been shown to produce methanol, including the abundant marine cyanobacteria *Synechococcus* and *Prochlorococcus* (Mincer and Aicher, 2016). Additionally, cyanobacterial and phytoplankton decomposition has been indicated to produce methanol (and other C1 compounds), which could likely be further facilitated by viral lysis (Hill *et al.*, 1998; Antony *et al.*, 2010; Mincer and Aicher, 2016). Therefore, when VLP2 abundance is higher due to associated increases in relevant host abundances (Figure 4.10a), methylotrophic organisms utilise the additional methanol made available from cell lysis and decomposition, and incorporate increased proportions of methanol into biomass. This seems probable upon consideration of a previous study which indicated the close association of the

methylophilic *OM43* clade of bacteria within a phytoplankton bloom of diatoms via “unknown mechanisms” (Morris, Longnecker and Giovannoni, 2006). Cell lysis could result naturally or from grazing, but is equally likely to occur from virus infection, with virus mortality estimations varying widely from ~14 to 90% in the marine environment (Steward, Smith and Azam, 1996; Winget *et al.*, 2005). Additionally, methanol metabolism within methylophilic bacteria could be directly driven by virus infection, where virus-encoded metabolic genes known as auxiliary metabolic genes (AMGs) alter the metabolic capacity of an infected host. Although no AMGs have been identified which augment methanol metabolism, AMGs involved in photosynthesis, photosynthetic electron transport, nucleotide metabolism, sulphur metabolism, phosphate metabolism and carbon metabolism have been identified (Rohwer *et al.*, 2000; Sullivan *et al.*, 2005, 2010; Millard *et al.*, 2009; Thompson *et al.*, 2011; Breitbart, 2012; Anantharaman *et al.*, 2014). Although it is interesting that the same relationship is not indicated at L4 as well.

To determine what environmental factors best describe the seasonal variability in methanol dissimilation across all depths at Station L4, a ‘BIO-ENV’ analysis was carried out using PRIMER (Section 2.13.2). This suggested that bacterial production, fluorescence, salinity and concentrations of ammonia best explained the temporal variation in methanol dissimilation rates ($r = 0.382$, $n = 84$, $P < 0.1$). The suggestion of bacterial production and fluorescence agrees with the analysis of Sargeant (2013), who found a slightly stronger correlation, but only with surface samples, the similarities and differences between the findings of Sargeant (2013) and this project are indicated in Figure 4.20. When the ‘BIO-ENV’ analysis was repeated with methanol assimilation rates, there were no combination of environmental variables which best described the temporal variation at a significant level of confidence, which is contrary to the findings of Sargeant (2013); who suggested that HNA bacteria, bacterial production, picoeukaryotes and *Synechococcus* sp. best explain the temporal variation in methanol assimilation rates. When the same ‘BIO-ENV’ analysis was carried out with the methanol dissimilation rates and other environmental data at station E1, this indicated ammonia concentrations, salinity and fluorescence as the best drivers at E1, although this was not significant ($r = 0.224$, $n =$, $P < 0.3$). Interestingly the ‘BIO-

ENV' analysis of the methanol assimilation rates at E1 indicated that VLP2, VLP3, LNA bacteria and cryptophyte abundances best described the temporal variation ($r = 0.383$, $n = 19$, $P < 0.4$), where the VLP2 and LNA abundance agrees with the previous correlation findings.

This analysis indicates the seasonal variability of methanol dissimilation to be higher during winter months at both stations in the WCO during periods of low bacterial productivity, when the colder coastal seawater is fresher and nutrient rich from the increased riverine input and water column mixing. Therefore, methanol could be preferentially utilised throughout the water column due to a limitation in dissolved organic matter (originating from phytoplankton) available to methylotrophic bacteria. At this point of the year, fluorescence and ammonia (from bacterial remineralisation) concentrations are also lower, since these factors are driven by the plankton blooms that occur later in the year. Ammonia-oxidation by bacteria has also been indicated to peak during winter months, further reducing the ammonia concentrations in the water column (Tait *et al.*, 2014). When the spring and summer blooms in plankton begin to increase the organic matter available to the water column, there is a subsequent increase in the bacterial abundance and productivity. As stated previously and backed up by Pearson correlations, this bacterial productivity is facilitated by the increased virus pressure (specifically the bacteriophage associated subfraction), which in turn leads to a reduction in methanol dissimilation and a gradual switch by facultative methylotrophic bacteria to more readily available sources of energy.

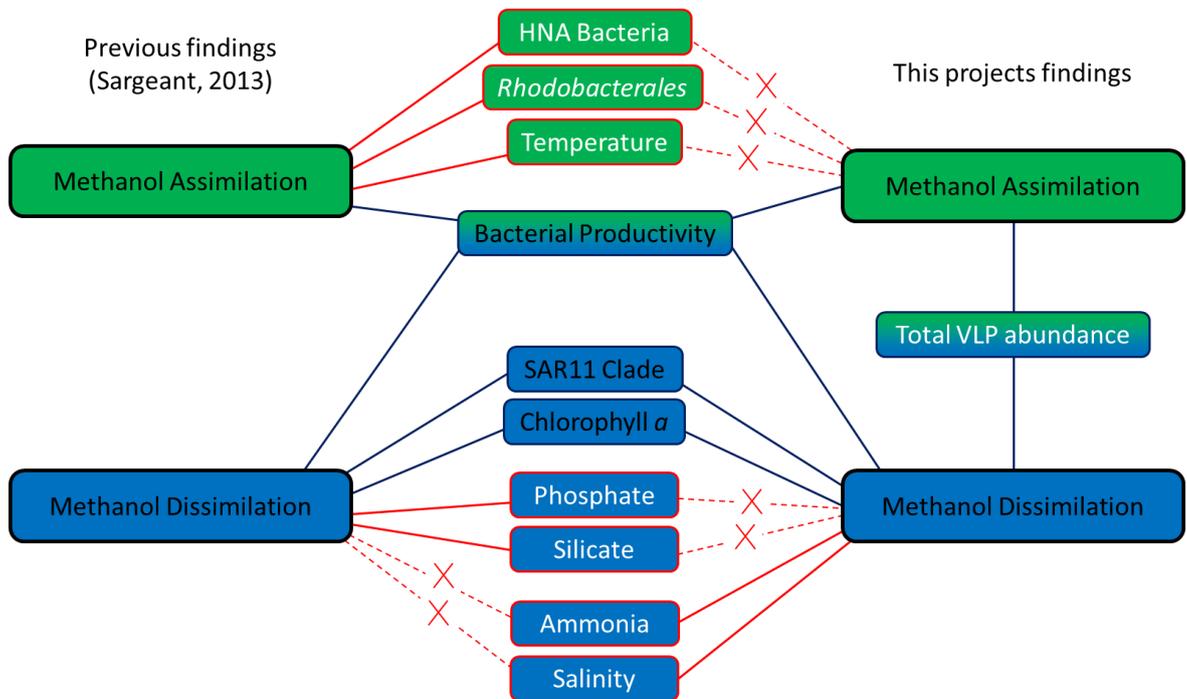


Figure 4.20 Schematic indicating the environmental and abundance factors which best describe methanol assimilation (green) and dissimilation (blue) seasonal variability at station L4 in this and previous work. Connections indicate what factors have been associated with which uptake rates and are coloured if only associated in current or previous work (red) or in current and previous work (blue).

Of note is the variation in correlations between stations L4 and E1, where the same significant correlations are not determined at each station, except for the virus abundance data which indicated significant negative correlations with dissimilation rates at both stations. Given the high significance of correlations at station L4, it is surprising that the same correlations are not evident at station E1, however this can perhaps be best explained by differences in the hydrography of both stations, where L4 is more coastal and subject to greater terrestrial influences from riverine input, whereas E1 is more consistent with an open shelf location (Smyth *et al.*, 2010). Additionally, less samples were taken from station E1 due to the greater distance and weather dependability of successful sampling from this station (Smyth *et al.*, 2015). Correlations were in some cases less significant than that found by Sargeant (2013), although this is also likely a direct result of the increased number of samples, which was much lower in the previous study (13 – 20 samples), due to only investigating the surface water of station L4.

4.5.2. Microbial community in the Western Channel Observatory

The *Alpha*-, *Gammaproteobacteria*, *Bacteroidia* and *Oxyphotobacteria* were the only classes to be detected at all sequence sample dates and depths of both stations. Station L4 and E1 shared a lot of the same taxonomic classes, however the variation came in the relative abundances between the stations. On average, the *Gammaproteobacteria* dominated station L4 at the surface ($42 \pm 9\%$) and bottom of the water column ($45 \pm 16\%$), whereas it was less dominant at station E1 where it constituted $27 \pm 5\%$ of the surface bacterial community and $36 \pm 10\%$ at 60 m depth. Instead it was the *Alphaproteobacteria* which was also dominant at station E1 ($28 \pm 4\%$ and $37 \pm 7\%$ at the surface and 60 m, respectively), when compared to station L4 ($19 \pm 8\%$ and $18 \pm 6\%$ at the surface and 50 m, respectively), this dominance of the *Alphaproteobacteria* and *Gammaproteobacteria* within the WCO reflects that found by Gilbert *et al.* (2009). The dominance of the *Alphaproteobacteria* at E1 is driven by the increased relative abundance of the SAR11 clade, which was a major order within the E1 samples, compared to L4. The SAR11 clade has been widely reported in the marine environment (Giovannoni *et al.*, 2008), and has previously been indicated to have a significant positive relationship with methanol dissimilation rates (Sargeant *et al.*, 2016), therefore it is surprising that in this study SAR11 sequences were not as abundant in L4 samples. The *Alphaproteobacteria* mainly consisted of the SAR11 clade and *Rhodobacterales* at both depths and displayed statistically significant opposing cycles of dominance in surface E1 samples ($r = -0.791$, $n = 9$, $P < 0.02$). The opposing cycles of SAR11 and *Rhodobacteraceae* has been demonstrated previously by Gilbert *et al.* (2012), and in that same study it was stated that this family is dominated by *Roseobacter* in the WCO; however, our analysis instead indicates that it mainly consists of the *Amylibacter* and *Ascidiahabitans* genera.

The *coxF5* gene sequences detected in samples from L4 and E1 are overwhelmingly *Alphaproteobacterial* sequences ($\sim 85 - 91\%$ of the total *coxF5* relative abundance) and predominantly align with the *Rhodobacteraceae* family at both stations and depths ($\sim 64 - 74\%$), which are known to contain multiple methylophs and reflects the findings of Taubert *et al.* (2015). The

Alphaproteobacteria class also mainly consisted of the *Methylobacteriaceae* and *Bradyrhizobiaceae* families within the *Rhizobiales* order, of which the *Methylobacteriaceae* was the more abundant taxonomic family within *Rhizobiales*, at both stations and depths in the WCO. *Gammaproteobacterial* sequences were also detected and mainly consisted of the *Chromatiaceae* family at the surface of E1, and the *Piscirickettsiaceae* family at 60 m of E1. The *Piscirickettsiaceae* sequences aligned with *Methylophaga thiooxydans*, which belongs to the highly researched methylophilic *Methylophaga* genus (Janvier *et al.*, 1985; Urakami and Komagata, 1987; Villeneuve *et al.*, 2012; Boden, 2019). Interestingly, *Burkholderiales* increased in relative abundance in March, when there was a corresponding input of riverine freshwater at the surface, which may be driving a change in the *Rhizobiales*.

Non-assigned sequences at the order taxonomic level were highest at the bottom of the L4 water column (~19%), and also higher at the surface of the E1 water column (~19%) (Figure 4.17b and Figure 4.18b). This indicates that using a manually curated sequence database has not characterised the majority of the *xoxF5* diversity and further implies that there are still a significant number of unidentified organisms with the *xoxF5* gene, which was similarly stated by Taubert *et al.* (2015). This is often the case with functional genes due to a lack of reference sequences, and for *xoxF* genes in particular the presence of multiple *xoxF* copies in many organisms compounds this issue, as the copies are often less closely related to one another in one organism than to other corresponding copies in a different organism, making assignment difficult (Taubert, *pers. comm.*). This could result from the extensive horizontal gene transfer which has been noted in some methylophilic genomes, such as the *Methylobacterium* and the *Methylophilaceae* (Vuilleumier *et al.*, 2009; Salcher *et al.*, 2019). This is an interesting note, due to the association of horizontal gene transfer with viruses, indicating another potential link for further exploration between methylophilic organisms and viruses (Breitbart, 2012; Breitbart *et al.*, 2018). This could be investigated further to determine if viruses are potentially mediating horizontal gene transfer in methylophilic organisms and to what degree.

Pearson correlations were conducted between methanol uptake rates and bacterial taxonomic groups identified within 16S sequences which have been associated with methylotrophy, to link the active metabolism of methanol to different methylotrophic taxa. Using this approach, no significant correlations could be determined utilising the combined relative abundances from both stations, and this was also the case with the methanol uptake rates at station L4. However, the methanol dissimilation rates at station E1 correlated positively with *Rhodobacterales* ($r = 0.597$, $n = 21$, $P < 0.01$), *Flavobacteriales* ($r = 0.542$, $n = 21$, $P < 0.01$), *Alphaproteobacteria* ($r = 0.552$, $n = 21$, $P < 0.01$) and *Bacteroidia* ($r = 0.536$, $n = 21$, $P < 0.01$). The potential link with the *Bacteroidia* is interesting, as this order was more relatively abundant in the station L4 samples at both depths than when compared to E1 samples, although no correlation was determined with the L4 data. At the bottom of the water column of station L4, the *Bacteroidia* and *Gammaproteobacteria* had statistically significant opposing relative abundances ($r = -0.834$, $n = 16$, $P < 0.01$), when the *Bacteroidia* were most abundant in spring. Methylotrophy is reportedly widespread in the *Rhodobacterales* order, with accounts of methanol utilisation in multiple *Rhodobacteraceae* genera, such as *Paracoccus*, *Rhodobacter* and *Sagittula* (Kalyuzhnaya, Lapidus, *et al.*, 2008; Chistoserdova, Kalyuzhnaya and Lidstrom, 2009; Kolb, 2009; Schlueter *et al.*, 2015; Taubert *et al.*, 2015). Conversely, there are few indications of *Flavobacteriales* isolates having methylotrophic capabilities, and reports of methanol utilisation do not exist for *Flavobacteriales*, typically more associated with utilising organic matter released during phytoplankton blooms (Gómez-Consarnau *et al.*, 2012; Kolb and Stacheter, 2013; Chistoserdova and Kalyuzhnaya, 2018). *Rhodobacterales* and *Flavobacteriales* are consistently linked with phytoplankton blooms, both using non-streamlined genomes with wider metabolic arrays to respond to the pulses of nutrients and organic matter associated with bloom events (Buchan *et al.*, 2014; Giovannoni, Thrash and Temperton, 2014). This would appear to indicate that the methanol dissimilation rates can be more attributed to the *Rhodobacterales* than the *Flavobacteriales*, increasing in relative abundance when methanol uptake is higher.

Additionally, the methanol assimilation rates at station E1 indicated significant positive correlations with *Thiotrichales* relative abundance ($r = 0.524$, $n = 21$, $P <$

0.01) implicating this bacterial group with methanol assimilation rates. Dissimilation rates have been shown to positively correlate with grouped 16S sequences which included *Thiotrichales* sequences (Sargeant, 2013), but this relationship has not previously been shown independently from other sequences. This is interesting, as *Thiotrichales* is the order which the methylophilic genus, *Methylophilus*, belongs to and has been implicated in methanol assimilation in the marine environment via DNA-SIP experiments (Neufeld et al., 2007; Boden, 2019). As certain taxonomic groups of bacteria correlated with methanol uptake rates at one station, but not at another, this potentially indicates a decoupling of the methylophilic community between these two stations, which is also indicated by the differences in SAR11 and OM43 relative abundance at both stations.

In addition to the correlations with the 16S relative abundances, the *xoxF5* relative abundances were also compared with the methanol uptake rates. This included *xoxF5* sequences of *Rhodobacterales*, *Methylobacteriaceae*, *Thiotrichales*, *Bradyrhizobiaceae*, *Rhizobiales*, *Nitrosomonadales* and *Alphaproteobacteria*, which have all been indicated to be involved in methanol metabolism (Chen, 2012; Dixon and P D Nightingale, 2012; Kolb and Stacheter, 2013; Boden, 2019). However, only *Thiotrichales* indicated any correlation, which was positive with methanol dissimilation rates at stations L4 and E1, although this was not statistically significant. As *xoxF5* has been shown to be essential for methanol oxidation in some methylophilic but has also been detected in non-methylophilic, this lack of correlation again highlights a disconnect and the mostly unknown role that *xoxF*-type methanol dehydrogenases may play in the marine microbial community (Taubert et al., 2015; Howat, 2017; Chistoserdova, 2019). Additionally, none of the VLP subgroup abundances indicated correlations with any of the *xoxF5* bacterial relative abundances.

To determine what microbial taxa best described the seasonal variability in methanol dissimilation in the WCO; a 'BV-STEP' analysis was carried out using PRIMER (Section 2.13.2). Methanol dissimilation rates were best explained by the SAR11 16S sequences combined with the *xoxF5* sequences of the *Rhodobacterales* ($r = 0.361$, $P < 0.3$). The implication of a link between SAR11 and methanol dissimilation aligns with the findings of Sargeant et al. (2016), however these findings were not

statistically significant, perhaps due to the increased sample size. It is perhaps also interesting that *Rhodobacterales* was also linked along with SAR11 to best describe the seasonality of the methanol dissimilation rates, as Sargeant *et al.* (2016) stated that *Rhodobacterales* aligned closer with methanol assimilation rates, although not to any statistically significant extent. Although, both SAR11 and multiple genera belonging to *Rhodobacterales* have been attributed to methanol oxidation, this finding is likely significant and further implicates the SAR11 clade with methanol dissimilation rates (Kalyuzhnaya, Lapidus, *et al.*, 2008; Sun *et al.*, 2011; Schlueter *et al.*, 2015).

The same analysis was also carried out using the methanol assimilation rates, which were best explained by the relative abundances of *Thiotrichales* 16S sequences, combined with the *Bradyrhizobiaceae* and *Nitrosomonadaceae* *xoxF5* sequences ($r = 0.398$, $P < 0.1$). Although these were the correlations which best described the methanol assimilation rates, these were again statistically insignificant. The *Thiotrichales* group of bacteria have frequently been associated with methanol oxidation, and the genus *Methylophaga* includes many of the type strains used in much of the fundamental research into marine methylotrophy, as do some members of the *Bradyrhizobiaceae* family have the capacity to oxidise methanol (Fitriyanto *et al.*, 2011; Boden, 2019). The *Nitrosomonadaceae* bacterial family has been associated with methylotrophy, but not with methanol oxidation in the past, this could be an indication that the *xoxF5* containing *Nitrosomonadaceae* fraction of the microbial community can assimilate methanol (Nercessian *et al.*, 2005).

It is interesting that between station L4 and E1, a main difference between the two stations is defined by slightly higher relative proportions of the *Alphaproteobacteria* and the *Betaproteobacteriales* order within the *xoxF5* sequences. As station L4 is subject to greater riverine input, this likely results in increased concentrations of lanthanum (and other rare earth elements) at L4 compared to E1 (Goldstein and Jacobsen, 1988; Sholkovitz, 1993). This may be important, as *xoxF5* has been determined to be upregulated under greater concentrations of lanthanum by *Methylorubrum extorquens*, which belongs to the *Methylobacteriaceae* family of the *Alphaproteobacteria* (Vu *et al.*, 2016). This indicates that greater lanthanum

concentrations could be affecting the *xoxF5* diversity at L4 but is limited to less relatively abundant taxonomies, suggesting that upregulation of *xoxF5* under increased lanthanum concentrations is a limited metabolic capability of the *xoxF5* community.

A notable observation of this study was the low abundance of SAR11 clade sequences detectable in samples from station L4, which was comparatively lower than previous indications at the same sampling site (Sargeant, 2013; Sargeant *et al.*, 2016), but does reflect findings (<1%) at coastal sites along the California coastline (Cottrell and Kirchman, 2000). These observed differences at L4 could be driven by different sequencing approaches (i.e. 454 pyrosequencing), which can be differentially sensitive to various taxa and is a recognized fundamental problem (McLaren, Willis and Callahan, 2019). However, the relative abundance of SAR11 at station E1 was more comparable with previous findings (Sargeant *et al.*, 2016), suggesting that the observed variation does not originate from differences in methodological approach. Furthermore, SAR11 clone libraries have previously indicated variability between coastal sites ranging from 1 – 12%, which is a similar range in variation as presented here between stations L4 and E1 (Cottrell and Kirchman, 2000). A significant positive correlation between methanol dissimilation rates and SAR11 sequences at L4 has been highlighted previously by Sargeant *et al.* (2016). However, as a similar range and seasonal trend in dissimilation rates was indicated in this study where the SAR11 was less abundant, this suggests that SAR11 may not be as significant to methanol dissimilation rates as previously proposed. This is further highlighted by the disparity in SAR11 between stations L4 and E1, where little difference in the range or seasonal trends of methanol dissimilation rates was determined between the two stations.

At station L4, the SAR86 clade averaged $22 \pm 8\%$ of the microbial community throughout the water column, this prevalence drops at station E1, but still accounts for $\sim 12 \pm 3\%$ of the community. Additionally, the SAR11 bacterial clade is found to contribute a similar proportion as the SAR86 clade in the E1 water column, and peaks in winter as with previous surveys (Gilbert *et al.*, 2012; Sargeant, 2013). These two bacterial clades have been indicated to share an ecological niche, where

SAR86 utilises more complex carbon, whilst SAR11 exploits simpler carbon compounds (Dupont *et al.*, 2012; West *et al.*, 2016). This would suggest that at station L4, the SAR86 clade is outcompeting SAR11 or that SAR11 growth is inhibited, whereas at station E1 these bacterial clades are in a state of equilibrium, which would occur in oligotrophic environments (Giovannoni, Thrash and Temperton, 2014; West *et al.*, 2016). This could be explained by a larger concentration of more complex carbons available to the SAR86 clade at L4. Alternatively, Dupont *et al.* (2012) suggested that the presence of nitroaromatics could select for SAR86 over SAR11. These compounds can be introduced by anthropogenic sources via riverine input and are also naturally occurring in marine environments (Ju and Parales, 2010). Additionally, these observations could result from a weaker stratification at station L4 in 2018 than at E1 (Figure 4.5a), where a weaker stratification has been shown to influence SAR11 abundance relative to SAR86 at other time series stations (Giovannoni and Vergin, 2012).

4.6. Conclusion

This seasonal study has yielded a more comprehensive understanding of the coastal seasonal dynamics of microbial methanol uptake rates at the western channel observatory. Virus-like particle abundances were also determined, and a statistically significant negative relationship existed with the seasonal changes of methanol uptake rates, specifically the methanol dissimilation rates. This strengthens the hypothesis that microbial methanol uptake rates will decrease during periods of increased VLP abundance, although the direct mechanisms by which a virus infection will affect a methanol-utilising methylotroph remains unknown due to the shortcomings of Chapter 3. However, this research does highlight the direct role of viruses upon facultative methylotroph metabolism and how viruses likely drive a shift from methanol to another carbon source by facilitating the release of organic matter. This is an important step towards understanding the unknown sources of methanol in the oceans and further emphasises the importance of viruses within the marine environment. This finding also potentially indicates that facultative

methylophils, are the primary drivers of seasonal variation in methanol dissimilation rates in seawater.

This research project indicates that the changes in methanol dissimilation are best explained by the SAR11 clade and the *Rhodobacterales* *soxB5* bacterial group. This agrees with past findings of SAR11 and methanol dissimilation, and further implicates the *soxB5* gene with methanol oxidation as many *Rhodobacterales* species have been determined to oxidise methanol. Alternatively, this casts doubt over previous assertions that the *Rhodobacterales* is associated with methanol assimilation rates. These are instead best explained by the bacterial orders: *Thiotrichales* and the *Bradyrhizobiaceae*; with indications that *Nitrosomonadaceae* is also involved, which has not previously been indicated to assimilate methanol. Collectively, this implies the involvement of different taxonomic groups with methanol dissimilation and assimilation rates.

This is the first use of amplicon sequence variants (ASVs) with sequencing data derived from the WCO and is an important step away from the more limiting operational taxonomic units used in past microbial community assessments of the English Channel. However, this approach was not used to the fullest potential and with more resources, the interrogation of the individual ASVs alongside environmental and abundance records would likely yield greater insight. Furthermore, *soxB5* covers a broader taxonomic range, but is only one clade of the *soxB* gene, with further homologs being associated with various proportions of the microbial community (i.e. *soxB1* - *Xanthomonas* and *Beijerinckia*, *soxB2* - *Verrucomicrobia*, *soxB3* - *Rhizobiales* and *soxB4* - *Methylophilaceae*). Therefore, a future study could take a more holistic approach and incorporate all five gene homologs along with a quantitative PCR approach to gain further insight to *soxB* gene expression in the water column.

Chapter 5

Basin scale variability of microbial and *xoxF5* diversity

5. Basin scale variability of microbial and *xoxF5* diversity

5.1. Introduction

Large scale variability in methanol uptake rates have been shown between the different geographic basins of the Atlantic Ocean (Figure 5.1), indicating significantly higher methanol dissimilation rates in the Northern Temperate (NT) and Northern Subtropical Gyre (NSG) regions (0.69 ± 0.35 and 0.99 ± 0.41 nmol L⁻¹ h⁻¹, respectively) than in the southern latitudes (averaging 0.18 ± 0.06 nmol L⁻¹ h⁻¹; Sargeant *et al.*, 2018). Additionally, methanol assimilation rates and bacterial productivity were found to be higher in the Equatorial upwelling (EQU) regions, and lower in the northern and southern Atlantic gyres (Sargeant, 2013; Sargeant *et al.*, 2018). Methanol concentrations were also found to vary drastically in the Atlantic, being lowest in the NT and Southern Gyre (SG; 110 ± 80 nM), and highest in the NSG and Northern Tropical Gyre (NTG) provinces (198 ± 42 nM; Beale *et al.*, 2013; Sargeant, 2013). Although the surface bacterial communities at these locations indicated little variability in composition between regions, they did allude to the potential significance of the SAR11 clade and *Deltaproteobacteria*, which positively correlated with methanol dissimilation and assimilation rates, respectively (Sargeant, 2013; Sargeant *et al.*, 2018).

Functionally, the *mxoF* gene has yielded much about the diversity of methylotrophic bacteria in the marine environment, however this primer set has been shown to have limitations and does not successfully characterise the methylotrophic community, indicated by the SAR11 genome which does not contain the *mxoF* gene, but is linked with methanol dissimilation rates (Giovannoni *et al.*, 2008; Sun *et al.*, 2011; Sargeant *et al.*, 2016, 2018). Alternatively, interest in the rare earth element-dependent *xoxF* genes has increased rapidly and they have been shown to be essential for methanol utilisation in some methylotrophs, and have also been implicated in the genomes of non-methylotrophs, raising further questions as to their function (Pol *et al.*, 2014; Chistoserdova, 2016; Howat *et al.*, 2019; Huang *et al.*, 2019). XoxF has been phylogenetically classified into five major clades, and of particular interest is the *xoxF5* clade, which comprises the majority of known *xoxF* sequences (Taubert *et al.*, 2015). Due to its presence in a wide range of

Proteobacteria orders, one of the dominant bacterial phyla in the marine environment, it therefore has as yet unknown diversity and distribution throughout the marine environment (Taubert *et al.*, 2015; Chistoserdova, 2019; Huang *et al.*, 2019).

Chapter 4 investigated *16S rRNA (V4)* and *xoxF5* diversity at two stations in the Western Channel Observatory (WCO) between August 2017 and September 2018. This gave further insights to the temporal diversity of the *xoxF5* gene in the coastal marine environment and at depth, but very little was gained regarding the spatial diversity of this gene. This chapter attempts to clarify this point by using surface samples from a research cruise along the Atlantic Meridional Transect (AMT) in 2009 (AMT cruise number 19, JC039, RRS James Cook) and additional samples from a cruise to the Barents Sea in the Arctic in 2017 (JR16006, RRS James Clark Ross) to give a holistic overview of *16S* and *xoxF5* diversity in surface waters of Atlantic and Arctic ocean regions (Figure 5.1). An assortment of physicochemical and abundance measurements is also presented for the Arctic Ocean cruise transects, this includes salinity, temperature, fluorometry defined chlorophyll-*a*, total bacterial abundance and nutrients (ammonium, phosphate, silicate, nitrate and nitrite). The salinity, temperature and nutrient values are important in the Arctic Ocean for driving changes in the microbial community and influences community structure, especially in the shelf and coastal regions of the Barents Sea where multiple currents exchange and mix. Whereas the chlorophyll-*a* and bacterial abundance measurements are more holistic indicators of regional productivity.

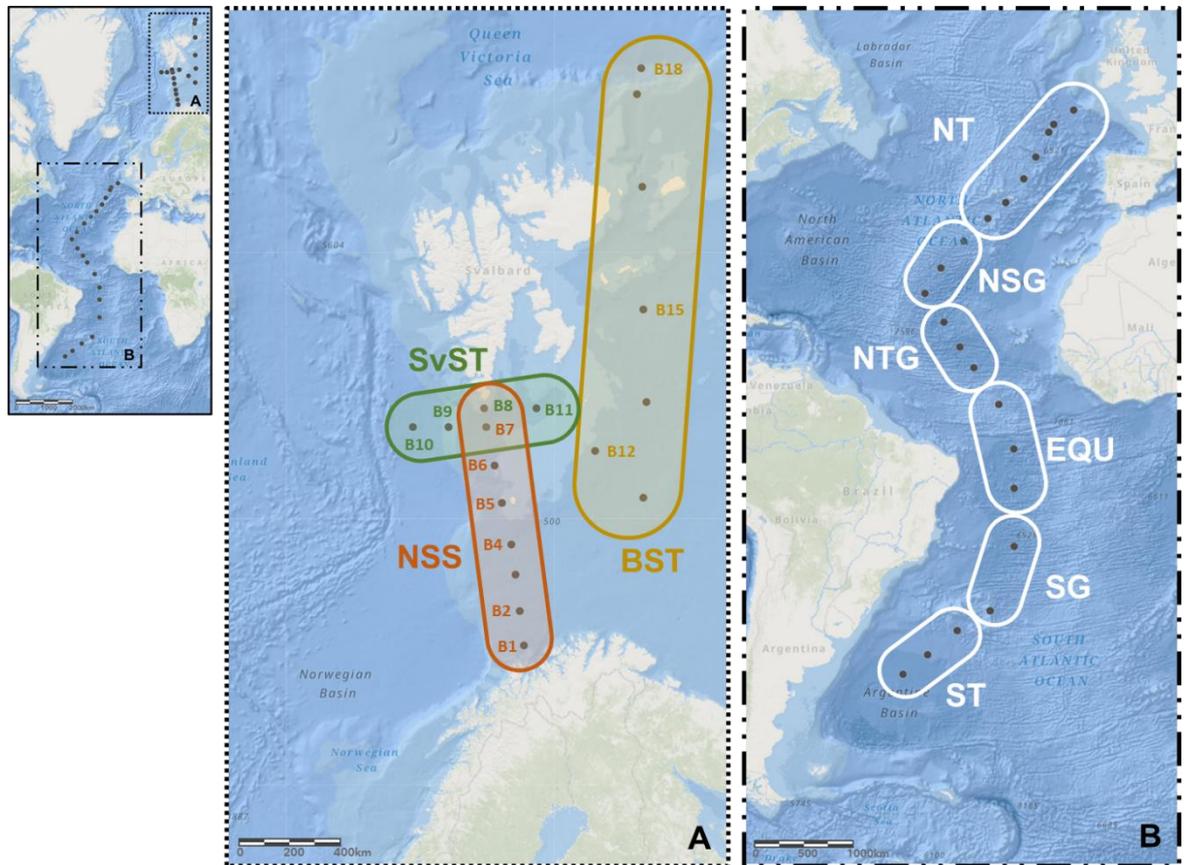


Figure 5.1 A map of A) the sampling station locations of JR16006 in the Arctic Ocean (black circles) overlaid with three regions and B) the sampling points from the JC039 cruise (black circles) and overlaid with the Atlantic regions. Atlantic province names: NT (northern temperate), NSG (northern subtropical gyre), NTG (northern tropical gyre), EQU (equatorial upwelling), SG (southern gyre) and ST (southern temperate). Arctic regions: Norway to Svalbard shelf (NSS; red), Svalbard shelf transect (SvST; green) and the Barents Shelf transect (BST; yellow). Produced using ArcGIS Online (Available at <https://www.esri.com>).

5.2. Results

5.2.1. *Arctic biogeochemical properties*

The seawater surface temperature (SST; Figure 5.2a) ranged between 10.3 and -1.5°C, and averaged 4.9 ± 3.8 °C, where the warmest section of the cruise transect was along the Norway to Svalbard Shelf (NSS) averaging 7.5 ± 2.4 °C and the coldest section was the Barents Shelf Transect (BST) where the SST averaged 1.5 ± 3.5 °C. The coldest temperatures occurred at the most northerly latitudes of the sampling cruise at stations B16 – B18. Chlorophyll-a concentrations (Figure 5.2b) were determined to range between 1.7 and 0.1 mg m⁻³ and shared an average of 0.5 ± 0.4 mg m⁻³ across all cruise sections. The highest chlorophyll-a concentration occurred at station B17, and the lowest at station B15, both within the BST, indicating a highly variable section. Measurements of seawater salinity (Figure 5.2c) did not vary greatly and indicated an average of 34.3 ± 0.8 PSU; however, it was noticed that where the SST was lowest was also where the salinity was lower, averaging 33 ± 0.5 PSU. Total bacterial abundance (Figure 5.2d) was also determined for each of the sampled stations and revealed that bacterial abundance varies from 20 – 5.1×10^5 cells mL⁻¹ and averages $9.9 - 3.6 \times 10^5$ cells mL⁻¹ across all sampled stations. The most abundant bacterial concentration was determined at station B2 and the lowest at B15. On average the BST had the lowest average bacterial abundance ($8.8 \pm 3.1 \times 10^5$ cells mL⁻¹), and the highest average abundance was the Svalbard Shelf Transect (SvST; $11 \pm 0.9 \times 10^5$ cells mL⁻¹).

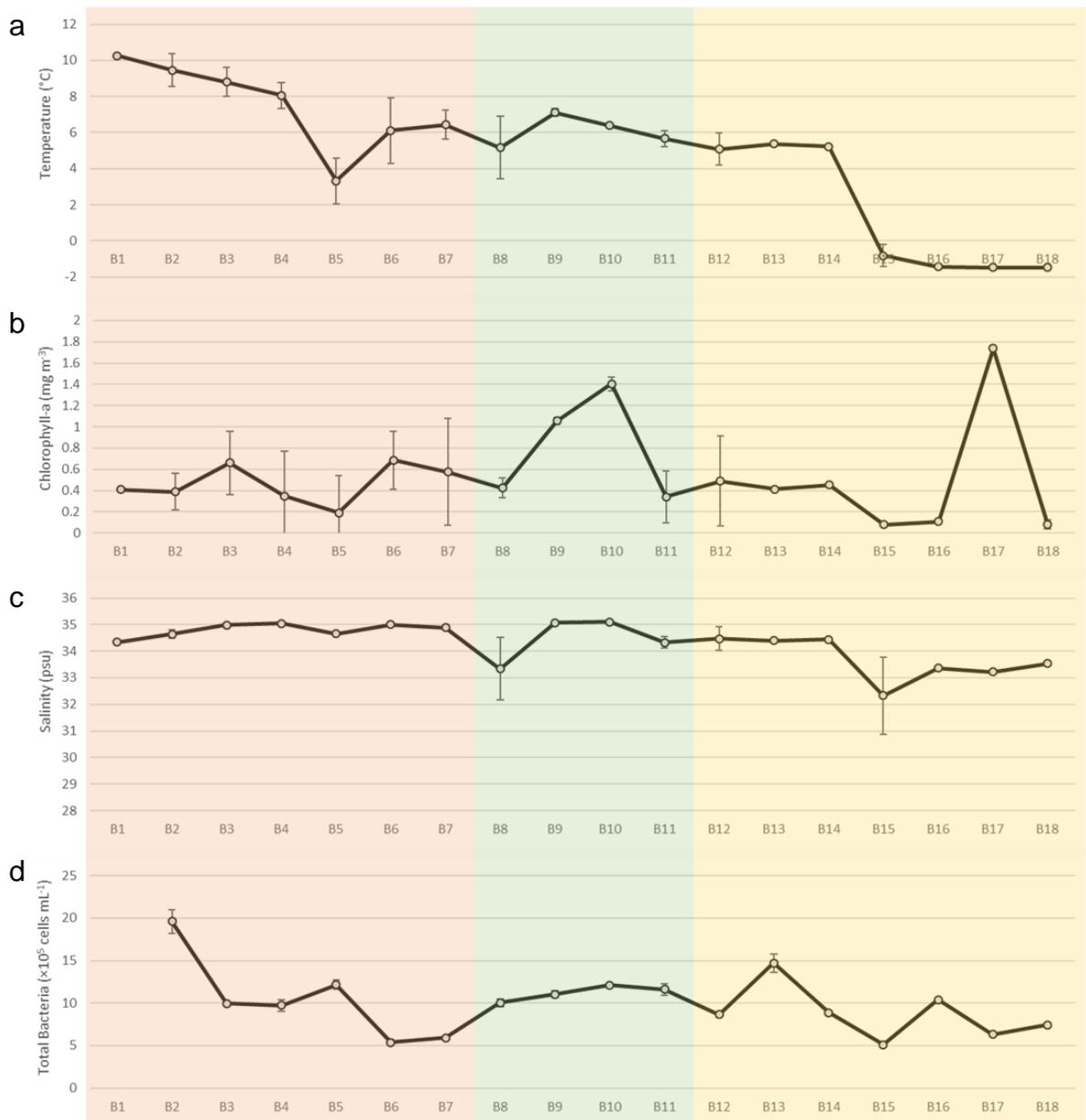


Figure 5.2 Measurements of (a) temperature, (b) chlorophyll-a, (c) salinity and (d) total bacterial abundance at the surface of the sampled stations from cruise JR16006. Arctic regions: Norway to Svalbard shelf (red), Svalbard shelf transect (green), and the Barents Shelf transect (yellow). Data collected June – July 2017 and available from the British Oceanographic Data Centre (<https://www.bodc.ac.uk/>).

Nutrient concentrations were also determined for the surface water of each sampled station. Ammonium concentrations indicated a range of 1.4 – 0.1 $\mu\text{mol L}^{-1}$ with an average of $0.4 \pm 0.4 \mu\text{mol L}^{-1}$ (Figure 5.3a). The different cruise sectors indicated that on average the NSS had the highest ammonium concentrations ($0.7 \pm 0.5 \mu\text{mol L}^{-1}$), and the BST had the lowest concentrations ($0.2 \pm 0.1 \mu\text{mol L}^{-1}$). Phosphate concentrations ranged from 0.7 to 0 $\mu\text{mol L}^{-1}$ and were on average lowest across the BST ($0.1 \pm 0.2 \mu\text{mol L}^{-1}$; Figure 5.3b) and peaked at station B7, however no phosphate could be detected at stations B12 and B14. Concentrations of silicate indicated a range of 4.5 – 0 $\mu\text{mol L}^{-1}$, and on average the BST had the lowest measurements ($0.6 \pm 0.9 \mu\text{mol L}^{-1}$; Figure 5.3c) and averaged $1 \pm 1.2 \mu\text{mol L}^{-1}$ across all stations. Values for nitrate (Figure 5.3d) were found to range from 11.2 – 0 $\mu\text{mol L}^{-1}$ and averaged $2.4 \pm 2.3 \mu\text{mol L}^{-1}$, with the highest average across the NSS ($3.8 \pm 5.2 \mu\text{mol L}^{-1}$) and lowest detected across the BST ($1.8 \pm 2.8 \mu\text{mol L}^{-1}$). Nitrate was very limited across 7 of the stations, but variable with a peak at station B7 and an immediate drop to 0.6 $\mu\text{mol L}^{-1}$ at station B8. Finally, is nitrite which yielded values between 0.7 – 0.0 $\mu\text{mol L}^{-1}$ (Figure 5.3e), starting off at a peak at station B2 before dropping and averaging $0.04 \pm 0.05 \mu\text{mol L}^{-1}$ across all the other stations. All nutrient regimes were limited between stations B11 and B14, except ammonium, although it also indicated a lower concentration across these stations.

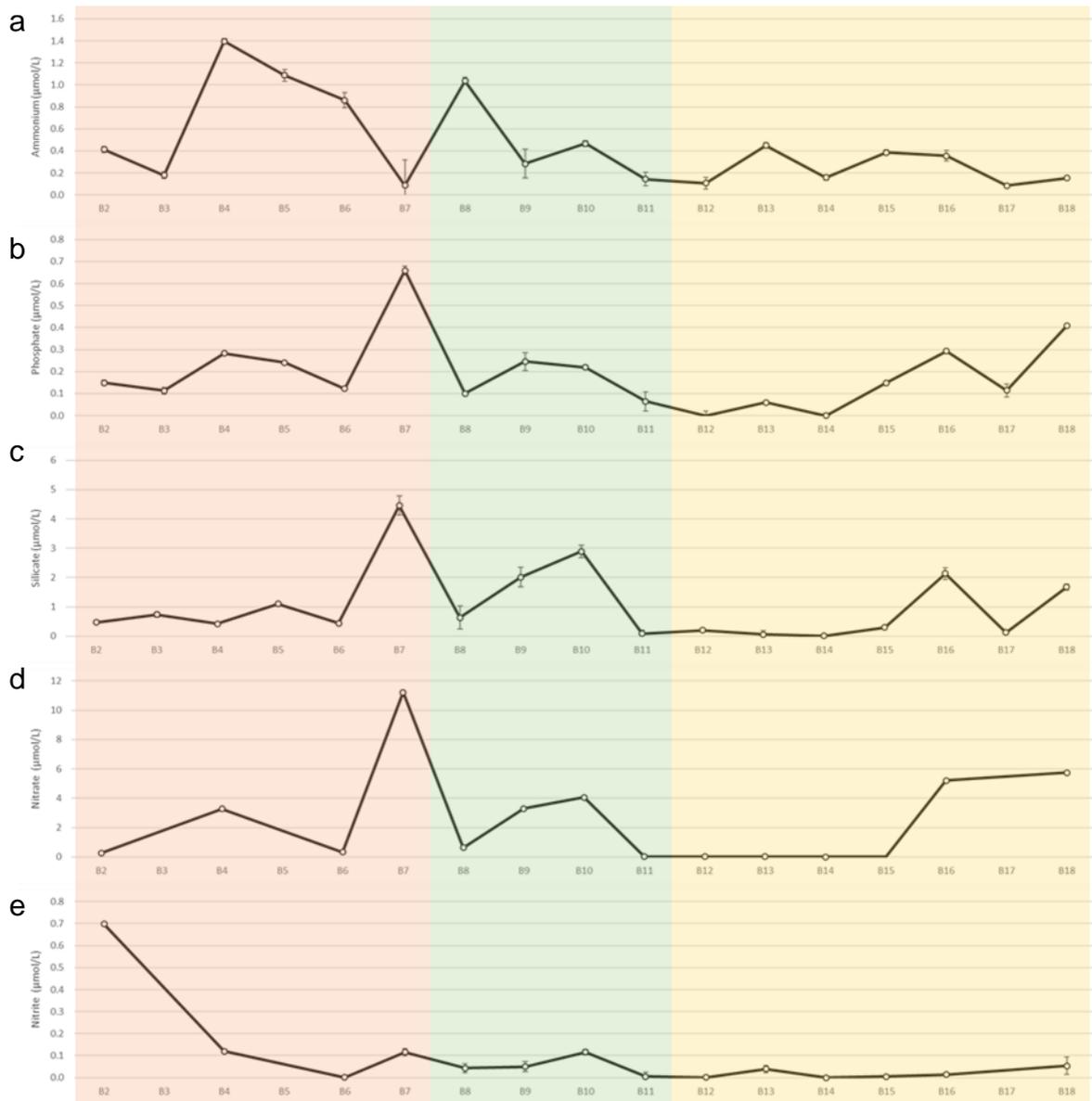


Figure 5.3 Concentrations of (a) ammonium, (b) phosphate, (c) silicate, (d) nitrate and (e) nitrite at the surface of the sampled stations from cruise JR16006. Arctic regions: Norway to Svalbard shelf (red), Svalbard shelf transect (green), and the Barents Shelf transect (yellow). Data collected June – July 2017 and available from the British Oceanographic Data Centre (<https://www.bodc.ac.uk/>).

5.2.2. Microbial diversity in Arctic and Atlantic regions

Spatial variations in bacterial community composition in different regions of the Atlantic and Arctic were determined by sequencing *16S (V4) rRNA* genes amplified from DNA samples (Section 2.10.7 and 2.10.8). 13 samples were taken from the surface of different stations along the polar front of the Barents Sea within the Arctic Ocean, and 16 samples were used from six different Atlantic provinces, which are represented from north to south (Figure 5.1 and Figure 5.4).

Samples taken from stations in the Barents Sea in the Arctic indicated that the surface microbial communities were typically dominated by the *Proteobacteria* ($56 \pm 13\%$), *Bacteroidetes* ($35 \pm 17\%$), *Cyanobacteria* ($8 \pm 8\%$) and, to a lesser extent, the *Verrucomicrobia* ($2 \pm 3\%$) (Figure 5.4a). The majority of *Proteobacteria* sequences were split between the *Gammaproteobacteria* ($46 \pm 12\%$), and the *Alphaproteobacteria* ($11 \pm 4\%$) class, whereas the *Bacteroidetes* (*Bacteroidia*), *Verrucomicrobia* (*Verrucomicrobiae*) and *Cyanobacteria* (*Oxyphotobacteria*) only consisted of one class each. The *Gammaproteobacteria* primarily comprised of the *Cellvibrionales* ($20 \pm 9\%$), *Oceanospirillales* ($17 \pm 10\%$) and SAR86 clade ($6 \pm 8\%$). The *Cellvibrionales* relative abundance between B1 and B11 averaged $16 \pm 6\%$, before increasing and averaging $31 \pm 7\%$ from B12 to B15. The *Oceanospirillales* relative abundance indicated variability amongst the sampled stations and was highest at stations B5, B8, B11 and B12 (averaging $30 \pm 2\%$), and averaged $11 \pm 5\%$ amongst the other stations. The SAR86 clade was only abundant at stations B1 and B2, and then B9 and B10 ($17 \pm 3\%$), similarly *Synechococcus* were also only abundant at these same four stations ($9 \pm 9\%$). Within the *Gammaproteobacteria*, there were also orders of *Thiomicrospirales*, *Alteromonadales*, *Betaproteobacteriales* and the *KI89A* clade, which across all Arctic stations only comprised between $<1 - 4\%$. The less abundant *Alphaproteobacteria* class was primarily made up of *Rhodobacterales* and SAR11 sequences, where the *Rhodobacterales* accounted for $8 \pm 3\%$ of the surface community across all stations and SAR11 made up $2 \pm 2\%$, and was more relatively abundant at stations B9 and B10. The *Bacteroidetes* were dominated by the *Flavobacteriales*, and increased gradually from B1 to B6, ranging from $9 - 74\%$ of the surface relative abundance,

respectively, and then at the rest of the stations, the *Flavobacteriales* averaged $34 \pm 12\%$.



Figure 5.4 Relative abundance (> 2%) of the bacterial community (*16S V4 rRNA*) in surface samples of (a) the Arctic (n=13), and (b) Atlantic Ocean (n=16) at the Class level. Arctic regions: Norway to Svalbard shelf (red), Svalbard shelf transect (green), and the Barents Shelf transect (yellow). Atlantic regions: NT (northern temperate), NSG (northern subtropical gyre), NTG (northern tropical gyre), EQU (equatorial upwelling), SG (southern gyre) and ST (southern temperate).

The northern temperate (NT) province indicated that *Proteobacteria* and *Cyanobacteria* dominated the surface microbial communities, consisting of $60 \pm 22\%$ and $31 \pm 25\%$ of the total relative abundance. The *Cyanobacteria* sequences primarily aligned with the *Synechococcales* order across all Atlantic regions, and showed a decrease in relative abundance from 49°N to 37°N . This coincided with an increase in the *Proteobacteria* abundance, which consisted mainly of the *Gamma*- and *Alphaproteobacteria* classes. The *Gammaproteobacteria* were the primary component of the *Proteobacteria*, making up $53 \pm 20\%$ of the NT surface community, and mainly comprised of the *Alteromonadales*, *Oceanospirillales* and the SAR86 clade orders (40 ± 17 , 7 ± 6 and $2 \pm 1\%$, respectively). The *Alphaproteobacteria* were not as relatively abundant as the *Gammaproteobacteria* in the NT region, but still made up $7 \pm 3\%$ of the surface community, consisting of the *Rhodobacterales* order and SAR11 clade (4 ± 1 and $1 \pm 1\%$). The *Bacteroidia* class of bacteria were primarily *Flavobacteriales*, and indicated an increase in relative abundance with decreasing latitude, making up 2% of the relative abundance at 49°N and increasing to 7% at 37°N . The *Acidomicrobia* class (within the *Actinobacteria* phylum) increased alongside the *Bacteroidia* between 49 and 37°N , going from 1 to 8%, respectively.

The northern subtropical gyre (NSG) region indicated the same dominant bacterial classes, where the *Gammaproteobacteria* dominated with $55 \pm 18\%$ of the surface community, especially 31°N where *Gammaproteobacteria* made up 72% of the community. The *Gammaproteobacteria* had a variable order composition of *Alteromonadales*, *Cellvibrionales*, *Oceanospirillales* and the SAR86 clade. The *Alteromonadales* and *Oceanospirillales* drove the increase in *Gammaproteobacteria* at 31°N , where they consisted of 28 and 35% of the relative bacterial abundance, respectively. However, at 34 and 27°N , these abundances dropped, and they averaged $8 \pm 4\%$. The *Alphaproteobacteria* indicated a higher relative abundance than the NT region, averaging $12 \pm 5\%$, of which *Rhodobacterales* and SAR11 were again the main orders (6 ± 7 and $2 \pm 3\%$, respectively). The *Synechococcales* and *Actinomarinales* orders (belonging to the *Cyanobacteria* and *Actinobacteria* phyla, respectively) were the other major taxonomic groups in this region which both peaked at 34°N (20 and 26%,

respectively) before dropping off to $3 \pm 2\%$ of the microbial relative abundance at 31 and 27 °N.

Samples in the northern tropical gyre (NTG) indicated that the *Proteobacteria*, *Cyanobacteria* and *Actinobacteria* were the dominant bacterial phyla (61 ± 23 , 23 ± 17 and $11 \pm 6\%$, respectively). The *Alpha*- and *Gammaproteobacteria* classes made up 21 ± 7 and $40 \pm 16\%$ of the microbial relative abundance, respectively. The *Alphaproteobacteria* sequences showed that the *Rhodobacterales* order contributed $8 \pm 1\%$ to the microbial community, and the *Caulobacterales* order increased from $<1\%$ at 24 °N to 10% at 14 °N, however SAR11 only represented $<1\%$ of the sequences in this region. *Gammaproteobacteria* sequences were mainly represented by *Oceanospirillales*, *Salinisphaerales*, *Cellvibrionales*, *Alteromonadales* and SAR86 clade sequences. The *Oceanospirillales* order consisted of $17 \pm 2\%$ of the relative abundance in the NTG region, whereas the *Alteromonadales* and *Cellvibrionales* varied from $<1 - 2\%$ at 24 °N and 12 - 18% at 14 °N. The *Salinisphaerales* and SAR86 sequences averaged $4 \pm 2\%$ in the NTG. The *Synechococcales* (*Cyanobacteria*) were the most abundant order for the region reaching 34% at 24 °N and then dropping to 11% at 14 °N, and the *Actinomarinales* averaged 7%.

Within the equatorial (EQU) region of the Atlantic, sequences of *Cyanobacteria* were dominant in two of the EQU latitudes (10° and -4°) with overall relative average abundance of $42 \pm 30\%$ on that region, and again, primarily consisted of *Synechococcales* sequences. There was a decrease in the proportion of *Proteobacteria* sequences to $39 \pm 21\%$, which were made up of *Gamma*- and *Alpha*- and *Deltaproteobacteria* (24 ± 8 , 9 ± 4 and $7 \pm 10\%$, respectively). The *Gammaproteobacteria* indicated variability in abundance, which is made up of *Oceanospirillales* ($7 \pm 3\%$), *Alteromonadales* ($5 \pm 1\%$) and the *Cellvibrionales*, *Salinisphaerales* and SAR86 clade (3%) orders. The *Alphaproteobacteria* indicated that sequences of *Rhodobacterales*, *Caulobacterales*, SAR11 and *Rhizobiales* all ranged between $<1 - 3\%$. The *Deltaproteobacteria* were primarily *Desulfobacterales* and *Desulfovibrionales*, contributing 7% to the surface relative abundance and

appeared to only prevalent in the southern EQU latitude and in the southern gyre samples.

The southern gyre (SG) also indicated variability, where the latitude closest to the southern temperate region (-29 °N; Figure 5.4b) had the largest average relative proportion of *Alphaproteobacteria* ($24 \pm 15\%$), primarily consisting of *Rhodobacterales* ($11 \pm 11\%$) and *Caulobacterales* ($4 \pm 1\%$). As with the other gyre regions (NSG and NTG), the *Proteobacteria* were the largest phylum on average ($63 \pm 8\%$), followed by the *Cyanobacteria* ($12 \pm 4\%$) and *Actinobacteria* ($18 \pm 19\%$). The *Proteobacteria* consisted of the *Gamma*- and *Alphaproteobacteria* classes (31 ± 8 and $24 \pm 15\%$, respectively); however, the *Deltaproteobacteria* were also a significant class (8%). Within the *Gammaproteobacteria*, the *Alteromonadales* and *Oceanospirillales* orders were most relatively abundant ($11 \pm 1\%$ and $11 \pm 8\%$, respectively), followed by *Salinisphaerales* and the SAR86 clade (<5%). The *Alphaproteobacteria* mainly consisted of *Rhodobacterales* ($11 \pm 11\%$) and *Caulobacterales*, *Rhizobiales* and SAR11 sequences (<4%). The relative abundance increase in *Deltaproteobacteria* was derived from *Bradymonadales* and *Myxococcales* sequences (<6%). As mentioned previously, the *Synechococcales* were very relatively abundant in the EQU and NT regions and lower in the gyre regions and is a trend which continues in the southern Atlantic averaging $7 \pm 3\%$, the lowest average abundance of all the regions. Finally, the *Actinobacteria* were primarily sequences of the *Actinomarinales*, as in other regions, and consisted of $18 \pm 20\%$ of the microbial relative abundance in SG.

The southern temperate (ST) region was the most southerly sampled area of the Atlantic and indicated a significant dominance of the surface microbial communities by the *Cyanobacteria* ($83 \pm 17\%$), especially at -39 °N where *Synechococcales* sequences accounted for 86 % of the relative abundance. This was followed by the *Proteobacteria* ($8 \pm 8\%$), which mainly constituted of *Cellvibrionales* and SAR86 clade sequences (<4%).

5.2.3. Variability in *coxF5* gene diversity

Spatial variations in *coxF5* diversity in different regions of the Atlantic and Arctic were investigated by sequencing *coxF5* genes, where amplification from DNA samples was possible. Eight out of the 16 sampled latitudes in the Atlantic yielded *coxF5* amplicons (Figure 5.5), and only three stations in the Arctic region, however these Arctic amplicons could not be sequenced due to low amplicon DNA concentrations and quality.

Within the Atlantic, *coxF5* sequences overwhelmingly aligned with the *Alphaproteobacteria* class from the NTG to the ST regions ($97 \pm 4\%$), whereas in the NT and NSG regions an unassigned class accounted for $49 \pm 3\%$ of the *coxF5* relative abundance, followed by the *Alphaproteobacteria* ($35 \pm 18\%$) and then the *Gammaproteobacteria* ($16 \pm 16\%$). The *Alphaproteobacteria* were almost exclusively attributed to the *Rhodobacteraceae* family of bacteria (Figure 5.5a), which made up $94 \pm 4\%$ of the *coxF5* sequences from the NTG down south to the ST region. The *Rhodobacteraceae* sequences that could be resolved to the genus level, primarily aligned with the *Salipiger* and *Sagittula* genera (Figure 5.5b). There were also sequences of *Methylobacteriaceae*, which accounted for $4 \pm 3\%$ and $4 \pm 5\%$ of the *coxF5* diversity in the SG and ST regions, and could be identified as genera belonging to the *Methylobacterium* and *Methylobacterium*.

The *Gammaproteobacteria* sequences aligned with *Piscirickettsiaceae* and were only a major component of the *coxF5* diversity in the NT region, where they accounted for 26% of the sequences, but were also minor groups in the NSG and EQU regions ($4 \pm 2\%$). All of the *Gammaproteobacteria* *coxF5* sequences aligned with the *Methylophaga thiooxydans* species (Figure 5.5b). The unassigned sequences in the two northern regions (NT and NSG) indicated highest degrees of similarity with *Alphaproteobacteria* sequences when using BLASTn (sequence identity $<80\%$), which would reflect the other regions, indicating that *Alphaproteobacteria* averaged $92 \pm 8\%$ across all Atlantic regions.

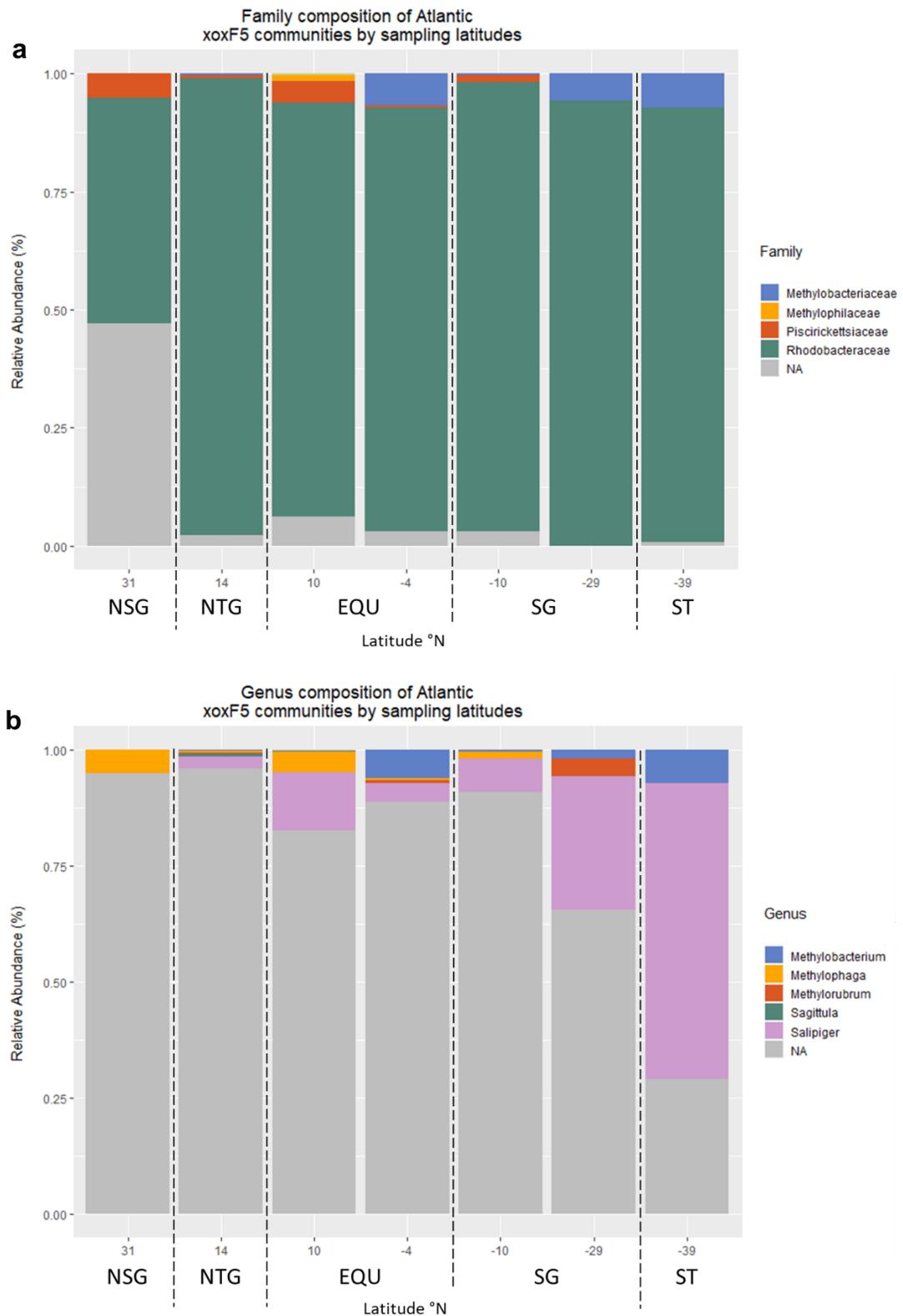


Figure 5.5 Relative abundance of the *coxF5* community in surface samples of the Atlantic Ocean at the (a) Class and at the (b) Family level. NSG (northern subtropical gyre), NTG (northern tropical gyre), EQU (equatorial upwelling), SG (southern gyre) and ST (southern temperate). NA indicates non-assigned sequences.

5.3. Discussion

We present the first basin wide open ocean spatial study of the *xoxF5* gene and its diversity. The *xoxF5* gene was successfully amplified from surface samples in all sampled Atlantic regions and sequenced, however samples from the Barents Sea region of the Arctic did not yield any amplifiable PCR products. In the successful amplicons, the vast majority of *xoxF5* sequences were derived from the *Rhodobacterales* order of bacteria and indicated little variability in the diversity of *xoxF5* at the family taxonomic level between the different regions of the Atlantic, primarily consisting of *Rhodobacteraceae* sequences (Figure 5.5a). This builds upon the findings of Taubert *et al.* (2015), who indicated that there was major taxonomic variability in the *xoxF5* gene amongst different coastal sites, whereas this novel open ocean study suggests that in offshore surface water, the diversity of *xoxF5* is much more restricted. Furthermore, there is still unknown *xoxF5* diversity within the *Rhodobacteraceae* family, indicated by the larger proportion of unassigned sequences at lower taxonomic levels (Figure 5.5b).

Rhodobacteraceae is a large family of bacteria with over 180 different genera and contains many identified methylotrophs, such as *Rhodobacter*, *Roseobacter*, *Roseovarius* and *Sagittula* (Gonzalez *et al.*, 1997; Barber and Donohue, 1998; Grob *et al.*, 2015; Taubert *et al.*, 2015). However, within the Atlantic *xoxF5* samples only two *Rhodobacteraceae* genera could be identified - *Sagittula*, a known methanol utilising marine methylotroph, and also *Salipiger*, which has no previous indication of methylotrophic capabilities (Gonzalez *et al.*, 1997; Kolb, 2009). As the *Salipiger* genus is the most identified *Rhodobacteraceae* genus, this further supports the suggestion that organisms containing *xoxF5* may have unidentified methylotrophic potential or that the gene is involved in some other metabolic role. This proposal originates from indications that *xoxF5* is the functional methanol dehydrogenase (MDH) in *Rhodobacter sphaeroides*, shown by deletion mutants losing the ability to metabolise methanol (Wilson, Gleisten and Donohue, 2008), furthermore, *xoxF5* is required by *Methylorubrum extorquens* AM1 for expression of the MxaFI derived MDH (Skovran *et al.*, 2011), highlighting different scenarios where *xoxF5* is important for methanol metabolism.

Rhodobacterales xoxF5 sequences were dominant across most of the Atlantic, however there were sequences in the northern Atlantic regions which could not be assigned at the class taxonomic level. Although these sequences related closest with *Alphaproteobacteria* when BLASTn was used (>80% identity). This is the class that the *Rhodobacterales* belong to and further indicates the dominance of the *Alphaproteobacteria* regarding *xoxF5* diversity in the open ocean and suggests that there are other microbes belonging to this class yet to be sequenced and identified which also contain the *xoxF5* gene.

In addition to the *xoxF5* diversity study, the microbial community was also investigated in all samples using the *16S (V4) rRNA* gene. The diversity of microbial community along the Atlantic transect highlights known community features, such as the greater relative abundance of *Synechococcus* at the equator and in the temperate regions, which is associated with relatively higher levels of primary productivity and driven by upwelling of cold, nutrient rich waters (Sigman and Hain, 2012). Additionally, is the high proportion of *Gamma*- and *Alphaproteobacteria* throughout the Atlantic (Sargeant *et al.*, 2018; Zorz *et al.*, 2019), and similarly to Sargeant *et al.* (2018), this study found that the SG is a more diverse environment (Figure 5.4).

Samples taken from the Barents Sea in the Arctic Ocean, could not produce any sequenceable PCR products for *xoxF5*, but did for *16S (V4) rRNA*, thereby providing insight to the microbial community. The *Bacteroidetes* group of bacteria have a similar average relative abundance ($30 \pm 13\%$) to another study (14 – 21%; Kirchman, Cottrell and Lovejoy, 2010), however the abundance increases with latitude along the Norway to Svalbard Shelf (NSS; Figure 5.1a) and then dominates at station B6 (74%). The sampling stations in the Arctic are located along a front where significant parcels of water originating from the polar regions and the Atlantic interact (Laukert *et al.*, 2019). Station B6 has a shallow bathymetry (~140 m), where the East Spitsbergen and Hopen-Bjørnøya Currents (cold, fresh polar water) mix with the Norwegian Atlantic Current (warm, saline Atlantic water), and the station was sampled at the end of the spring bloom (June-July). Under these conditions, *Flavobacteriales* could dominate a microbial community of a bloom in decay phase

(Buchan *et al.*, 2014) - indicated by the high ammonium, nutrient depleted water (Figure 5.3), however without phytoplankton data it is difficult to comment further.

Cellvibrionales is another bacterial group, which indicated more abundance in the northern most sampled stations, B12 to B15 ($31 \pm 7\%$), when compared to the average across the other stations ($16 \pm 6\%$). These three stations constitute the Barents Shelf Transect (BST) and are highly influenced by currents transporting nutrient-rich, cold water from the northern pole. This likely explains why there was an increase in *Cellvibrionales* in those stations alone, as Spring *et al.* (2015) explains that some of the *Cellvibrionales* lineages prefer nutrient-rich conditions for growth, specifically highlighting *Porticoccaceae*, *Haliaceae* and *Spongiibacteraceae* which happen to be the three most abundant *Cellvibrionales* families in the Atlantic samples. Conversely to the *Cellvibrionales*, the SAR86 bacterial clade is only abundant at stations B1, B2, B9 and B10 ($17 \pm 3\%$). Nitrite levels are much higher at the start of the NSS stations (B1 and B2), however the SAR86 clade has no apparent capacity to assimilate nitrite, instead producing nitrite via a reduction of nitro-aromatic compounds (Dupont *et al.*, 2012). Therefore, this observation could also be influenced by the currents in this region. As these four stations are all located along known passages for Atlantic Water masses, specifically the Norwegian Atlantic Current (NWAC), which follows the coast of Norway and then splits and continues along the coastline to stations B1 and B2. The NWAC also follows the edge of the shelf break along the NSS, eventually reaching stations B9 and B10, southwest of Svalbard (Figure 5.1a). Thus, it is likely that the similarities observed with SAR86 abundance are strongly influenced by the ocean currents, similar to recent work which indicated that currents resulted in distinct microbial communities on and off-shelf in the Atlantic (Zorz *et al.*, 2019).

The *Alphaproteobacteria*, which comprises *Rhodobacterales*, is an important bacterial group for *xoxF5* diversity. The 16S (V4) *rRNA* data revealed that *Oceanibulbus* genera provided the majority of the *Rhodobacterales* community abundance ($44 \pm 35\%$), although it does not contain a *xoxF5* gene. Additionally, this genera was much more dominant in the samples taken north of the equator, making up $76 \pm 23\%$ of the *Rhodobacterales* sequences, compared to the other more

southern samples which were more diverse in regards to *Rhodobacterales* sequences. Interestingly, *xoxF5* samples were only successfully amplified in the regions where the relative abundance of *Oceanibulbus* was lower. Conversely, *Roseovarius* is another genera belonging to the *Rhodobacteraceae* (*Rhodobacterales*) which does contain *xoxF5*, and was only detectable in the samples where *xoxF5* products were successfully amplified - averaging $52 \pm 32\%$ of the *Rhodobacterales* community. This perhaps indicates that the unassigned *xoxF5* sequences are *Roseovarius* homologs.

A correlation approach was used to investigate links between the methanol dissimilation rates in the Atlantic with the microbial and *xoxF5* communities. Uptake rates were determined in the Atlantic during the same cruise that the DNA samples were collected, and indicated statistically significant positive correlations for methanol dissimilation with *Flavobacteriaceae*, *Pseudoalteromonadaceae* ($r = 0.502$, $n = 31$, $P \leq 0.05$ and $r = 0.531$, $n = 31$, $P \leq 0.03$, respectively) and *Alcanivoracaceae* ($r = 0.677$, $n = 31$, $P \leq 0.01$). Relationships have previously not been indicated for *Pseudoalteromonadaceae* and *Alcanivoracaceae* with methanol dissimilation, and interestingly neither of these bacterial groups have been identified as methylotrophic. Alternatively, the *xoxF5* sequences indicated only one correlation, a statistically significant negative relationship existing between methanol dissimilation and *Rhodobacteraceae* ($r = -0.356$, $n = 23$, $P \leq 0.03$). A negative relationship has been indicated previously between these two variables, with the addition of SAR11 sequences (Chapter 4), interestingly SAR11 was again not indicated as a relatively abundant bacterial group in Atlantic samples using this approach $1.3 \pm 1.2\%$, but were nonetheless detected in all samples and were most prevalent in the NSG (5%).

It is curious that there was less success in amplifying and sequencing *xoxF5* genes in samples taken from the Barents Sea region of the Arctic. There is limited rare earth element (REE) data for the Arctic Ocean, however a recent study has indicated that lanthanum concentrations throughout the water column in this region, are equivalent to samples from the SE Atlantic open ocean ($\sim 26 \text{ pmol kg}^{-1}$; Crockett *et al.*, 2018; Laukert *et al.*, 2019). Additionally, concentrations of lanthanum are

reportedly higher closer to the coast (64 pmol kg^{-1} , Icelandic Shelf), and are generally attributed to terrestrial runoff and glacier meltwater (Crocket *et al.*, 2018; Laukert *et al.*, 2019). However, as no *xoxF5* sequences were detected in samples closer to coastlines (i.e. B1 and B8), the perceived relationship indicated between lanthanum concentrations and *xoxF5* offers little explanation regarding unamplified *xoxF5* products from Arctic samples, and is more likely a direct result of lower concentrations of DNA on the filter following extraction. As *xoxF5* is associated with methanol oxidation, it is also worth considering the methanol concentrations as another potential factor influencing *xoxF5* gene amplification, however, much the same as REEs, the methanol concentration measurements in the Arctic are also limited. However, recent analysis in the Canadian Arctic has indicated methanol concentrations averaging $17 \pm 6 \text{ nM}$ (Wohl *et al.*, 2019), which is lower than the concentration range reported in the Atlantic ($48 - 361 \text{ nM}$; Sargeant, 2013) and Western Channel Observatory (WCO; $97 \pm 8 \text{ nM}$; Dixon, Beale and Nightingale, 2011), but is still an acceptable concentration used in methanol utilising lab cultures.

Alternatively, a more likely explanation is that the *xoxF5* gene abundance is too low in the Arctic samples to be detected or that there is a potential bias introduced by this PCR-based approach, which has been indicated previously by Taubert *et al.* (2015). *Rhodobacteraceae* have been implicated as a major reservoir of *xoxF5* diversity, and this could typically indicate absence or low abundance of that taxonomic family in the Arctic samples (Taubert *et al.*, 2015). However, as *Rhodobacteraceae* sequences were detected in all Arctic samples using a 16S (V4) *rRNA* primer set, it is more likely that it is absence or low abundance at a lower taxonomic level. In samples from the Atlantic and the WCO (Chapter 4), 173 ASVs are associated with *Rhodobacteraceae*, compared with 36 ASVs in Arctic samples (where 30% and 8% are unassigned at the genus level, respectively). Therefore, there is a less diverse pool of *Rhodobacteraceae* genera identified in the Arctic samples and also indicates that although *Rhodobacteraceae* has been found to be a reservoir of *xoxF5* diversity, it should not be used as an indicator of *xoxF5* presence.

5.4. Conclusion

Samples taken from cruise transects in the Atlantic and Arctic Ocean were successfully used to give the first indications of the distribution and diversity of the *xoxF5* gene clade in the open ocean. This study determined that *xoxF5* is detectable in surface samples from all sampled regions and was also amplifiable in all regions of the Atlantic, apart from the northern temperate samples. The *Rhodobacteraceae* family of bacteria appear to be the primary reservoir for this gene in the surface waters of the open ocean regions, however, below this taxonomic level there is still ambiguity with most sequences remaining unassigned.

Samples taken from the Barents Sea region of the Arctic Ocean did contain sequenceable PCR products, therefore although the diversity of *xoxF5* remains unknown in this region, the gene could be amplified and sequenced with more resources. This would be of further interest if combined with microbial methanol uptake rates determined in the same regions, as extremely little is known regarding methylotrophy in the more productive polar oceans. This could be an important metabolic approach during the ice-covered winter months when chemotrophic-based food webs dominate (Bunse and Pinhassi, 2017).

As indicated in Chapter 4, *xoxF5* covers the broadest taxonomic range (*Alpha*-, *Beta*- and *Gammaproteobacteria*), which is why it was the gene clade targeted for this study. However, the other gene homologs have been linked with various other taxonomic groups of the microbial community, which the *xoxF5* gene clade does not represent. As already stated, facultative and obligate methylotrophs could be an important metabolic group in the Arctic, attributing more importance to all clades of the *xoxF* gene.

Chapter 6

Synthesis and conclusions

6. Synthesis and conclusions

6.1. Virus-like particles associated with an obligate methylotroph

Methanol provides a crucial source of carbon to methylotrophic microbes in the environment, although almost nothing is known about the viruses which infect the microorganisms which occupy an important niche in marine carbon cycling. Various approaches were utilised to identify enveloped virus-like particles using electron microscopy, which suggested that the life cycle associated with the virus-host system may be distinct from the typically reported lytic and lysogenic interactions. The aim of this work was to isolate a lytic virus infection, however this was not achieved and could have been repeatedly prevented at multiple stages of infection by the host cell, such as resisting virus adsorption, blocking phage genome insertion or the use of restriction enzymes (Figure 1.9) (Hyman and Abedon, 2010). Alternatively, although *Methylophaga* is predominant amongst marine methylotrophs and isolatable from a range of marine environments, it has a low relative abundance within the WCO, which is likely a limiting factor on the abundance of any lytic viruses associated with this genus (Wommack and Colwell, 2000; Neufeld *et al.*, 2007). In future studies, this issue could be mitigated by targeting specific periods of the year when virus abundance is higher (i.e. summer months as highlighted in Chapter 4), when the *Methylophaga* group is more relatively abundant (October to November according to this survey), or concentrating and filtering a larger volume of seawater (>20 L) for potential lysate. If these approaches (independently or combined) were to successfully result in a lytic infection of an isolate of the *Methylophaga* genus, then this would yield more virus protein and genomic material for further virus characterisation studies (Figure 6.1). Furthermore, with a repeatable infection system in place, direct studies on the influences of virus infection on methanol uptake rates could be determined, thereby yielding crucial information regarding the direct impacts at the cellular level.

With a sequenced genome, a plethora of additional analysis and experimentation can be undertaken, one approach could involve identifying conserved regions within the virus genome and creating customised primer sets to target and then investigate seawater samples (Short, Chen and Wilhelm, 2010). Simplistic presence and

absence sampling could lead to eventual sequencing and diversity studies of primer targets, much like the way the *xoxF5* clade has been targeted in this project using previously established primer sets (Taubert *et al.*, 2015).

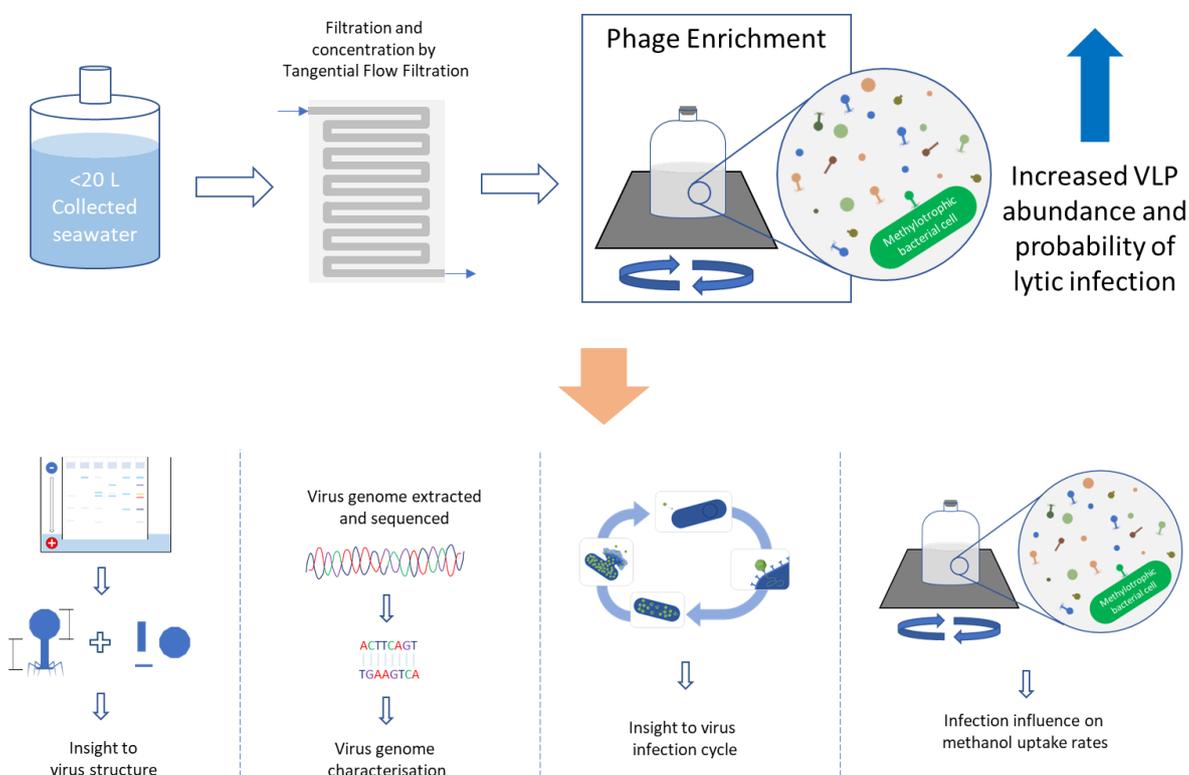


Figure 6.1 Schematic indicating how the phage enrichment approach could be better optimised to increase the probability of infecting a lab strain with a lytically infecting virus. The orange arrow then highlights the different avenues that could be pursued if a lytic system could be established.

If these approaches were unsuccessful, then recent findings have indicated the successful isolation of lytic viruses associated with the methylophilic OM43 clade of bacteria (Buchholz and Temperton, *pers. comms.*). Time course experiments with this virus-host system could also be combined with methanol uptake rate experiments to determine the impact of lytic infection upon host methanol uptake rates. The other approaches mentioned earlier would be suitable in this instance as well and sequencing of the genome of these isolated viruses could yield exciting insights to ‘methylophage’ and their likely impacts on methanol metabolism.

6.2. Viruses and methanol uptake rates in seawater

This study included the most detailed study of virus-like particle (VLP) abundances in the WCO to date (in combination with microbial methanol uptake rates) and confirms the global understanding of temporal and depth variation in virus abundance in marine systems. These VLP abundances were analysed alongside methanol uptake rates and identified a significant inverse relationship with methanol dissimilation rates and total VLP abundances. This relationship implicates viruses directly within methanol cycling in the marine environment but raises further questions regarding the direct effects upon host metabolism, potentially changing the role of methylotrophs as a source or sink of methanol when infected (Figure 6.2).

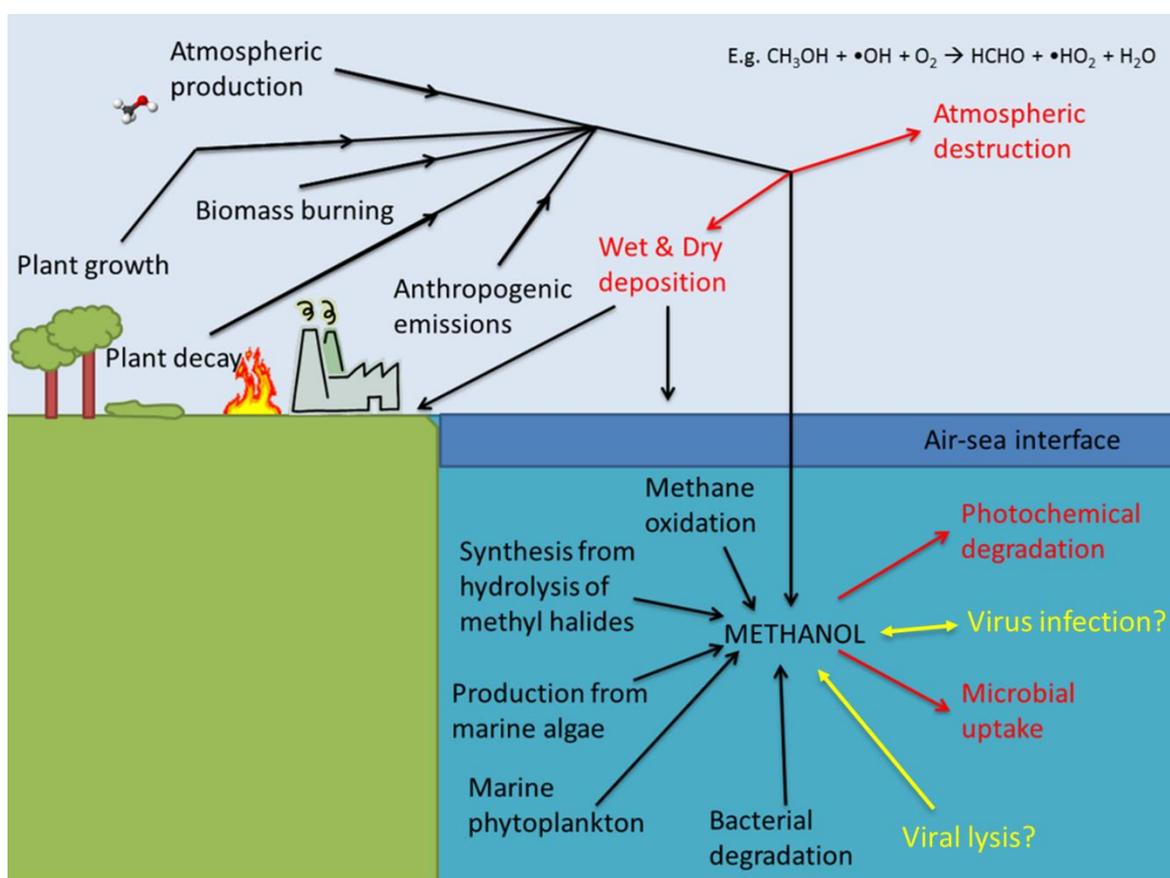


Figure 6.2 Updated simplified overview of different sources and sinks in the environment. Sources are indicated by black lines, whereas potential sinks are represented by red lines and text. Yellow lines and text represent proposed links to virus infection and viral lysis. Figure taken from Sargeant (2013).

The indirect role of viruses upon methanol uptake rates was highlighted in this project (particularly methanol dissimilation rates), where an increased virus pressure upon the varying trophic levels of the marine food web potentially drives a shift away from methanol to other available carbon sources by facilitating the release of organic matter via lytic activity. This switching in methanol metabolism is most likely attributed to facultative methylotrophs and may also indicate that facultative methylotrophs comprise a significant proportion of the methanol utilisers in the water column and are the primary drivers of seasonal variation in methanol uptake rates (Figure 6.3).

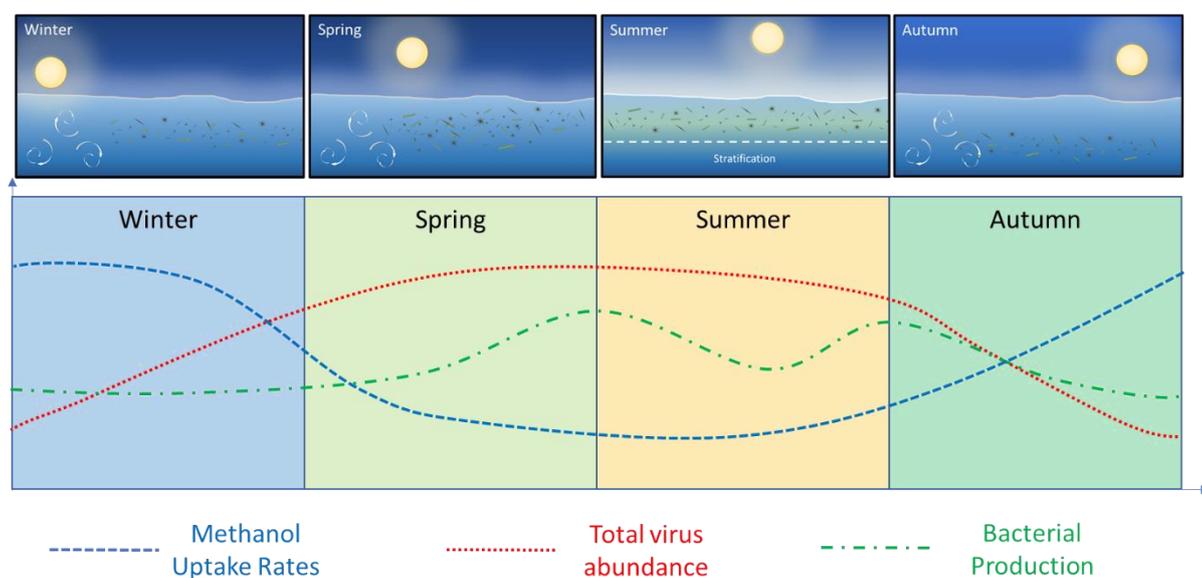


Figure 6.3 Simple schematic overview of the seasonal variation in methanol uptake rates (blue), virus abundance (red) and the bacterial production (green). Additional plots of the change in the water column that occurs throughout the year during the different seasons aligned with the findings of this project.

Therefore, during spring and summer months the methanol uptake rates can be attributed to obligate methanol utilisers, as virus abundance is higher and facilitating an alternate carbon source to facultative methylotrophs during bloom periods. Then during the colder, less productive autumn and winter months the facultative methylotrophs begin to preferentially scavenge methanol, thereby driving an increase in methanol dissimilation rates when other carbon sources are scarcer in the environment as virus pressure is lower. If this can be corroborated then it would

be an important finding for the greater understanding of microbial methanol utilisation in the marine environment, implicating viruses as a key control of microbial methanol metabolism.

The role that phytoplankton play in methanol cycling is likely also very significant, especially as producers of methanol (previously identified as a source in Figure 6.2), which is then utilised by methylotrophic bacteria, as already indicated by the OM43 clade association with diatom blooms (Morris, Longnecker and Giovannoni, 2006). This could be a source of methanol facilitation by viruses to methanol utilisers, via the release of methanol from phytoplankton cells upon viral mediated lysis, as viruses do exhibit a degree of control over phytoplankton blooms (Wilson *et al.*, 2002). The experiments of Mincer and Aicher (2016) implicated cultures of *Synechococcus* sp., *Prochlorococcus* sp. and *Emiliana huxleyi* with methanol production, all of which have established virus-host systems and would be logical candidates for any future culture experiments (Suttle and Chan, 1993; Wilson *et al.*, 2002; Sullivan, Waterbury and Chisholm, 2003). Additionally, the industrially utilised *Phaeodactylum tricornutum* was also implicated in methanol production, and indicated one of the highest methanol production rates in the study by Mincer and Aicher (2016), however this organism has also unsuccessfully yielded an infecting virus after previous extensive efforts, reiterating the difficulties associated with establishing a virus-host system, even with a well-established lab strain (Allen and Rooks, *pers. comm.*). Efforts to determine this direct association were attempted, but are not presented in this thesis due to several issues involved in adapting dilution and reduction methodologies to this system (Winget *et al.*, 2005; Kimmance and Brussaard, 2010; Wommack *et al.*, 2010). However, preliminary results did indicate that changes in the virus-to-bacteria ratio in natural seawater assemblages can influence the methanol uptake rates and may be worth pursuing further.

6.3. XoxF5 diversity and distribution in the marine environment

Environmental expression of the xoxF gene clades and relevant proteins has suggested their widespread importance to biogeochemical processes (including

methylophony), where the *xoxF5* clade exists across the broadest taxonomic range of all the homologs potentially indicating an increased importance. This study has provided crucial insight to the diversity and distribution of the *xoxF5* gene clade, which was previously unknown across much of the marine environment. This revealed that in the western English Channel, the *Rhodobacteraceae* family dominated the majority of *xoxF5* sequences, although there was a brief period at the end of spring when the *Rhizobiales* were the dominant order in relative abundance. The dominance of these two groups over *xoxF5* diversity has previously been indicated at the same sampling location (Taubert *et al.*, 2015); however this study expands those results with the first ever timeseries analysis of this gene, and indicates little seasonal variation in diversity at this coastal location. Interestingly, there was an increase in the relative abundance of *Burkholderiales* in March, corresponding with significant riverine input, potentially implicating a response in *xoxF5* diversity driven by an increase in terrestrially derived rare earth elements (REEs) (Goldstein and Jacobsen, 1988; Sholkovitz, 1993). The dominance of the *Rhodobacterales* sequences amongst the *xoxF5* community likely makes it difficult to link seasonal changes with the methanol uptake rates. A better approach could be a similar study using quantitative PCR to determine *xoxF5* gene abundance in the marine environment, similar to what has been successfully carried out using a *xoxF4* primer set, which revealed spatiotemporal variation in abundance and expression levels (Ramachandran and Walsh, 2015). Ideally concentrations of REEs could also be determined, giving a more holistic understanding of the suspected role that REEs (or specifically lanthanum) play in *xoxF* function in the environment.

The spatial diversity of *xoxF5* in different regions of the Atlantic and within a region of the Arctic Ocean once again indicated the relative dominance of the *Rhodobacteraceae* family of bacteria. However, within the *Rhodobacteraceae* family only two genera were identified in the Atlantic samples (*Sagittula* and *Salipiger*), whereas within the *xoxF5* timeseries data the *Rhodobacteraceae* family was indicated to include these two genera plus the *Yangia*, *Roseobacter*, *Polymorphum*, *Paracoccus*, *Pelagibaca* and the *Rhodobacter* genera. Indicating an increased *xoxF5* gene diversity at a coastal location, compared to an open ocean

transect - even amongst one taxonomic family. The low abundance of *xoxF5* sequences in Arctic samples was surprising as *Rhodobacterales* was detected in 16S sequences at all Arctic sampling sites and are known to be distributed and diverse in the Arctic (Fu *et al.*, 2013). This indicates that *xoxF5* is not conserved across the *Rhodobacterales* order, and further highlights the lack of knowledge of *xoxF5* at the genus level in the marine environment, as pointed out by Taubert *et al.* (2015).

Regardless of the link between the *xoxF* gene and methanol metabolism, methanol utilisation may be of significant interest within the polar regions, as extremely little is known regarding methylotrophy in the productive Arctic and Antarctic oceans. This could be an important metabolic approach in locations of expansive ice-coverage during winter months when chemotrophic-based food webs dominate the water column (Bunse and Pinhassi, 2017). This was hinted at in this study by the higher methanol uptake rates during the colder months highlighted earlier (Figure 6.3), and could leave the diversity of any methylotrophic community susceptible to climate change driven reductions in ice coverage at the poles (Evans *et al.*, 2017). This would present some interesting challenges regarding sampling and determining methanol uptake rates in the water column beneath ice cover but could establish the importance of methylotrophy within the arctic food web.

List of abbreviations

ADH	Alcohol dehydrogenase
AFC	Analytical flow cytometry
AMT	Atlantic Meridional Transect
ASV	Amplicon sequence variants
BCD	Bacterial carbon demand
BGE	Bacterial growth efficiency
BGE _M	Bacterial growth efficiency of methanol
BLI	Bacterial leucine incorporation
BP	Bacterial production
BSA	Bovine serum albumin
C1	One-carbon
CF	Conversion factor
Ci	Curies
CsCl	Caesium chloride
CTAB	Cetyltrimethylammonium bromide
CTD	Conductivity-temperature-depth probe
DADA	Divisive amplicon denoising algorithm
dH ₂ O	Distilled H ₂ O
DMS	Dimethyl sulphide
DMSO	Dimethyl sulfoxide
DMSP	Dimethylsulfoniopropionate
DNA	Deoxyribose nucleic acid
DOC	Dissolved organic carbon
DOM	Dissolved organic matter
DPM	Disintegrations per minute
ECF	Empirical conversion factor
EDTA	Ethylenediaminetetraacetic acid
GFCA	Glass fibre cellulose acetate

HGT	Horizontal gene transfer
HNA	High nucleic acid bacteria
KDPG	2-keto 3-deoxy 6-phosphogluconate
K_M	Michaelis constant
LNA	Low nucleic acid bacteria
LSC	Liquid scintillation counter vials
MAMS	Marine ammonium mineral salts
MB	Marine broth
MBM	Marine basal media
MDH	Methanol dehydrogenase
NMDS	Non-metric multidimensional scaling

References

- Abedon, S. T. and Yin, J. (2009) 'Bacteriophage Plaques: Theory and Analysis', in *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*, pp. 161–174. doi: 10.1007/978-1-60327-164-6.
- Ackermann, H. (2009) 'Basic Phage Electron Microscopy', in Clokie, M. R. J. and Kropinski, A. M. (eds) *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*. Humana Press, pp. 113–126. doi: 10.1007/978-1-60327-164-6.
- Ackermann, H. W. and DuBow, M. S. (1987) 'Viruses of Prokaryotes Vol. 1', in *General Properties of Bacteriophages*.
- Anantharaman, K. *et al.* (2014) 'Sulfur Oxidation Genes in Diverse Deep-Sea Viruses', *Science*, 145, p. 7. doi: 10.3354/meps145269.
- Angly, F. E. *et al.* (2006) 'The marine viromes of four oceanic regions', *PLoS Biology*, 4(11), pp. 2121–2131. doi: 10.1371/journal.pbio.0040368.
- Anthony, C. (1982) *The Biochemistry of Methylotrophs*.
- Anthony, C. (2011) 'How half a century of research was required to understand bacterial growth on C1 and C2 compounds; the story of the serine cycle and the ethylmalonyl-CoA pathway', *Science Progress*, 94(2), pp. 109–137. doi: 10.3184/003685011X13044430633960.
- Anthony, C. and Ghosh, M. (1998) 'The structure and function of the PQQ-containing quinoprotein dehydrogenases', *Progress in Biophysics and Molecular Biology*, 69(1), pp. 1–21. doi: 10.1016/S0079-6107(97)00020-5.

Anthony, C. and Williams, P. (2003) 'The structure and mechanism of methanol dehydrogenase', *Biochimica et Biophysica Acta - Proteins and Proteomics*, 1647(1–2), pp. 18–23. doi: 10.1016/S1570-9639(03)00042-6.

Anthony, C. and Zatman, L. J. (1964) 'The microbial oxidation of methanol. 1. Isolation and properties of *Pseudomonas* sp. M27.', *The Biochemical journal*, 92(3), pp. 609–14.

Anthony, C. and Zatman, L. J. (1967) 'The microbial oxidation of methanol. The prosthetic group of the alcohol dehydrogenase of *Pseudomonas* sp. M27: A new oxidoreductase prosthetic group', *Biochemical Journal*, 104(3), pp. 960–969.

Antony, C. P. *et al.* (2010) 'Active methylotrophs in the sediments of Lonar Lake, a saline and alkaline ecosystem formed by meteor impact', *The ISME Journal*, 4(11), pp. 1470–1480. doi: 10.1038/ismej.2010.70.

Apprill, A. *et al.* (2015) 'Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton', *Aquatic Microbial Ecology*, 75(2), pp. 129–137. doi: 10.3354/ame01753.

Arfman, N. *et al.* (1989) 'Methanol metabolism in thermotolerant methylotrophic *Bacillus* strains involving a novel catabolic NAD-dependent methanol dehydrogenase as a key enzyme', *Archives of Microbiology*, 152, pp. 280–288.

Arndt, D. *et al.* (2016) 'PHASTER: a better, faster version of the PHAST phage search tool', *Nucleic Acids Research*, 44(W1), pp. W16–W21. doi: 10.1093/nar/gkw387.

Azam, F. *et al.* (1983) 'The Ecological Role of Water-Column Microbes in the Sea', *Marine Ecology Progress Series*, 10, pp. 257–263. doi: 10.3354/meps010257.

- Bamford, D. H., Palva, E. T. and Lounatmaa, K. (1976) 'Ultrastructure and Life Cycle of the Lipid-containing Bacteriophage $\phi 6$ ', *Journal of General Virology*, 32, pp. 249–259.
- Barber, R. D. and Donohue, T. J. (1998) 'Function of a glutathione-dependent formaldehyde dehydrogenase in *Rhodobacter sphaeroides* formaldehyde oxidation and assimilation', *Biochemistry*, 37(2), pp. 530–537. doi: 10.1021/bi971463t.
- Barone, R. and Naselli-Flores, L. (2003) 'Distribution and seasonal dynamics of Cryptomonads in Sicilian water bodies', 502(i), pp. 325–329. doi: 10.1023/B.
- Bassham, J. A., Benson, A. A. and Calvin, M. (1950) 'The path of carbon in photosynthesis', *Journal of Biological Chemistry*, 185(2), pp. 781–787.
- Beale, R. *et al.* (2011) 'Quantification of oxygenated volatile organic compounds in seawater by membrane inlet-proton transfer reaction/mass spectrometry', *Analytica Chimica Acta*. Elsevier B.V., 706(1), pp. 128–134. doi: 10.1016/j.aca.2011.08.023.
- Beale, R. *et al.* (2013) 'Methanol, acetaldehyde, and acetone in the surface waters of the Atlantic Ocean', *Journal of Geophysical Research: Oceans*, 118(10), pp. 5412–5425. doi: 10.1002/jgrc.20322.
- Beale, R. *et al.* (2015) 'Annual study of oxygenated volatile organic compounds in UK shelf waters', *Marine Chemistry*. Elsevier B.V., 171, pp. 96–106. doi: 10.1016/j.marchem.2015.02.013.
- Beck, D. A. C. C. *et al.* (2015) 'Multiphyletic origins of methylotrophy in Alphaproteobacteria, exemplified by comparative genomics of Lake Washington isolates', *Environmental Microbiology*, 17(3), pp. 547–554. doi: 10.1111/1462-2920.12736.

- Bergh, Ø. *et al.* (1989) 'High abundance of viruses found in aquatic environments', *Nature*, 340, pp. 467–468. doi: 10.1038/340301a0.
- Boden, R. *et al.* (2010) 'Oxidation of dimethylsulfide to tetrathionate by *Methylophaga thiooxidans* sp. nov.: A new link in the sulfur cycle', *Environmental Microbiology*, 12(10), pp. 2688–2699. doi: 10.1111/j.1462-2920.2010.02238.x.
- Boden, R. (2019) 'Methylophaga', in Bergey's Manual Trust (ed.) *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Inc., p. 13. doi: 10.1002/9781118960608.gbm01218.pub2.
- Breitbart, M. *et al.* (2007) 'Exploring the Vast Diversity of Marine Viruses', *Oceanography*, 20(2), pp. 135–139. doi: 10.5670/oceanog.2007.58.
- Breitbart, M. (2012) 'Marine Viruses: Truth or Dare', *Annual Review of Marine Science*, 4(1), pp. 425–448. doi: 10.1146/annurev-marine-120709-142805.
- Breitbart, M. *et al.* (2018) 'Phage puppet masters of the marine microbial realm', *Nature Microbiology*. Springer US, 3(7), pp. 754–766. doi: 10.1038/s41564-018-0166-y.
- Brussaard, C. P. D. *et al.* (1996) 'Virus-like particles in a summer bloom of *Emiliana huxleyi* in the North Sea', *Aquatic Microbial Ecology*, 10(1992), pp. 105–113. doi: 10.3354/ame010105.
- Brussaard, C. P. D. (2006) 'Enumeration of bacteriophages using flow cytometry', in *Virus Ecology Workshop*, pp. 32–45.
- Brussaard, C. P. D. *et al.* (2008) 'Virioplankton dynamics and virally induced phytoplankton lysis versus microzooplankton grazing southeast of the Kerguelen (Southern Ocean)', *Deep-Sea Research Part II: Topical Studies in Oceanography*,

55(5–7), pp. 752–765. doi: 10.1016/j.dsr2.2007.12.034.

Brussaard, C. P. D. (2009) 'Enumeration of bacteriophages using flow cytometry', *Methods Mol.Biol.*, 501(1064-3745 (Print)), pp. 97–111.

Brussaard, C. P. D., Marie, D. and Bratbak, G. (2000) 'Flow cytometric detection of viruses', *Journal of Virological Methods*, 85(1–2), pp. 175–182. doi: 10.1016/S0166-0934(99)00167-6.

Buchan, A. *et al.* (2014) 'Master recyclers: features and functions of bacteria associated with phytoplankton blooms.', *Nature reviews. Microbiology*, 12(10), pp. 686–698. doi: 10.1038/nrmicro3326.

Bunse, C. and Pinhassi, J. (2017) 'Marine Bacterioplankton Seasonal Succession Dynamics', *Trends in Microbiology*. Elsevier Ltd, 25(6), pp. 494–505. doi: 10.1016/j.tim.2016.12.013.

Callahan, B. J. *et al.* (2016) 'DADA2: High-resolution sample inference from Illumina amplicon data', *Nature Methods*, 13(7), pp. 581–583. doi: 10.1038/nmeth.3869.

Calvo-Díaz, A. and Morán, X. A. G. (2009) 'Empirical leucine-to-carbon conversion factors for estimating heterotrophic bacterial production: Seasonality and predictability in a temperate coastal ecosystem', *Applied and Environmental Microbiology*, 75(10), pp. 3216–3221. doi: 10.1128/AEM.01570-08.

Caporaso, J. G. *et al.* (2010) 'QIIME allows analysis of high-throughput community sequencing data', *Nature Methods*, 7(5), pp. 1–12. doi: 10.1038/nmeth.f.303.QIIME.

Casjens, S. R. and Hendrix, R. W. (2015) 'Bacteriophage lambda: Early pioneer and still relevant', *Virology*, 479–480, pp. 310–330. doi: 10.1016/j.virol.2015.02.010.

Cenens, W. *et al.* (2013) 'Phage-host interactions during pseudolysogeny: Lessons from the Ptd/dgo interaction.', *Bacteriophage*, 3(1), p. e25029. doi: 10.4161/bact.25029.

Cervený, K. E. *et al.* (2002) 'Phage Therapy of Local and Systemic Disease Caused by *Vibrio vulnificus* in Iron-Dextran-Treated Mice', *Clinical and diagnostic laboratory immunology*, 70(11), pp. 6251–6262. doi: 10.1128/IAI.70.11.6251–6262.2002.

Charlson, R. J. *et al.* (1987) 'Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate', *Nature*, 326, pp. 655–661. doi: 10.1038/326655a0.

Chen, F. *et al.* (2001) 'Application of Digital Image Analysis and Flow Cytometry To Enumerate Marine Viruses Stained with SYBR Gold', *Applied and Environmental Microbiology*, 67(2), pp. 539–545. doi: 10.1128/AEM.67.2.539-545.2001.

Chen, Y. (2012) 'Comparative genomics of methylated amine utilization by marine *Roseobacter* clade bacteria and development of functional gene markers (*tmm*, *gmaS*)', *Environmental Microbiology*. doi: 10.1111/j.1462-2920.2012.02765.x.

Chistoserdova, L. (2011) 'Modularity of methylotrophy, revisited', *Environmental Microbiology*, 13(10), pp. 2603–2622. doi: 10.1111/j.1462-2920.2011.02464.x.

Chistoserdova, L. (2016) 'Lanthanides: New life metals?', *World Journal of Microbiology and Biotechnology*. Springer Netherlands, 32(8), pp. 1–7. doi: 10.1007/s11274-016-2088-2.

Chistoserdova, L. (2019) 'Methylotrophs and Methylotroph Communities', *Methylotrophs and Methylotroph Communities*, 5. doi: 10.21775/9781912530045.

Chistoserdova, L. and Kalyuzhnaya, M. G. (2018) 'Current Trends in Methylotrophy',

Trends in Microbiology. Elsevier Ltd, 26(8), pp. 703–714. doi: 10.1016/j.tim.2018.01.011.

Chistoserdova, L., Kalyuzhnaya, M. G. and Lidstrom, M. E. (2009) 'The Expanding World of Methylophilic Metabolism', *Annual review of microbiology*, 63(May), pp. 477–499. doi: 10.1146/annurev.micro.091208.073600.

Chistoserdova, L. and Lidstrom, M. E. (1997) 'Molecular and mutational analysis of a DNA region separating two methylophilic gene clusters in *Methylobacterium extorquens* AM1', *Microbiology*, (143), pp. 1729–1736.

Chu, F., Beck, D. A. C. C. and Lidstrom, M. E. (2016) 'MxaY regulates the lanthanide-mediated methanol dehydrogenase switch in *Methylomicrobium buryatense*', *PeerJ*, 2016(9), pp. 1–16. doi: 10.7717/peerj.2435.

Clarke, K. R. and Gorley, R. N. (2015) *Getting started with PRIMER v7*. Primer-E, Plymouth, UK. doi: 10.22201/ib.20078706e.2018.3.2409.

Clokier, M. R. J. *et al.* (2011) 'Phages in nature', *Bacteriophage*, 1(1), pp. 31–45. doi: 10.4161/bact.1.1.14942.

Clokier, M. R. J. and Kropinski, A. M. (eds) (2009) *Bacteriophages Methods and Protocols Volume 2: Molecular and Applied Aspects*, *Bacteriophages. Methods and Protocols. Volume 2: Molecular and Applied Aspects*. Humana Press, a part of Springer Science+Business Media. doi: 10.1007/978-1-60327-565-1.

Cotmore, S. F. *et al.* (2019) 'ICTV virus taxonomy profile: Parvoviridae', *Journal of General Virology*, 100(3), pp. 367–368. doi: 10.1099/jgv.0.001212.

Cottrell, M. T. and Kirchman, D. L. (2000) 'Community composition of marine bacterioplankton determined by 16S rRNA gene clone libraries and fluorescence in

situ hybridization', *Applied and Environmental Microbiology*, 66(12), pp. 5116–5122.

doi: 10.1128/AEM.66.12.5116-5122.2000.

Crocket, K. C. *et al.* (2018) 'Rare Earth Element Distribution in the NE Atlantic: Evidence for Benthic Sources, Longevity of the Seawater Signal, and Biogeochemical Cycling', *Frontiers in Marine Science*, 5(April), pp. 1–22. doi: 10.3389/fmars.2018.00147.

Deng, L. *et al.* (2014) 'Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space.', *Nature*. Nature Publishing Group, 513(7517), pp. 242–5. doi: 10.1038/nature13459.

van Dijken, J. P., Otto, R. and Harder, W. (1976) 'Growth of *Hansenula polymorpha* in a methanol-limited chemostat. Physiological responses due to the involvement of methanol oxidase as a key enzyme in methanol metabolism.', *Archives of microbiology*, 111(1–2), pp. 137–44.

Dixon, J. L., Beale, R. and Nightingale, P. D. (2010) 'Microbial methanol uptake in northeast Atlantic waters.', *The ISME journal*. Nature Publishing Group, 5(4), pp. 704–16. doi: 10.1038/ismej.2010.169.

Dixon, J. L., Beale, R. and Nightingale, P. D. (2011) 'Rapid biological oxidation of methanol in the tropical Atlantic: Significance as a microbial carbon source', *Biogeosciences*, 8(8), pp. 2707–2716. doi: 10.5194/bg-8-2707-2011.

Dixon, J. L., Beale, R. and Nightingale, P. D. (2013) 'Production of methanol, acetaldehyde, and acetone in the Atlantic Ocean', *Geophysical Research Letters*, 40(17), pp. 4700–4705. doi: 10.1002/grl.50922.

Dixon, J. L. and Nightingale, Philip D. (2012) 'Fine-scale variability in methanol

- uptake and oxidation: from the microlayer to 1000 m', *Biogeosciences*, 9(8), pp. 2961–2972. doi: 10.5194/bg-9-2961-2012.
- Dixon, J. L. and Nightingale, P D (2012) 'Fine scale variability in methanol uptake and oxidation in the micro-layer and near-surface waters of the Atlantic', *BGD Biogeosciences Discuss*, 9(9), pp. 4513–4542. doi: 10.5194/bgd-9-4513-2012.
- Doron, S. *et al.* (2016) 'Transcriptome dynamics of a broad host-range cyanophage and its hosts', pp. 1437–1455. doi: 10.1038/ismej.2015.210.
- Doyle, J. J. and Doyle, J. L. (1987) 'A rapid DNA isolation procedure for small quantities of fresh leaf tissue.', *Phytochemical Bulletin*, 19, pp. 11–15.
- Drake, L. A. *et al.* (1998) 'Vertical profiles of virus-like particles and bacteria in the water column and sediments of Chesapeake Bay, USA', *Aquatic Microbial Ecology*, 16(1), pp. 17–25. doi: 10.3354/ame016017.
- Ducklow, H. W. (2000) 'Bacterioplankton production and biomass in the oceans', *Microbial Ecology of the Oceans, 1st edition*, pp. 1–47.
- Duine, J. A., Frank, J. and van Zeeland, J. K. (1979) 'Glucose dehydrogenase from *Acinetobacter calcoaceticus*', *FEBS Letters*, 108(2), pp. 3–6.
- Dumont, M. G. (2010) 'Primers -functional genes for methylotrophs and methanotrophs', in *Handbook of Hydrocarbon and Lipid Microbiology*, p. 25.
- Duncan, B. N. (2003) 'Interannual and seasonal variability of biomass burning emissions constrained by satellite observations', *Journal of Geophysical Research*, 108(D2), p. 4100. doi: 10.1029/2002JD002378.
- Dupont, C. L. *et al.* (2012) 'Genomic insights to SAR86, an abundant and

uncultivated marine bacterial lineage', *The ISME Journal*. Nature Publishing Group, 6(6), pp. 1186–1199. doi: 10.1038/ismej.2011.189.

Elliott, S. and Rowland, F. S. (1995) 'Methyl halide hydrolysis rates in natural waters', *Journal of Atmospheric Chemistry*, 20(3), pp. 229–236. doi: 10.1007/BF00694495.

Eren, A. M. *et al.* (2013) 'Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data', *Methods in Ecology and Evolution*, 4(12), pp. 1111–1119. doi: 10.1111/2041-210X.12114.

Evans, C. *et al.* (2003) 'Direct estimates of the contribution of viral lysis and microzooplankton grazing to the decline of a *Micromonas* spp. population', *Aquatic Microbial Ecology*, 30(3), pp. 207–219. doi: 10.3354/ame030207.

Evans, C. *et al.* (2011) 'Potential climate change impacts on microbial distribution and carbon cycling in the Australian Southern Ocean', *Deep-Sea Research Part II: Topical Studies in Oceanography*. Elsevier, 58(21–22), pp. 2150–2161. doi: 10.1016/j.dsr2.2011.05.019.

Evans, C. *et al.* (2017) 'Drivers of interannual variability in virioplankton abundance at the coastal western Antarctic peninsula and the potential effects of climate change', *Environmental Microbiology*, 19(2), pp. 740–755. doi: 10.1111/1462-2920.13627.

Firsova, Y. E., Torgonskaya, M. L. and Trotsenko, Y. A. (2015) 'Functionality of the *xoxF* gene in *Methylobacterium dichloromethanicum* DM4', *Microbiology*, 84(6), pp. 796–803. doi: 10.1134/S002626171506003X.

Fitriyanto, N. A. *et al.* (2011) 'Molecular structure and gene analysis of Ce3+-

induced methanol dehydrogenase of *Bradyrhizobium* sp. MAFF211645', *Journal of Bioscience and Bioengineering*. The Society for Biotechnology, Japan, 111(6), pp. 613–617. doi: 10.1016/j.jbiosc.2011.01.015.

Flombaum, P. *et al.* (2013) 'Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*', *Proceedings of the National Academy of Sciences of the United States of America*, 110(24), pp. 9824–9829. doi: 10.1073/pnas.1307701110.

Fogg, P. C. M. *et al.* (2011) 'Characterization of a newly discovered Mu-like bacteriophage, RcapMu, in *Rhodobacter capsulatus* strain SB1003', *Virology*. Elsevier Inc., 421(2), pp. 211–221. doi: 10.1016/j.virol.2011.09.028.

Fu, Y. *et al.* (2013) 'Water mass and depth determine the distribution and diversity of Rhodobacterales in an Arctic marine system', *FEMS Microbiology Ecology*, 84(3), pp. 564–576. doi: 10.1111/1574-6941.12085.

Fuhrman, J. A. (1992) 'Bacterioplankton roles in cycling of organic matter: the microbial food web', *In Primary productivity and biogeochemical cycles in the sea*, 43, pp. 361–383.

Fuhrman, J. A. (1999) 'Marine viruses and their biogeochemical and ecological effects.', *Nature*, 399(6736), pp. 541–8. doi: 10.1038/21119.

Fuhrman, J. A. and Azam, F. (1980) 'Bacterioplankton secondary production estimates for coastal waters of British Columbia, Antarctica, and California', *Applied and Environmental Microbiology*, 39(6), pp. 1085–1095.

Fuhrman, J. a. and Noble, R. T. (1995) 'Viruses and protists cause similar bacterial mortality in coastal seawater', *Limnology and Oceanography*, 40(7), pp. 1236–1242.

doi: 10.4319/lo.1995.40.7.1236.

Fuhrman, J. A. and Suttle, C. A. (1993) 'Viruses in marine planktonic systems', *Oceanography*, 6(2), pp. 51–63.

Galbally, I. E. and Kirstine, W. (2002) 'The production of methanol by flowering plants and the global cycle of methanol', *Journal of Atmospheric Chemistry*, 43, pp. 195–229. doi: 10.1023/A:1020684815474.

Garoff, H., Hewson, R. and Opstelten, D. J. (1998) 'Virus Maturation by Budding', *Microbiology and molecular biology reviews : MMBR*, 62(4), pp. 1171–90.

Ghabrial, S. A. *et al.* (2018) 'ICTV Virus taxonomy profile: Chrysoviridae', *Journal of General Virology*, 99(1), pp. 19–20. doi: 10.1099/jgv.0.000994.

Ghosh, M. *et al.* (1995) 'The refined structure of the quinoprotein methanol dehydrogenase from *Methylobacterium extorquens* at 1.94 Å', *Structure*, 3(2), pp. 177–187. doi: 10.1016/S0969-2126(01)00148-4.

Gilbert, J. A. *et al.* (2009) 'The seasonal structure of microbial communities in the Western English Channel', *Environmental Microbiology*, 11(12), pp. 3132–3139. doi: 10.1111/j.1462-2920.2009.02017.x.

Gilbert, J. A. *et al.* (2012) 'Defining seasonal marine microbial community dynamics', *ISME Journal*, 6(2), pp. 298–308. doi: 10.1038/ismej.2011.107.

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

Giovannoni, S. J. *et al.* (1990) 'Genetic diversity in Sargasso Sea bacterioplankton',

Nature, 345, pp. 60–63.

Giovannoni, S. J. *et al.* (2005) 'Genome Streamlining in a Cosmopolitan Oceanic Bacterium', 309, pp. 1242–1245. doi: 10.1126/science.1114057.

Giovannoni, S. J. *et al.* (2008) 'The small genome of an abundant coastal ocean methylotroph', *Environmental Microbiology*, 10(7), pp. 1771–1782. doi: 10.1111/j.1462-2920.2008.01598.x.

Giovannoni, S. J., Thrash, J. C. and Temperton, B. (2014) 'Implications of streamlining theory for microbial ecology.', *The ISME journal*. Nature Publishing Group, 8(8), pp. 1–13. doi: 10.1038/ismej.2014.60.

Giovannoni, S. J. and Vergin, K. L. (2012) 'Seasonality in Ocean Microbial Communities', *Science*, 335(6069), pp. 671–676. doi: 10.1126/science.1198078.

Gobler, C. J. *et al.* (1997) 'Release and bioavailability of C, N, P, Se and Fe following viral lysis of a marine chrysophyte', *Limnology and Oceanography*, 42(7), pp. 1492–1504.

Goddard, V. J. *et al.* (2005) 'Temporal distribution of viruses, bacteria and phytoplankton throughout the water column in a freshwater hypereutrophic lake', *Aquatic Microbial Ecology*, 39(3), pp. 211–223. doi: 10.3354/ame039211.

Goldsmith, D. B. *et al.* (2011) 'Development of *phoH* as a novel signature gene for assessing marine phage diversity.', *Applied and environmental microbiology*, 77(21), pp. 7730–9. doi: 10.1128/AEM.05531-11.

Goldstein, S. J. and Jacobsen, S. B. (1988) 'REE in the Great Whale River estuary, northwest Quebec', *Earth and Planetary Science Letters*, 88(3–4), pp. 241–252. doi: 10.1016/0012-821X(88)90081-7.

Gómez-Consarnau, L. *et al.* (2012) 'Structuring of bacterioplankton communities by specific dissolved organic carbon compounds', *Environmental Microbiology*. doi: 10.1111/j.1462-2920.2012.02804.x.

Gonzalez, J. M. *et al.* (1997) 'Sagittula stellata gen. nov., sp. nov., a lignin-transforming bacterium from a coastal environment', *International Journal of Systematic Bacteriology*, 47(3), pp. 773–780. doi: 10.1099/00207713-47-3-773.

Goodwin, K. D. *et al.* (2001) 'Consumption of Tropospheric Levels of Methyl Bromide by C1 Compound-Utilizing Bacteria and Comparison to Saturation Kinetics', *Applied and Environmental Microbiology*, 67(12), pp. 5437–5443. doi: 10.1128/AEM.67.12.5437-5443.2001.

Grob, C. *et al.* (2013) 'Elemental composition of natural populations of key microbial groups in Atlantic waters', *Environmental Microbiology*, 15(11), pp. 3054–3064. doi: 10.1111/1462-2920.12145.

Grob, C. *et al.* (2015) 'Combining metagenomics with metaproteomics and stable isotope probing reveals metabolic pathways used by a naturally occurring marine methylotroph', *Environmental Microbiology*. doi: 10.1111/1462-2920.12935.

Guenther, A. (2002) 'The contribution of reactive carbon emissions from vegetation to the carbon balance of terrestrial ecosystems', *Chemosphere*, 49(8), pp. 837–844. doi: 10.1016/S0045-6535(02)00384-3.

Halsey, K. H. *et al.* (2017) 'Biological cycling of volatile organic carbon by phytoplankton and bacterioplankton', *Limnology and Oceanography*, 62(6), pp. 2650–2661. doi: 10.1002/lno.10596.

Halsey, K. H., Carter, A. E. and Giovannoni, S. J. (2012) 'Synergistic metabolism of

a broad range of C1 compounds in the marine methylotrophic bacterium HTCC2181', *Environmental Microbiology*, 14(3), pp. 630–640. doi: 10.1111/j.1462-2920.2011.02605.x.

Hara, S., Terauchi, K. and Koike, I. (1991) 'Abundance of viruses in marine waters: Assessment by epifluorescence and transmission electron microscopy', *Applied and Environmental Microbiology*, 57(9), pp. 2731–2734.

Heikes, B. G. (2002) 'Atmospheric methanol budget and ocean implication', *Global Biogeochemical Cycles*, 16(4), pp. 1–13. doi: 10.1029/2002GB001895.

Heldal, M. and Bratbak, G. (1991) 'Production and decay of viruses in aquatic.pdf', *Marine Ecology Progress Series*, 72, pp. 205–212.

Heptinstall, J. and Quayle, J. R. (1970) 'Pathways leading to and from serine during growth of *Pseudomonas* AM1 on C1 compounds or succinate.', *The Biochemical journal*, 117(3), pp. 563–572.

Hibi, Y. *et al.* (2011) 'Molecular structure of La³⁺-induced methanol dehydrogenase-like protein in *Methylobacterium radiotolerans*', *Journal of Bioscience and Bioengineering*. The Society for Biotechnology, Japan, 111(5), pp. 547–549. doi: 10.1016/j.jbiosc.2010.12.017.

Higgins, J. L. (2005) 'Virus dynamics in high-nutrient, low-chlorophyll marine surface waters', (May), p. 80.

Hill, R. W. *et al.* (1998) 'Virus-mediated total release of dimethylsulfoniopropionate from marine phytoplankton: A potential climate process', *Aquatic Microbial Ecology*, 14(1), pp. 1–6. doi: 10.3354/ame014001.

Hobbs, Z. and Abedon, S. T. (2016) 'Diversity of phage infection types and

associated terminology: the problem with “Lytic or lysogenic”, *FEMS microbiology letters*, 363(7), pp. 1–8. doi: 10.1093/femsle/fnw047.

Holmfeldt, K. *et al.* (2013) ‘Twelve previously unknown phage genera are ubiquitous in global oceans’, *Proceedings of the National Academy of Sciences*, 110(31), pp. 12798–12803. doi: 10.1073/pnas.1305956110.

Hoppe, H. G. *et al.* (2002) ‘Bacterial growth and primary production along a north-south transect of the Atlantic Ocean’, *Nature*, 416(6877), pp. 168–171. doi: 10.1038/416168a.

Howard-varona, C. *et al.* (2017) ‘Regulation of infection efficiency in a globally abundant marine Bacteriodes virus’. Nature Publishing Group, pp. 284–295. doi: 10.1038/ismej.2016.81.

Howard-Varona, C. *et al.* (2017) ‘Lysogeny in nature: mechanisms, impact and ecology of temperate phages’, *The ISME Journal*. Nature Publishing Group, 11(7), pp. 1–10. doi: 10.1038/ismej.2017.16.

Howat, A. M. (2017) *Characterisation of novel methylotrophs and the role of xoxF in coastal marine environments*. University of East Anglia.

Howat, A. M. *et al.* (2019) ‘Comparative genomics and mutational analysis reveals a novel XoxF-utilising methylotroph in the Roseobacter group isolated from the marine environment’, *Frontiers in Microbiology*, 9(APR), pp. 1–12. doi: 10.3389/fmicb.2018.00766.

Huang, J. *et al.* (2019) ‘Rare earth element alcohol dehydrogenases widely occur among globally distributed, numerically abundant and environmentally important microbes’, *The ISME Journal*. Springer US, pp. 20–22. doi: 10.1038/s41396-019-

0414-z.

Hurwitz, B. L., Brum, J. R. and Sullivan, M. B. (2015) 'Depth-stratified functional and taxonomic niche specialization in the " core " and " flexible " Pacific Ocean Virome', pp. 472–484. doi: 10.1038/ismej.2014.143.

Hurwitz, B. L. and Ren, J. M. U. (2016) 'Viral metabolic reprogramming in marine ecosystems', *Current Opinion in Microbiology*. Elsevier Ltd, 31, pp. 161–168. doi: 10.1016/j.mib.2016.04.002.

Hyman, P. and Abedon, S. T. (2010) 'Bacteriophage host range and bacterial resistance.', in *Advances in applied microbiology*. 1st edn. Elsevier Inc., pp. 217–248. doi: 10.1016/S0065-2164(10)70007-1.

Ichikawa, T., Tahara, T. and Hoshino., J. (1977) 'Isolation and characterisation of bacteriophages active against methanol assimilating bacteria', *Niigata Research Laboratory, Mitsubishi Gas Chemical Company Ltd, Japan*.

Iguchi, H., Yurimoto, H. and Sakai, Y. (2015) 'Interactions of Methylootrophs with Plants and Other Heterotrophic Bacteria', *Microorganisms*, 3(2), pp. 137–151. doi: 10.3390/microorganisms3020137.

Iriberry, J. *et al.* (1985) 'Heterotrophic bacterial activity in coastal waters: Functional relationship of temperature and phytoplankton population', *Ecological Modelling*, 28(1–2), pp. 113–120. doi: 10.1016/0304-3800(85)90016-X.

Janda, J. M. and Abbott, S. L. (2007) '16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls', *Journal of Clinical Microbiology*, 45(9), pp. 2761–2764. doi: 10.1128/JCM.01228-07.

Janvier, M. *et al.* (1985) 'Methylophaga marina gen. nov., sp. nov. and

Methylophaga thalassica sp. nov., Marine Methylophages', *International Journal of Systematic Microbiology*, 35(2), pp. 131–139.

Janvier, M., Regnault, B. and Grimont, P. (2003) 'Development and use of fluorescent 16S rRNA-targeted probes for the specific detection of Methylophaga species by in situ hybridization in marine sediments', *Research in Microbiology*, 154(7), pp. 483–490. doi: 10.1016/S0923-2508(03)00146-3.

Joint, I. and Smale, D. A. (2017) 'Marine heatwaves and optimal temperatures for microbial assemblage activity', *FEMS Microbiology Ecology*. Edited by R. Laanbroek, 93(2), p. fiw243. doi: 10.1093/femsec/fiw243.

Ju, K.-S. and Parales, R. E. (2010) 'Nitroaromatic Compounds, from Synthesis to Biodegradation', *Microbiology and Molecular Biology Reviews*, 74(2), pp. 250–272. doi: 10.1128/mnbr.00006-10.

Kalyuzhnaya, M. G., Hristova, K. R., *et al.* (2008) 'Characterization of a Novel Methanol Dehydrogenase in Representatives of Burkholderiales: Implications for Environmental Detection of Methylophagy and Evidence for Convergent Evolution', *JOURNAL OF BACTERIOLOGY*, 190(11), pp. 3817–3823. doi: 10.1128/JB.00180-08.

Kalyuzhnaya, M. G., Lapidus, A., *et al.* (2008) 'High-resolution metagenomics targets specific functional types in complex microbial communities', *Nature Biotechnology Letters*, 26(9), pp. 1029–1034. doi: 10.1038/nbt.1488.

Kalyuzhnaya, M. G. *et al.* (2012) 'Novel methylophagic isolates from lake sediment, description of Methylophaga versatilis sp. nov. and emended description of the genus Methylophaga', *INTERNATIONAL JOURNAL OF SYSTEMATIC AND*

EVOLUTIONARY MICROBIOLOGY, 62(1), pp. 106–111. doi: 10.1099/ijms.0.029165-0.

Kanagawa, T., Dazai, M. and Fukuoka, S. (1982) 'Degradation of 0, 0-Dimethyl Phosphorodithioate by Thiobacillus thioparus TK-1 and Pseudomonas AK-21 " DMDTP was degraded to orthophosphate by activated', 46(10), pp. 2571–2578.

Kelly, D. P., Ardley, J. K. and Wood, A. P. (2015) 'Cultivation of Methylotrophs', in *Hydrocarbon and Lipid Microbiology Protocols*, p. 43.

Keltjens, J. T. *et al.* (2014) 'PQQ-dependent methanol dehydrogenases: rare-earth elements make a difference', *Applied Microbiology and Biotechnology*, 98(14), pp. 6163–6183. doi: 10.1007/s00253-014-5766-8.

Kiene, R. P. and Bates, T. S. (1990) 'Biological removal of dimethyl sulphide from sea water', *Letters to Nature*, 345, pp. 702–705.

Kimman, S. A. and Brussaard, C. P. D. (2010) 'Estimation of viral-induced phytoplankton mortality using the modified dilution method', *Manual of Aquatic Viral Ecology*. Edited by S. Wilhelm, M. Weinbauer, and C. Suttle. American Society of Limnology and Oceanography, 7, pp. 65–73. doi: 10.4319/mave.2010.978-0-9845591-0-7.

Kirchman, D. L., Cottrell, M. T. and Lovejoy, C. (2010) 'The structure of bacterial communities in the western Arctic Ocean as revealed by pyrosequencing of 16S rRNA genes', *Environmental Microbiology*, 12(5), pp. 1132–1143. doi: 10.1111/j.1462-2920.2010.02154.x.

Kirchman, D. L. and Ducklow, H. W. (1993) 'Estimating Conversion Factors for the Thymidine and Leucine Methods for Measuring Bacterial Production', in Kemp, P.

F. et al. (eds) *Handbook of Methods in Aquatic Microbial Ecology*. Lewis Publishers, pp. 513–517. doi: 10.1201/9780203752746-60.

Kirchman, D. L., K'nees, E. and Hodson, R. E. (1985) 'Leucine incorporation and its potential as a measure of protein synthesis by bacteria in natural aquatic systems', *Applied and Environmental Microbiology*, 49(3), pp. 599–607.

Kist, J. and Tate, R. L. (2013) 'Phylogeny of bacterial methylotrophy genes reveals robustness in *Methylobacterium* mxaF sequences and mxa operon construction', *Soil Biology and Biochemistry*. Elsevier Ltd, 59, pp. 49–57. doi: 10.1016/j.soilbio.2012.12.010.

Kolb, S. (2009) 'Aerobic methanol-oxidizing Bacteria in soil', *FEMS Microbiology Letters*, 300(1), pp. 1–10. doi: 10.1111/j.1574-6968.2009.01681.x.

Kolb, S. and Stacheter, A. (2013) 'Prerequisites for amplicon pyrosequencing of microbial methanol utilizers in the environment', *Frontiers in Microbiology*, 4(SEP), pp. 1–12. doi: 10.3389/fmicb.2013.00268.

Kouno, K. et al. (1973) 'Isolation of new Methanol-utilising bacteria and its thiamine-requirement for growth', *Journal of General Applied Microbiology*, 19, pp. 11–21.

Kumar, S. et al. (2018) 'MEGA X: Molecular evolutionary genetics analysis across computing platforms', *Molecular Biology and Evolution*, 35(6), pp. 1547–1549. doi: 10.1093/molbev/msy096.

Laghdass, M. et al. (2012) 'High Contribution of SAR11 to Microbial Activity in the North West Mediterranean Sea', *Microbial Ecology*, 63(2), pp. 324–333. doi: 10.1007/s00248-011-9915-7.

Lamy, D. et al. (2010) 'Temporal changes of major bacterial groups and bacterial

heterotrophic activity during a *Phaeocystis globosa* bloom in the eastern English Channel', *Aquatic Microbial Ecology*, 58(1), pp. 95–107. doi: 10.3354/ame01359.

Lane, D. J. (1991) '16s/23S rRNA Sequencing', in *Nucleic Acid Techniques in Bacterial Systematic*, pp. 115–175.

Laukert, G. *et al.* (2019) 'Water mass transformation in the Barents Sea inferred from radiogenic neodymium isotopes, rare earth elements and stable oxygen isotopes', *Chemical Geology*. Elsevier, 511(August 2018), pp. 416–430. doi: 10.1016/j.chemgeo.2018.10.002.

Leak, D. J. and Dalton, H. (1986) 'Growth yields of methanotrophs', *Applied Microbiology and Biotechnology*, 23(6), pp. 470–476. doi: 10.1007/BF02346062.

Legendre, M. *et al.* (2014) 'Thirty-thousand-year-old distant relative of giant icosahedral DNA viruses with a pandoravirus morphology', *Proceedings of the National Academy of Sciences of the United States of America*, 111(11), pp. 4274–4279. doi: 10.1073/pnas.1320670111.

Lewis, A. C. *et al.* (2005) 'Sources and sinks of acetone, methanol, and acetaldehyde in North Atlantic marine air', *Atmospheric Chemistry and Physics*, 5(7), pp. 1963–1974. doi: 10.5194/acp-5-1963-2005.

Lidstrom, M. E. *et al.* (1994) 'New unified nomenclature for genes involved in the oxidation of methanol in Gram-negative bacteria', *FEMS Microbiology Letters*, (117), pp. 103–106.

Lidstrom, M. E., Fulton, G. L. and Ann E. Wopat (1983) "'Methylobacterium ethanolicum": a Syntrophic Association of Two Methylotrophic Bacteria', *Journal of General Microbiology*, 129, pp. 3139–3148.

- Lin, S. B., Wang, H. H. and Wu, C. S. (1983) 'Bdellovibrio in methylotrophic bacteria', *Biotechnology Letters*, 5(8), pp. 549–554.
- Lindell, D. *et al.* (2005) 'Photosynthesis genes in marine viruses yield proteins during host infection.', *Nature*, 438(7064), pp. 86–9. doi: 10.1038/nature04111.
- Macey, M. C. (2017) *Characterisation of methylotrophs in the rhizosphere*. University of East Anglia.
- Mackinder, L. C. M. *et al.* (2009) 'A unicellular algal virus, *Emiliana huxleyi* virus 86, exploits an animal-like infection strategy', *Journal of General Virology*, 90(9), pp. 2306–2316. doi: 10.1099/vir.0.011635-0.
- Madhaiyan, M. *et al.* (2010) 'Flavobacterium glycines sp. nov., a facultative methylotroph isolated from the rhizosphere of soybean', *International Journal of Systematic and Evolutionary Microbiology*, 60(9), pp. 2187–2192. doi: 10.1099/ijs.0.014019-0.
- Malin, G. *et al.* (1998) 'Elevated production of dimethylsulfide resulting from viral infection of cultures of *Phaeocystis pouchetii*', *Limnology and Oceanography*, 43(6), pp. 1389–1393. doi: 10.4319/lo.1998.43.6.1389.
- Marie, D. *et al.* (1999) 'Enumeration of Marine Viruses in Culture and Natural Samples by Flow Cytometry', *Applied and environmental microbiology*, 65(1), pp. 45–52.
- Martínez, J. M., Swan, B. K. and Wilson, W. H. (2014) 'Marine viruses, a genetic reservoir revealed by targeted viromics', *ISME Journal*, 8(5), pp. 1079–1088. doi: 10.1038/ismej.2013.214.
- Mcdonald, I. R. and Murrell, J. C. (1997) 'The Methanol Dehydrogenase Structural

Gene *mxoF* and Its Use as a Functional Gene Probe for Methanotrophs and Methylotrophs', *APPLIED AND ENVIRONMENTAL MICROBIOLOGY*, 63(8), pp. 3218–3224.

McLaren, M. R., Willis, A. D. and Callahan, B. J. (2019) 'Consistent and correctable bias in metagenomic sequencing experiments', *eLife*, 8, pp. 1–31. doi: 10.7554/elife.46923.

McMurdie, P. J. and Holmes, S. (2013) 'phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data', 8(4). doi: 10.1371/journal.pone.0061217.

Middelboe, M. (2000) 'Bacterial Growth Rate and Marine Virus-Host Dynamics.', *Microbial ecology*, 40(2), pp. 114–124. doi: 10.1007/s002480000050.

Middelboe, M. *et al.* (2003) 'Virus-induced transfer of organic carbon between marine bacteria in a model community', *Aquatic Microbial Ecology*, 33(1), pp. 1–10.

Middelboe, M., Jørgensen, N. O. G. and Kroer, N. (1996) 'Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton', *Applied and Environmental Microbiology*, 62(6), pp. 1991–1997.

Millard, A. D. *et al.* (2009) 'Comparative genomics of marine cyanomyoviruses reveals the widespread occurrence of *Synechococcus* host genes localized to a hyperplastic region: Implications for mechanisms of cyanophage evolution', *Environmental Microbiology*, 11(9), pp. 2370–2387. doi: 10.1111/j.1462-2920.2009.01966.x.

Millet, D. B. *et al.* (2008) 'New constraints on terrestrial and oceanic sources of atmospheric methanol', *Atmospheric Chemistry and Physics Discussions*, 8(2), pp.

7609–7655. doi: 10.5194/acpd-8-7609-2008.

Mincer, T. J. and Aicher, A. C. (2016) 'Methanol Production by a Broad Phylogenetic Array of Marine Phytoplankton', *PLOS ONE*. Edited by R. Rinnan, 11(3), p. 17. doi: 10.1371/journal.pone.0150820.

Mojica, K. D. A. and Brussaard, C. P. D. (2014) 'Factors affecting virus dynamics and microbial host - virus interactions in marine environments', *FEMS Microbiology Ecology*, 89(3), pp. 495–515. doi: 10.1111/1574-6941.12343.

Mojica, K. D. A., Evans, C. and Brussaard, C. P. D. (2014) 'Flow cytometric enumeration of marine viral populations at low abundances', *Aquatic Microbial Ecology*, 71(3), pp. 203–209. doi: 10.3354/ame01672.

Morris, R. M. *et al.* (2002) 'SAR11 clade dominates ocean surface bacterioplankton communities.', *Nature*, 420(December), pp. 806–810. doi: 10.1038/nature01240.

Morris, R. M., Longnecker, K. and Giovannoni, S. J. (2006) 'Pirellula and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom', *Environmental Microbiology*, 8(8), pp. 1361–1370. doi: 10.1111/j.1462-2920.2006.01029.x.

Moussard, H. *et al.* (2009) 'Identification of active methylotrophic bacteria inhabiting surface sediment of a marine estuary', *Environmental Microbiology Reports*, 1(5), pp. 424–433. doi: 10.1111/j.1758-2229.2009.00063.x.

Mühlencoert, E. and Müller, P. (2002) 'A novel two-component system of *Bradyrhizobium japonicum*: ElmS and ElmR are encoded in diverse orientations', *Mitochondrial DNA*, 13(2), pp. 93–102. doi: 10.1080/10425170290030006.

Murray, A. G. and Jackson, G. A. (1992a) 'Viral dynamics: a model of the effects of

size, shape, motion and abundance of single-celled planktonic organisms and other particles', *Marine Ecology Progress Series*, 89(2–3), pp. 103–116. doi: 10.3354/meps089103.

Murray, A. G. and Jackson, G. A. (1992b) 'Viral dynamics: A model of the effects size, shape, motion and abundance of single-celled planktonic organisms and other particles', *Marine Ecology Progress Series*, 89(2), pp. 103–116.

Murrell, J. C. and McDonald, I. R. (2000) 'Methylotrophs', in *Encyclopaedia of Microbiology*. 3rd edn. New York: Academic Press, pp. 245–255.

Nagasaki, K. *et al.* (2009) 'Isolation and characterization of a novel virus infecting *Teleaulax amphioxeia* (Cryptophyceae)', *Plankton and Benthos Research*, 4(3), pp. 122–124. doi: 10.3800/pbr.4.122.

Nakagawa, T. *et al.* (2012) 'A Catalytic Role of XoxF1 as La³⁺-Dependent Methanol Dehydrogenase in *Methylobacterium extorquens* Strain AM1', *PLoS ONE*, 7(11), pp. 1–7. doi: 10.1371/journal.pone.0050480.

Nercessian, O. *et al.* (2005) 'Bacterial Populations Active in Metabolism of C 1 Compounds in the Sediment of Lake Washington , a Freshwater Lake Bacterial Populations Active in Metabolism of C 1 Compounds in the Sediment of Lake Washington , a Freshwater Lake', *Society*, 71(11), pp. 6885–6899. doi: 10.1128/AEM.71.11.6885.

Neufeld, J. D. *et al.* (2007) 'Stable-isotope probing implicates *Methylophaga* spp and novel Gammaproteobacteria in marine methanol and methylamine metabolism', *The ISME Journal*, 1(6), pp. 480–491. doi: 10.1038/ismej.2007.65.

Neufeld, J. D. *et al.* (2008) 'Marine methylotrophs revealed by stable-isotope

probing, multiple displacement amplification and metagenomics', *Environmental Microbiology*, 10(6), pp. 1526–1535. doi: 10.1111/j.1462-2920.2008.01568.x.

Nightingale, P. D. (1991) *Low Molecular Weight Halocarbons in Seawater*. University of East Anglia.

Nunn, D. N. and Lidstrom, M. E. (1986a) 'Isolation and complementation analysis of 10 methanol oxidation mutant classes and identification of the methanol dehydrogenase structural gene of *Methylobacterium* sp. strain AM1', *Journal of Bacteriology*, 166(2), pp. 581–590.

Nunn, D. N. and Lidstrom, M. E. (1986b) 'Phenotypic characterization of 10 methanol oxidation mutant classes in *Methylobacterium* sp. strain AM1', *Journal of Bacteriology*, 166(2), pp. 591–597.

Obeng, N., Pratama, A. A. and Elsas, J. D. van (2016) 'The Significance of Mutualistic Phages for Bacterial Ecology and Evolution', *Trends in Microbiology*. Elsevier Ltd, 24(6), pp. 440–449. doi: 10.1016/j.tim.2015.12.009.

Obernosterer, I. and Herndl, G. J. (1995) 'Phytoplankton extracellular release and bacterial growth: Dependence on the inorganic N:P ratio', *Marine Ecology Progress Series*, 116(1–3), pp. 247–258. doi: 10.3354/meps116247.

Ochsner, A. M. *et al.* (2015) 'Methylobacterium extorquens: methylotrophy and biotechnological applications', *Applied Microbiology and Biotechnology*, 99(2), pp. 517–534. doi: 10.1007/s00253-014-6240-3.

Oikawa, P. Y. *et al.* (2011) 'Leaf and root pectin methyltransferase activity and ¹³C/¹²C stable isotopic ratio measurements of methanol emissions give insight into methanol production in *Lycopersicon esculentum*', *New Phytologist*, 191(4), pp.

1031–1040. doi: 10.1111/j.1469-8137.2011.03770.x.

Oikawa, P. Y. and Lerdau, M. T. (2013) 'Catabolism of volatile organic compounds influences plant survival', *Trends in Plant Science*. Elsevier Ltd, 18(12), pp. 695–703. doi: 10.1016/j.tplants.2013.08.011.

Oki, T., Nishida, H. and Ozaki, A. (1972) 'Deoxyribonucleic Acid Bacteriophage of *Methanomonas methylovora*', *Journal of Virology*, 9(3), pp. 544–546.

Ortmann, A. C. and Suttle, C. A. (2009) 'Determination of Virus Abundance by Epifluorescence Microscopy', in Clokie, M. R. J. and Kropinski, A. M. (eds). Humana Press, pp. 87–95. doi: 10.1007/978-1-60327-164-6.

Øvreås, L. *et al.* (2003) 'Response of bacterial and viral communities to nutrient manipulations in seawater mesocosms', *Aquatic Microbial Ecology*, 31, pp. 109–121.

Paul, J. H. (2008) 'Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas?', *The ISME journal*, 2(6), pp. 579–589. doi: 10.1038/ismej.2008.35.

Paul, J. H. and Sullivan, M. B. (2005) 'Marine phage genomics: What have we learned?', *Current Opinion in Biotechnology*, 16(3 SPEC. ISS.), pp. 299–307. doi: 10.1016/j.copbio.2005.03.007.

Payet, J. P. and Suttle, C. A. (2008a) 'Physical and biological correlates of virus dynamics in the southern Beaufort Sea and Amundsen Gulf', *Journal of Marine Systems*. Elsevier B.V., 74(3–4), pp. 933–945. doi: 10.1016/j.jmarsys.2007.11.002.

Payet, J. P. and Suttle, C. A. (2008b) 'Physical and biological correlates of virus dynamics in the southern Beaufort Sea and Amundsen Gulf', *Journal of Marine*

- Systems*, 74(3–4), pp. 933–945. doi: 10.1016/j.jmarsys.2007.11.002.
- Personnic, S. *et al.* (2009) 'Seasonal and spatial variability of virio-, bacterio-, and picophytoplanktonic abundances in three peri-alpine lakes', *Hydrobiologia*, 627(1), pp. 99–116. doi: 10.1007/s10750-009-9718-8.
- Peyraud, R. *et al.* (2012) 'Co-Consumption of Methanol and Succinate by *Methylobacterium extorquens* AM1', *PLoS ONE*, 7(11), p. 11. doi: 10.1371/journal.pone.0048271.
- Pol, A. *et al.* (2014) 'Rare earth metals are essential for methanotrophic life in volcanic mudpots', *Environmental Microbiology*, 16(1), pp. 255–264. doi: 10.1111/1462-2920.12249.
- Poorvin, L. *et al.* (2004) 'Viral release of iron and its bioavailability to marine plankton', *Limnol. Oceanogr*, 49(5), pp. 1734–1741.
- Proctor, L. M. and Fuhrman, J. A. (1990) 'Viral mortality of marine bacteria and cyanobacteria', *Nature*, 343, pp. 60–62. doi: 10.1038/343060a0.
- Putzrath, R. M. and Maniloff, J. (1977) 'Growth of an enveloped mycoplasma virus and establishment of a carrier state.', *Journal of virology*, 22(2), pp. 308–14.
- Quast, C. *et al.* (2013) 'The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools', *Nucleic Acids Research*, 41(D1), pp. 590–596. doi: 10.1093/nar/gks1219.
- Radajewski, S. M. *et al.* (2002) 'Identification of active methylotroph populations in an acidic forest soil by stable-isotope probing', *Microbiology*, 148(8), pp. 2331–2342. doi: 10.1099/00221287-148-8-2331.

Ramachandran, A. and Walsh, D. A. (2015) 'Investigation of XoxF methanol dehydrogenases reveals new methylotrophic bacteria in pelagic marine and freshwater ecosystems', *FEMS Microbiology Ecology*, 91(10), pp. 1–10. doi: 10.1093/femsec/fiv105.

Rappé, M. S., Vergin, K. L. and Giovannoni, S. J. (2000) 'Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems', *FEMS Microbiology Ecology*, 33(3), pp. 219–232. doi: 10.1016/S0168-6496(00)00064-7.

Rastelli, E. *et al.* (2016) 'Quantification of viral and prokaryotic production rates in benthic ecosystems: A methods comparison', *Frontiers in Microbiology*, 7(September), pp. 1–11. doi: 10.3389/fmicb.2016.01501.

Redfield, A. C., Ketchum, B. H. and Richards, F. A. (1963) 'The influence of organisms on the composition of seawater', in Hill, M. N. (ed.) *The Sea*. 2nd edn. New York: Wiley, Interscience, pp. 26–77.

Riemer, D. D. (2000) *Marine and terrestrial sources of reactive volatile organic compounds and their impact on the tropospheric ozone chemistry of the earth*. University of Miami.

Robarts, R. D. (1998) *Incorporation of Radioactive Precursors Into Macromolecules as Measures of Bacterial Growth: Problems and Pitfalls*, *Molecular Approaches to the Study of the Ocean*. doi: 10.1007/978-94-011-4928-0_25.

Rodriguez, F. *et al.* (2000) 'Temporal variability of viruses, bacteria, phytoplankton and zooplankton in the western English Channel off Plymouth', *J. Mar Biol. Ass. U. K.*, 80(2000), pp. 575–586. doi: 10.1055/s-2008-1044745.

- Rohwer, F. *et al.* (2000) 'The complete genomic sequence of the marine phage Roseophage', *Limnol Oceanogr*, 45(2), pp. 408–418.
- Rohwer, F. and Thurber, R. V. (2009) 'Viruses manipulate the marine environment.', *Nature*, 459(7244), pp. 207–212. doi: 10.1038/nature08060.
- Russel, M. and Model, P. (2006) 'Filamentous phage.', in *The bacteriophages*, pp. 146–160.
- Salcher, M. M. *et al.* (2019) 'Evolution in action: habitat transition from sediment to the pelagial leads to genome streamlining in Methylophilaceae', *ISME Journal*. doi: 10.1038/s41396-019-0471-3.
- Sandaa, R. (2008) 'Burden or benefit? Virus-host interactions in the marine environment', *Research in Microbiology*, 159(5), pp. 374–381. doi: 10.1016/j.resmic.2008.04.013.
- Sandaa, R., Short, S. M. and Schroeder, D. C. (2010) 'Fingerprinting aquatic virus communities', *Manual of Aquatic Viral Ecology*, pp. 9–18. doi: 10.4319/mave.2010.978-0-9845591-0-7.9.
- Sargeant, S. L. (2013) *Microbial utilisation of methanol in seawater*. The University of Warwick.
- Sargeant, S. L. *et al.* (2016) 'Seasonal variability in microbial methanol utilisation in coastal waters of the western English Channel', *Marine Ecology Progress Series*, 550, pp. 53–64.
- Sargeant, S. L. *et al.* (2018) 'Basin-scale variability of microbial methanol uptake in the Atlantic Ocean', *Biogeosciences*, 15(16), pp. 5155–5167. doi: 10.5194/bg-15-5155-2018.

Sargeant, S. L. *et al.* (no date) 'Seasonal variability in microbial methanol utilisation in coastal waters of the western English Channel', *Marine Ecology Progress Series*.

Schliep, K. P. (2011) 'phangorn: phylogenetic analysis in R', 27(4), pp. 592–593. doi: 10.1093/bioinformatics/btq706.

Schlueter, A. *et al.* (2015) 'Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM Genome-guided insight into the methylotrophy of *Paracoccus aminophilus* JCM 7686', (July 2016). doi: 10.3389/fmicb.2015.00852.

Schoemann, V. *et al.* (2005) 'Phaeocystis blooms in the global ocean and their controlling mechanisms: A review', *Journal of Sea Research*, 53(1-2 SPEC. ISS.), pp. 43–66. doi: 10.1016/j.seares.2004.01.008.

Semrau, J. D., DiSpirito, A. A. and Murrell, J. C. (2008) 'Life in the extreme: thermoacidophilic methanotrophy', *Trends in Microbiology*, 16(5), pp. 190–193. doi: 10.1016/j.tim.2008.02.004.

Senior, P. J. and Windass, J. D. (1980) 'The 1Cl single cell protein process', *Biotechnology Letters*, 2(5), pp. 205–210.

Shibata, A. *et al.* (1997) 'Formation of submicron colloidal particles from marine bacteria by viral infection', *Marine Ecology Progress Series*, 155, pp. 303–307. doi: 10.3354/meps155303.

Sholkovitz, E. R. (1993) 'The geochemistry of rare earth elements in the Amazon River estuary', *Geochimica et Cosmochimica Acta*, 57(10), pp. 2181–2190. doi: 10.1016/0016-7037(93)90559-F.

Short, S. M., Chen, F. and Wilhelm, S. W. (2010) 'The construction and analysis of marker gene libraries', *Manual of Aquatic Viral Ecology*, pp. 82–91. doi:

10.4319/mave.2010.978-0-9845591-0-7.82.

Sigman, D. M. and Hain, M. P. (2012) *The Biological Productivity of the Ocean, Nature Education Knowledge*.

Simon, M. and Azam, F. (1989a) 'Protein content and protein synthesis rates of planktonic marine bacteria', *Marine Ecology Progress Series*, 51, pp. 201–213. doi: 10.3354/meps051201.

Simon, M. and Azam, F. (1989b) 'Protein content and protein synthesis rates of planktonic marine bacteria', *Marine Ecology Progress Series*, 51, pp. 201–213. doi: 10.3354/meps051201.

Singh, H. B. *et al.* (2000) 'Distribution and fate of selected oxygenated organic species in the troposphere and lower stratosphere over the Atlantic', *Journal of Geophysical Research*, 105(D3), p. 3795. doi: 10.1029/1999JD900779.

Siringan, P. *et al.* (2014) 'Alternative bacteriophage life cycles: the carrier state of *Campylobacter jejuni*.', *Open biology*, 4, p. 130200. doi: 10.1098/rsob.130200.

Skovran, E. *et al.* (2010) 'A Systems Biology Approach Uncovers Cellular Strategies Used by *Methylobacterium extorquens* AM1 During the Switch from Multi- to Single-Carbon Growth', *PLoS ONE*, 5(11), p. 16. doi: 10.1371/journal.pone.0014091.

Skovran, E. *et al.* (2011) 'XoxF Is Required for Expression of Methanol Dehydrogenase in *Methylobacterium extorquens* AM1', *JOURNAL OF BACTERIOLOGY*, 193(21), pp. 6032–6038. doi: 10.1128/JB.05367-11.

Smith, D. C. and Azam, F. (1992) 'A simple , economical method for measuring bacterial protein synthesis rates in seawater using ³H-leucine 1', 6(2), pp. 107–114.

Smyth, T. J. *et al.* (2010) 'A broad spatio-temporal view of the Western English Channel observatory', *Journal of Plankton Research*, 32(5), pp. 585–601. doi: 10.1093/plankt/fbp128.

Smyth, T. J. *et al.* (2015) 'The Western Channel Observatory', *Progress in Oceanography*, 137, pp. 335–341. doi: 10.1016/j.pocean.2015.05.020.

Spencer, R. (1955) 'A marine bacteriophage.', *Nature*, 175(4459), pp. 690–691. doi: 10.1038/175690a0.

Spring, S. *et al.* (2015) 'A taxonomic framework for emerging groups of ecologically important marine gammaproteobacteria based on the reconstruction of evolutionary relationships using genome-scale data', *Frontiers in Microbiology*, 6(APR), pp. 1–17. doi: 10.3389/fmicb.2015.00281.

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 Bacteriology*, 44(4), pp. 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', *Annual Review Microbiology*, 39(1), pp. 321–346. doi: 10.1146/annurev.micro.39.1.321.

Steward, G. F., Smith, D. C. and Azam, F. (1996) 'Abundance and production of bacteria and viruses in the Bering and Chukchi Seas', *Marine Ecology Progress Series*, 131(1–3), pp. 287–300. doi: 10.3354/meps131287.

Strøm, T., Ferenci, T. and Quayle, J. R. (1974) 'The carbon assimilation pathways of *Methylococcus capsulatus*, *Pseudomonas methanica* and *Methylosinus*

trichosporium (OB3B) during growth on methane.', *The Biochemical journal*, 144(3), pp. 465–476.

Sullivan, M. B. *et al.* (2005) 'Three Prochlorococcus cyanophage genomes: Signature features and ecological interpretations', *PLoS Biology*, 3(5), pp. 0790–0806. doi: 10.1371/journal.pbio.0030144.

Sullivan, M. B. *et al.* (2010) 'Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments', *Environmental Microbiology*, 12(11), pp. 3035–3056. doi: 10.1111/j.1462-2920.2010.02280.x.

Sullivan, M. B., Waterbury, J. B. and Chisholm, S. W. (2003) 'Cyanophages infecting the oceanic cyanobacterium Prochlorococcus', *Nature*, 426(6966), p. 584.

Sun, J. *et al.* (2011) 'One carbon metabolism in SAR11 pelagic marine bacteria', *PLoS ONE*, 6(8). doi: 10.1371/journal.pone.0023973.

Sun, S. *et al.* (2012) 'Structure and function of the small terminase component of the DNA packaging machine in T4-like bacteriophages', *PNAS*, 109(3), pp. 817–822. doi: 10.1073/pnas.1110224109.

Sundquist, W. I. and Krausslich, H.-G. (2012) 'HIV-1 Assembly, Budding, and Maturation', *Cold Spring Harbour Perspectives in Medicine*, 2, pp. 1–24.

Suttle, C. A. (1994) 'The significance of viruses to mortality in aquatic microbial communities', *Microbial Ecology*, 28(2), pp. 237–243. doi: 10.1007/BF00166813.

Suttle, C. A. (2005) 'Viruses in the sea.', *Nature*, 437, pp. 356–361.

Suttle, C. A. (2007) 'Marine viruses--major players in the global ecosystem.', *Nature*

Reviews Microbiology, 5, pp. 801–812.

Suttle, C. A. and Chan, A. M. (1993) 'Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-infectivity and growth characteristics', *Marine Ecology Progress Series*, 92(1–2), pp. 99–109. doi: 10.3354/meps092099.

Tait, K. *et al.* (2014) 'Spatio-temporal variability in ammonia oxidation and ammonia-oxidising bacteria and archaea in coastal sediments of the western English Channel', *Marine Ecology Progress Series*, 511, pp. 41–58. doi: 10.3354/meps10933.

Tarran, G. A. and Bruun, J. T. (2015) 'Nanoplankton and picoplankton in the Western English Channel: Abundance and seasonality from 2007-2013', *Progress in Oceanography*. Elsevier Ltd, 137, pp. 446–455. doi: 10.1016/j.pocean.2015.04.024.

Taubert, M. *et al.* (2015) 'XoxF encoding an alternative methanol dehydrogenase is widespread in coastal marine environments', *Environmental Microbiology*, 17(10), pp. 3937–3948. doi: 10.1111/1462-2920.12896.

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(6), pp. 1320–1328. doi: 10.4319/lo.2000.45.6.1320.

Thingstad, T. F. and Lignell, R. (1997) 'Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand', *Aquat. Microb. Ecol.*, 13(1), pp. 19–27. doi: 10.3354/ame013019.

- Thomas, R. *et al.* (2011) 'Viral abundance, production, decay rates and life strategies (lysogeny versus lysis) in Lake Bourget (France)', *Environmental Microbiology*, 13(3), pp. 616–630. doi: 10.1111/j.1462-2920.2010.02364.x.
- Thompson, L. R. *et al.* (2011) 'Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism', *Proceedings of the National Academy of Sciences*, 108(39), pp. E757–E764. doi: 10.1073/pnas.1102164108.
- Tie, X. (2003) 'Biogenic methanol and its impacts on tropospheric oxidants', *Geophysical Research Letters*, 30(17), pp. 3–6. doi: 10.1029/2003GL017167.
- Trotsenko, Y. A. and Murrell, J. C. (2008) 'Metabolic Aspects of Aerobic Obligate Methanotrophy', in *Metabolism of Aerobic Methanotrophs*, pp. 184–229. doi: 10.1016/S0065-2164(07)00005-6.
- Twist, R. Van and Kropinski, A. M. (2009) 'Bacteriophage Enrichment from Water and Soil', *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*, 501, pp. 15–21. doi: 10.1007/1-4020-3791-0_3.
- Tyutikov, F. M. *et al.* (1980) 'Bacteriophages of methanotrophic bacteria', *Journal of Bacteriology*, 144(1), pp. 375–381.
- Urakami, T. and Komagata, K. (1987) 'Characterization of Species of Marine Methylophages of the Genus Methylophaga', *International Journal of Systematic Bacteriology*, 37(4), pp. 402–406. doi: 10.1099/00207713-37-4-402.
- Villeneuve, C. *et al.* (2012) 'Complete genome sequences of Methylophaga sp. strain JAM1 and Methylophaga sp. strain JAM7.', *Journal of bacteriology*, 194(15), pp. 4126–7. doi: 10.1128/JB.00726-12.
- Vu, H. N. *et al.* (2016) 'Lanthanide-Dependent Regulation of Methanol Oxidation

Systems in *Methylobacterium extorquens* AM1 and Their Contribution to Methanol Growth', *Journal of Bacteriology*. Edited by W. W. Metcalf, 198(8), pp. 1250–1259. doi: 10.1128/JB.00937-15.

Vuilleumier, S. *et al.* (2009) 'Methylobacterium genome sequences: A reference blueprint to investigate microbial metabolism of C1 compounds from natural and industrial sources', *PLoS ONE*, 4(5). doi: 10.1371/journal.pone.0005584.

Wang, Q. *et al.* (2007) 'Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy', *Applied and Environmental Microbiology*, 73(16), pp. 5261–5267. doi: 10.1128/AEM.00062-07.

Watanabe, M., Kojima, H. and Fukui, M. (2015) 'Limnochorda pilosa gen. Nov, sp. Nov, a moderately thermophilic, facultatively anaerobic, pleomorphic bacterium and proposal of Limnochordaceae fam. Nov, Limnochordales ord. Nov. And Limnochordia classis nov. In the phylum Firmicutes', *International Journal of Systematic and Evolutionary Microbiology*, 65(8), pp. 2378–2384. doi: 10.1099/ijs.0.000267.

Weaver, C. A. and Lidstrom, M. E. (1985) 'Methanol Dissimilation in Xanthobacter H4-14: Activities, Induction and Comparison to Pseudomonas AM1 and Paracoccus denitrificans', *Journal of General Microbiology*, 131, pp. 2183–2197.

Weinbauer, M. G. (2004) 'Ecology of prokaryotic viruses', *FEMS Microbiology Reviews*, 28(2), pp. 127–181. doi: 10.1016/j.femsre.2003.08.001.

Weinbauer, M. G. and Suttle, C. A. (1999) 'Lysogeny and prophage induction in coastal and offshore bacterial communities', *Aquatic Microbial Ecology*, 18(3), pp. 217–225. doi: 10.3354/ame018217.

Weitz, J. S. *et al.* (2015) 'A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes', *The ISME Journal*. Nature Publishing Group, 9(6), pp. 1352–1364. doi: 10.1038/ismej.2014.220.

Weitz, J. S. and Wilhelm, S. W. (2012) 'Ocean viruses and their effects on microbial communities and biogeochemical cycles.', *F1000 biology reports*, 4(September), p. 17. doi: 10.3410/B4-17.

West, N. J. *et al.* (2016) 'Distinct spatial patterns of SAR11, SAR86, and Actinobacteria diversity along a transect in the ultra-oligotrophic South Pacific Ocean', *Frontiers in Microbiology*, 7(MAR), pp. 1–16. doi: 10.3389/fmicb.2016.00234.

White, S. *et al.* (1993) 'The Active Site Structure of the Calcium-Containing Quinoprotein Methanol Dehydrogenase', *Biochemistry*, 32(48), p. 4.

Wickham, H. (2009) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag, New York.

Widdel, F., Kohring, G.-W. and Mayer, F. (1983) 'Studies on dissimilatory sulfate-reducing bacteria that decompose fatty acids', *Archives of Microbiology*, 134(4), pp. 286–294. doi: 10.1007/BF00407804.

Wigington, C. H. *et al.* (2016) 'Re-examination of the relationship between marine virus and microbial cell abundances', *Nature Microbiology*. Nature Publishing Group, 1(3), pp. 4–11. doi: 10.1038/nmicrobiol.2015.24.

Wilhelm, S. W., Brigden, S. M. and Suttle, C. A. (2002) 'A Dilution Technique For The Direct Measurement Of Viral Production: A Comparison In Stratified And Tidally Mixed Coastal Waters', *Microbial Ecology*, 43(1), pp. 168–173. doi:

10.1007/s00248-001-1021-9.

Wilhelm, S. W. and Suttle, C. A. (1999) 'Viruses and Nutrient Cycles the aquatic food webs', *BioScience*, 49(10), pp. 781–788.

Williams, J. (2004) 'Measurements of organic species in air and seawater from the tropical Atlantic', *Geophysical Research Letters*, 31(October), p. L23S06. doi: 10.1029/2004GL020012.

Wilson, S. M., Gleisten, M. P. and Donohue, T. J. (2008) 'Identification of proteins involved in formaldehyde metabolism by *Rhodobacter sphaeroides*', *Microbiology*, 154(1), pp. 296–305. doi: 10.1038/jid.2014.371.

Wilson, W. H. *et al.* (2002) 'Isolation of viruses responsible for the demise of an *Emiliania huxleyi* bloom in the English Channel', *Journal of the Marine Biological Association of the United Kingdom*, (82), pp. 369–377. doi: 10.1017/S002531540200560X.

Wilson, W. H. *et al.* (2017) 'Genomic exploration of individual giant ocean viruses', *ISME Journal*. Nature Publishing Group, 11(8), pp. 1736–1745. doi: 10.1038/ismej.2017.61.

Windass, J. D. *et al.* (1980) 'Improved conversion of methanol to single-cell protein by *Methylophilus methylotrophus*.', *Nature*, 287(5781), pp. 396–401. doi: 10.1038/287396a0.

Winget, D. M. *et al.* (2005) 'Tangential flow diafiltration: an improved technique for estimation of virioplankton production', *Aquatic Microbial Ecology*, 41(3), pp. 221–232. doi: 10.3354/ame041221.

Wohl, C. *et al.* (2019) 'Segmented flow coil equilibrators coupled to a Proton Transfer

Reaction Mass Spectrometer for measurements of a broad range of Volatile Organic Compounds in seawater', *Ocean Science Discussions*, 15, pp. 925–940. doi: 10.5194/os-2019-5.

Wommack, K. E. *et al.* (2009) 'Methods for the Isolation of Viruses from Environmental Samples', in *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions*, pp. 3–14. doi: 10.1007/978-1-60327-164-6.

Wommack, K. E. *et al.* (2010) 'Filtration-based methods for the collection of viral concentrates from large water samples', *Manual of Aquatic Viral Ecology*, pp. 110–117. doi: 10.4319/mave.2010.978-0-9845591-0-7.110.

Wommack, K. E. and Colwell, R. R. (2000) 'Virioplankton: Viruses in Aquatic Ecosystems', *Microbiology and Molecular Biology Reviews*, 64(1), pp. 69–114. doi: 10.1128/MMBR.64.1.69-114.2000.

Wright, E. S. (2016) 'Using DECIPHER v2 . 0 to Analyze Big Biological Sequence Data in R', 8, pp. 352–359.

Wright, R. T. and Hobbie, J. E. (1965) 'The uptake of organic solutes in lake water', *Limnology & Oceanography*, 10, pp. 22–28.

Wu, L. *et al.* (2015) 'Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis', *BMC Microbiology*. *BMC Microbiology*, 15(1), pp. 1–12. doi: 10.1186/s12866-015-0450-4.

Yamamoto, M. *et al.* (1980) 'Identification of marine methanol-utilizing bacteria', *Journal of Fermentation Technology*, 58(2), pp. 99–106.

Yuan, Y. and Gao, M. (2017) 'Jumbo bacteriophages: An overview', *Frontiers in*

Microbiology, 8(403), pp. 1–9. doi: 10.3389/fmicb.2017.00403.

Zamora, J. L. R. and Aguilar, H. C. (2018) 'Flow virometry as a tool to study viruses', *Methods*. Elsevier Inc., 134–135, pp. 87–97. doi: 10.1016/j.ymeth.2017.12.011.

Zhao, Y. *et al.* (2013) 'Abundant SAR11 viruses in the ocean.', *Nature*. Nature Publishing Group, 494(7437), pp. 357–360. doi: 10.1038/nature11921.

Zhong, X. *et al.* (2014) 'Variations in Abundance, Genome Size, Morphology, and Functional Role of the Virioplankton in Lakes Annecy and Bourget over a 1-Year Period', *Microbial Ecology*, 67(1), pp. 66–82. doi: 10.1007/s00248-013-0320-2.

Zhou, Y. *et al.* (2011) 'PHAST: A Fast Phage Search Tool', *Nucleic Acids Research*, 39(SUPPL. 2), pp. 347–352. doi: 10.1093/nar/gkr485.

Zorz, J. *et al.* (2019) 'Drivers of regional bacterial community structure and diversity in the northwest atlantic ocean', *Frontiers in Microbiology*, 10(FEB), pp. 1–24. doi: 10.3389/fmicb.2019.00281.