MOLECULAR BASIS OF CYANOPHAGE RESISTANCE IN MARINE *SYNECHOCOCCUS*

BLANCA MICAELA PEREZ SEPULVEDA

A THESIS SUBMITTED FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

SCHOOL OF LIFE SCIENCES

UNIVERSITY OF WARWICK

NOVEMBER 2015

CONTENTS

FIGURES AND TABLES

ACKNOWLEDGEMENTS

I would like to thank my supervisor Prof. Dave Scanlan for all his support and good advice; I am eternally grateful for having been given this wonderful opportunity. I would also like to thank Andy Millard for all his patience and guidance; thanks for never giving up on encouraging me and for being a great mentor and friend.

Thanks to my advisory panel Kevin Purdy and Liz Wellington for the critical advice and encouragement when things were looking strange. Also, thanks to the University of Warwick and the Chancellor's International Scholarship for funding my PhD.

Many thanks to Selin, Fran, Branko, Joseph, Mar, Mark, Carolina, Rich and all past and current members of the Scanlan lab for all the good (and sometimes very challenging) criticism of my work, for have welcomed me in a strange country, in your lives and for accepting that sometimes I am 'too latina'.

I would like to thank my friends, who always encouraged and supported me when I decided to cross the Ocean, and to my family:

Mamá, Papá y Cochón: ha sido un camino largo y no siempre fácil, pero lo logramos! Por fin, luego de casi 30 años, creo que terminé mi educación. Gracias por aceptar mi tozudez y ayudarme a cumplir los más locos sueños.

Finally, I would like to thank Brian for resisting stoically the worst and more stressful stage with the most beautiful smile and sense of humour; I did it!

This thesis is dedicated to each and every one mentioned here.

DECLARATION

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

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

- *Synechococcus* sp. WH7803 and WH7805 cells used for qPCR assay optimisation were flow sorted by Dr. Frances Pitt, University of Warwick

- SNPs and Indels identified were parsed from mpileup file using a Perl script written by Dr. Andrew Millard, University of Warwick

Parts of this thesis have been published by the author:

Puxty RJ, **Perez-Sepulveda B**, Rihtman B, Evans DJ, Millard AD, Scanlan DJ (2015) Spontaneous deletion of an "ORFanage" region facilitates host adaptation in a "photosynthetic" cyanophage. *PLoS ONE* 10(7): e0132642

SUMMARY

Marine phytoplankton are responsible for ~50 % of global primary production and encompass a wide range of microorganisms characterized by being phototrophs. Within this group are the marine picocyanobacteria encompassing the phylogenetically closely related genera *Synechococcus* and *Prochlorococcus*. Bacteriophage lysis, as well as grazing by eukaryotic protists, play central roles as the major biotic causes of *Synechococcus* mortality in pelagic systems. However, *Synechococcus* populations show no extinction, suggesting high rates of production counteract this mortality, or that specific bacteriophage resistance and prey-selectivity mechanisms exist. This thesis set out to determine the molecular basis of cyanophage resistance in marine *Synechococcus*, using previously isolated cyanophage resistant *Synechococcus* sp. WH7803 mutants.

Whole genome sequencing (WGS) analysis revealed that these cyanophage-resistant *Synechococcus* sp. WH7803 mutants, as well as a re-sequenced wild type strain, possessed a distinctive mutation profile, with a high number of mutations present in each mutant and with mutations present at a variable frequency. Such a profile is in stark contrast to what was recently found in *Prochlorococcus* where specific mutations could be identified at 100 % frequency (Avrani et al., 2011). The mutation profile of *Synechococcus* sp. WH7803 prevented the precise identification of specific genes involved in cyanophage resistance in this strain. This profile was hypothesised to be related to *Synechococcus* sp. WH7803 being an oligoploid organism, i.e. possessing more than one chromosome copy. Indeed, a qPCR assay that was optimised showed this strain possesses on average four chromosome copies.

Subsequent isolation and WGS characterisation of cyanophage resistant mutants from the monoploid strain *Synechococcus* sp. WH7805 revealed a completely different mutation profile, most similar to that previously described for *Prochlorococcus*. This identified mutations in a single gene, encoding a possible glycosyltransferase, that were confirmed by Sanger sequencing, as being potentially responsible for cyanophage resistance in this strain.

LIST OF ABBREVIATIONS

AMG	Auxiliary metabolic gene
ASW	Artificial sea water
BAM	Binary alignment map format
bp	Base pair(s)
CDS	Coding DNA sequence
CT	Crossing threshold
CV	Coefficient of variation
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic Acid
g	grams
g	Relative centrifugal force (Earth's gravitational force)
GOS	Global Ocean Sampling expedition
indel	Insertion/deletion
kb	Kilo base pair(s)
LB	Lysogeny broth
LPS	Lipopolysaccharide
Mbp	Mega base pair(s)
MOI	Multiplicity of infection
NCBI	National Centre for Biotechnology Information
nt	Nucleotide(s)
OD	Optical density
ORF	Open reading frame
P	Phosphorous
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
PE	Phycoerythrin
PEB	Phycoerythrobilin
PEG	Polyethylene glycol
PFU	Plaque forming units
PUB	Phycourobilin
qPCR	Quantitative PCR
SAM	Sequence alignment map format
SNP	Single nucleotide polymorphism
TEM	Transmission electron microscopy
VLP	Virus-like particle
WGS	Whole genome sequencing
WT	Wild type

Chapter 1

INTRODUCTION

1.1. Introduction

The oceans cover approximately 70 % of the Earth's surface and play a central role in the global carbon cycle. Contained within are a plethora of marine microorganisms that play critical roles in all the major biogeochemical cycles (Arrigo, 2005).

Indeed, 2010 culminated a decade long international effort to assess the diversity, distribution and abundance of marine life via the Census of Marine Life (www.coml.org). This investigation helped identify more than 6,000 new marine species, including microbes, around the globe revealing the vast diversity and richness of marine life and highlighting our general lack of knowledge of these organisms.

Likewise, the multinational Tara Oceans consortium was the largest international effort to sequence DNA from marine plankton, resulting in over 35,000 samples of viruses, bacteria, protists and small animals collected from 210 stations in the major oceanic regions from 2009 to 2013, which revealed the vast diversity of marine plankton and hence resulted in several articles included in a recent special Science issue (Bork et al., 2015; Brum et al., 2015; Lima-Mendez et al., 2015; Sunagawa et al., 2015; de Vargas et al., 2015; Villar et al., 2015).

Despite the low density of marine life in oligotrophic gyre regions, it is due to their vast areal extent that they contribute the greatest to the marine CO_2 fixation budget, a process that directly influences surface ocean biogeochemistry, carbon dynamics, and potentially modulates climate change (Partensky et al. 1999). Notably, these oligotrophic gyres are dominated by marine picocyanobacteria.

1.2. Marine picocyanobacteria

Marine photosynthetic microorganisms, namely phytoplankton, encompass a wide range of microorganisms, characterized by being autotrophs, i.e. organisms that use inorganic compounds for both energy and growth. These organisms are responsible for ~50 % of global primary production, i.e. the fixation of CO_2 into biomass, due to their central position in the biological carbon pump (Behrenfeld et al., 2006; Field, 1998; Sanders et al., 2014). Likewise, marine picocyanobacteria have been estimated

to contribute between 8.5 to 65 % of the net marine carbon fixation in some oceanic regions (Flombaum et al., 2013; Jardillier et al., 2010; Veldhuis et al., 1997).

Within this group, *Synechococcus* spp. and *Prochlorococcus* spp. are the two major cyanobacterial genera. These marine picocyanobacteria are the most abundant and globally distributed photoautotrophs, which has been estimated to contribute between 32 % and 89 % of primary production in some oligotrophic regions of the oceans and hence are of key environmental relevance (Flombaum et al., 2013; Jardillier et al., 2010; Scanlan et al., 2009; Veldhuis et al., 1997).

The geographic distribution of *Prochlorococcus* and *Synechococcus* is complimentary (Fig. 1.1). Whilst *Synechococcus* prefers relatively high nutrient regions and has a broader distribution from pole to pole, *Prochlorococcus* favours oligotrophic areas, such as the oceanic gyres, and hence its distribution is largely restricted to ocean regions between ~ 40˚N and 40˚S (Partensky et al. 1999; Scanlan et al. 2009; Flombaum et al. 2013; Vincent et al. 2000; Zwirglmaier et al. 2008; Johnson et al. 2006; Martin et al. 2005; Sohm et al. 2015; Fuller et al. 2006).

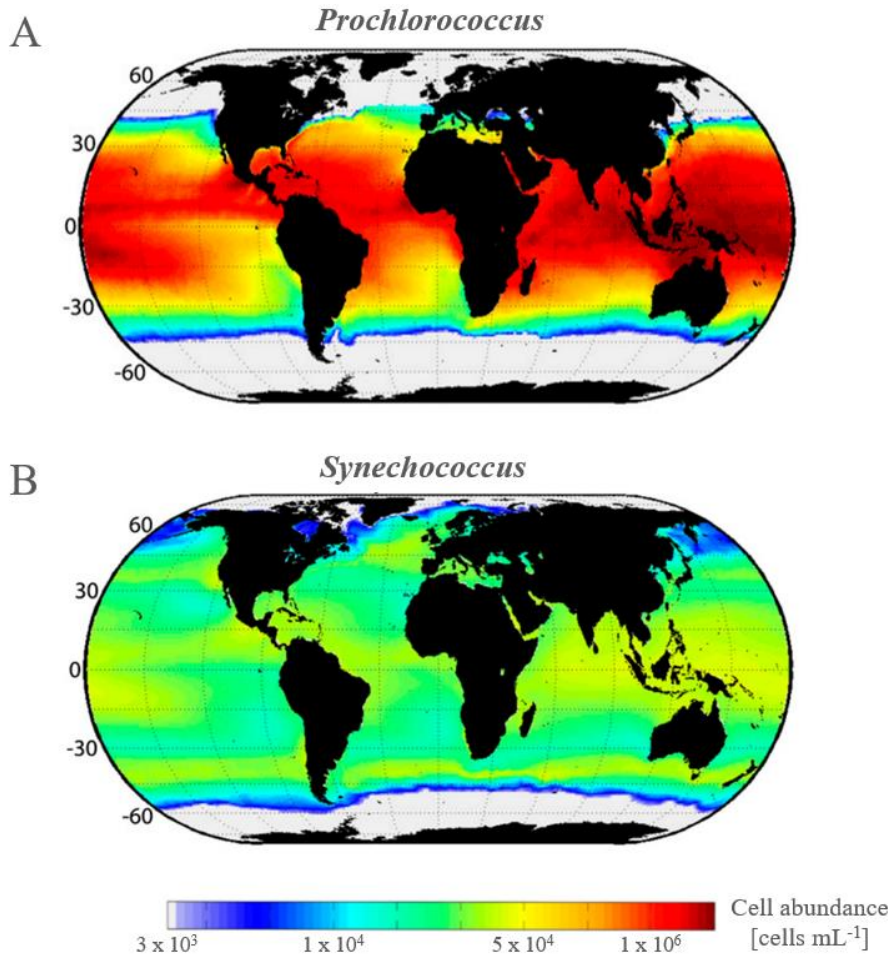


Figure 1.1. Global distribution and mean annual abundance of *Prochlorococcus* (A) and *Synechococcus* (B) (from Flombaum et al. 2013)

Marine *Synechococcus* and *Prochlorococcus* are phylogenetically closely related (Urbach et al., 1998). Several different conserved marker genes, including *rpoC1*, *ntcA*, *petB* and *narB*, as well as the 16S rRNA gene and the 16S-23S internal transcribed spacer region (ITS) have been investigated in order to determine their phylogenetic relationships (e.g. see Mazard et al. 2012). They cluster together forming what is known as the marine picophytoplankton clade (Fig. 1.2), with > 96 % identity in their 16S rRNA sequence (Urbach et al. 1998; Rocap et al. 2002; Mazard et al. 2012; Scanlan & West 2002; Fuller et al. 2003).

However, despite their similarity and that they co-exist in some oceanic regions, *Prochlorococcus* and *Synechococcus* have key physiological differences, including

elemental composition, cell size and pigmentation (Bertilsson et al., 2003; Heldal et al., 2003; Lindell, 2014; Moore et al., 2005)

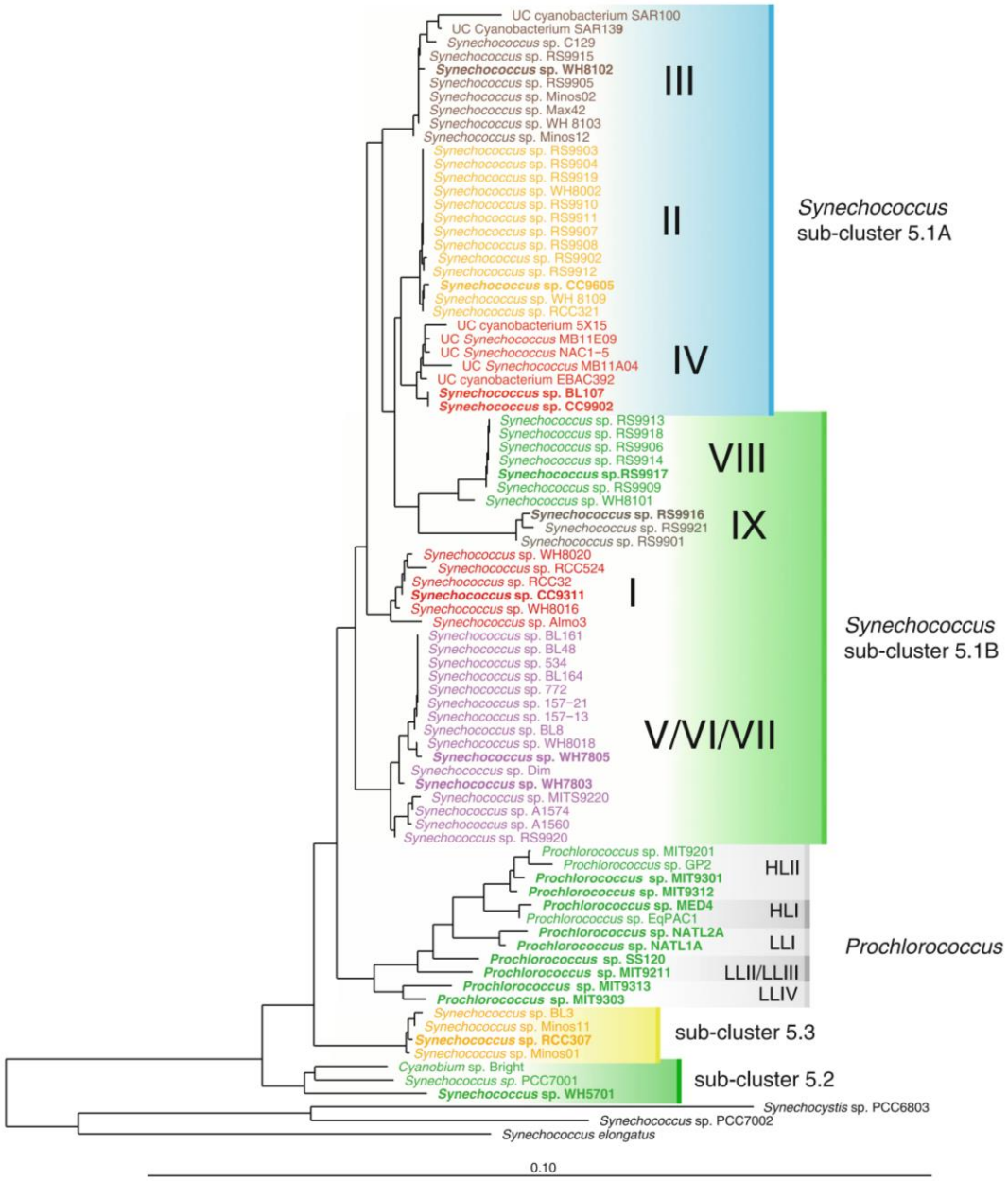


Figure 1.2. Neighbour-joining tree indicating the phylogenetic relationships of marine *Synechococcus* and *Prochlorococcus* (from Scanlan 2012)

The tree was constructed based on 16S rRNA gene sequences and bootstrap values of 70 %. Genome sequences are available for those strains indicated in bold.

The difference in global distribution of these two marine picocyanobacteria has been attributed to the adaptability of their light harvesting machinery and pigments. *Prochlorococcus* contains divinyl derivatives of chlorophyll *a* and *b* (Chl_{a2} and Chl_{b2}) as part of their light-harvesting membrane complex, but lacks a phycobilisome (Fig. 1.3). This light harvesting antenna is similar to that found in the chloroplasts of higher plants but a result of different evolutionary trajectories (Palenik and Haselkorn, 1992; Rippka et al., 2000; Urbach et al., 1992). Specific high light- and low light-adapted *Prochlorococcus* ecotypes exist that are genetically distinct and possess differences in their light harvesting accessory pigment content (Moore et al., 1998). Such variation in ecotypes has allowed *Prochlorococcus* to niche partition in vertically stratified water columns (Johnson et al., 2006; West and Scanlan, 1999; Zwirglmaier et al., 2008) resulting in greater integrated production than could be achieved by a single ecotype (Moore et al., 1998). Hence, *Prochlorococcus* can been found at greater depths than *Synechococcus*, sometimes down to depths > 200 m.

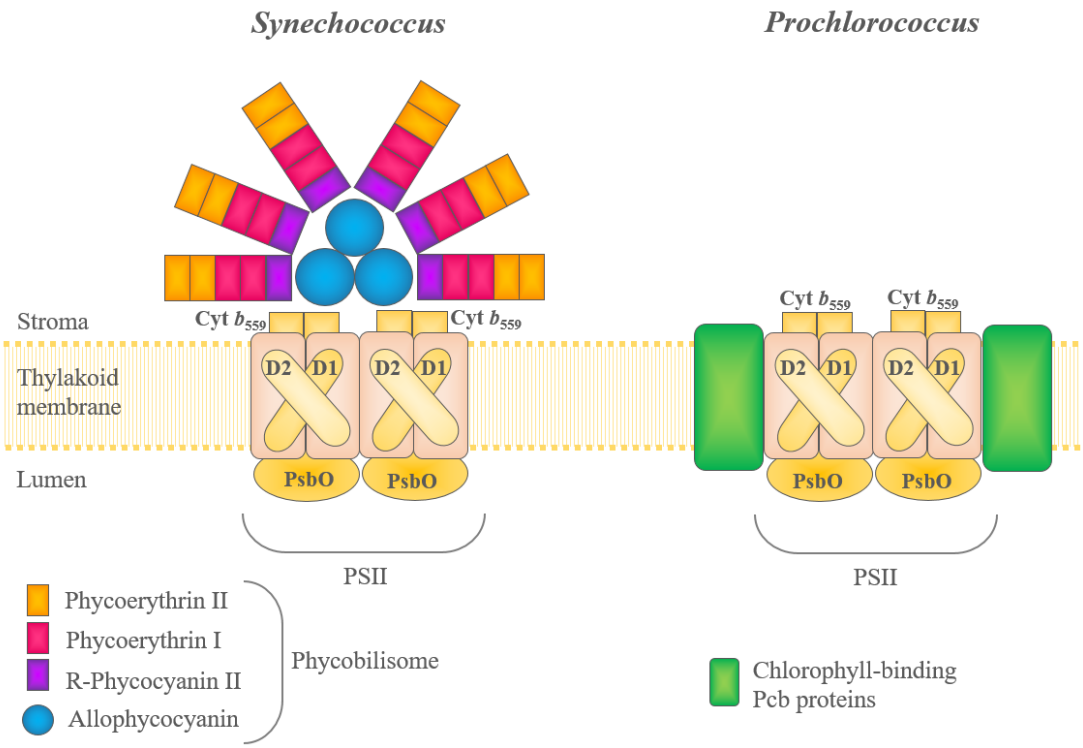


Figure 1.3. Schematic representation of the photosystem II (PSII) light-harvesting antenna of *Synechococcus* (left) and *Prochlorococcus* (right).

The phycobilisome composition shown is that found in *Synechococcus* sp. WH7803. Image based on Scanlan (2012); Ting et al. (2002) and Six et al. (2007). D1/D2: reaction centre protein; PsbO: manganese stabilising protein; Cyt b_{559}: cytochrome b_{559}

Synechococcus possesses a distinct light-harvesting accessory complex, the phycobilisome which can associate with photosystem I and II (see Fig. 1.3). This large protein complex consists of phycobiliproteins, dominated by phycoerythrin, phycocyanin and allophycocyanin, as well as a number of linker polypeptides that maintain structural integrity. Phycobiliproteins can account for half of the soluble protein content of cyanobacteria (Ong & Glazer 1991; Glazer & Clark 1986; Grossman et al., 1993).

In contrast to *Prochlorococcus*, which harvests mostly blue light (~400 – 500 nm), *Synechococcus* absorbs light of more green wavelengths (500 – 550 nm) and is potentially more flexible in the light wavelengths it can absorb due to the phycoerythrobilin (PEB; ~ 544 nm) and phycourobilin (PUB; ~ 493 nm)

chromophores attached to phycoerythrin present in the phycobilisome (see Six et al., 2007). The PUB:PEB ratio determines the light spectra that *Synechococcus* is able to absorb (Six et al. 2007; Scanlan et al. 2009; Ting et al. 2002; Morel et al. 1993; Grossman et al., 1993; Blot et al. 2009), with strains containing low PUB or PUB-lacking phycoerythrins dominating in mesotrophic waters whilst high PUB-containing phycoerythrin strains dominate in more transparent oligotrophic waters (Campbell and Iturriaga, 1988; Olson et al., 1988, 1990; Wood et al., 1999).

Depending on the specific phycobiliproteins present, *Synechococcus* can be classified into three pigment types. Pigment type 1 contains phycocyanin only, a representative being the 'green' strain *Synechococcus* sp. WH5701. Pigment type 2 consists of phycocyanin and phycoerythrin I and includes strain *Synechococcus* sp. WH7805 that is further studied in this thesis. Pigment type 3 comprises phycocyanin, phycoerythrin I and phycoerythrin II and can be subdivided into four sub-types depending on the PUB:PEB ratio. Thus, sub-type 3a includes strains possessing a low PUB:PEB ratio, type 3b a medium PUB:PEB ratio, type 3c a high PUB:PEB ratio or type 3d a variable PUB:PEB ratio (see Six et al. 2007). A widely studied type 3a pigment type representative is *Synechococcus* sp. WH7803, the major *Synechococcus* strain investigated in this thesis (Toledo et al., 1999; Waterbury et al., 1986).

1.3. Marine *Synechococcus*

Marine *Synechococcus* spp. are Gram negative bacteria around 1 µm in diameter (Morel et al. 1993; Olson et al. 1990) easily distinguished by their orange fluorescence detected using blue light epifluorescence microscopy or flow cytometry. The genus *Synechococcus* is a polyphyletic cluster including freshwater, obligatory marine and halotolerant strains, which has been divided into five clusters, with marine *Synechococcus* comprising sub-clusters 5.1, 5.2 and 5.3 (Fig. 1.2). Marine *Synechococcus* is genetically diverse hence the major sub-cluster (5.1) can be divided into at least 10 clades encompassing a wide range of pigment types (see section 1.1; Scanlan et al. 2009; Ahlgren & Rocap 2006; Mühling et al. 2006). Recently, the genomes of several marine *Synechococcus* strains have been sequenced (Dufresne et al., 2008; Palenik et al., 2003, 2006), including strains WH7803 and WH7805 which are the focus of this thesis, allowing unprecedented analysis of the molecular

mechanisms that have given rise to the ubiquitous distribution of this genus in oceanic waters (see Scanlan et al., 2009).

The distribution and abundance of *Synechococcus* is controlled both by biotic and abiotic factors. Several studies have investigated the bottom-up control shaping *Synechococcus* communities, including abiotic factors such as light (including UV light), temperature and nutrient availability (Six et al. 2007; Pittera et al. 2014; Sohm et al. 2015; Agawin et al. 2000; Agawin et al. 1998; Kretz et al. 2015; Fu et al. 2007; Zwirglmaier et al. 2008; Zwirglmaier et al. 2007; Fuller et al. 2005).

In contrast, predation of marine *Synechococcus* is the major biotic factor dictating *Synechococcus* community structure and mortality in pelagic marine systems. Such top-down factors include grazing (i.e. consumption) by eukaryotic protists, such as flagellates and ciliates, as well as bacteriophage (viral) lysis (Fuhrman, 1999; Mojica et al., 2015; Pernthaler, 2005; Rodriguez-Valera et al., 2009; Suttle, 2007; Weitz et al., 2015).

Whilst accurate estimates of virus-mediated mortality remain elusive being largely estimated using models (Suttle, 2005; Weitz et al., 2015), the microbial cell contents released after viral lysis can influence biogeochemical processes by increasing nutrient availability and hence being a source of recycled carbon, that can alter food webs via the viral shunt (Fig. 1.4; Breitbart, 2012; Suttle, 2005; Wilhelm and Suttle, 1999). Moreover, DNA transferred directly by bacteriophages or indirectly by releasing free DNA after host-cell lysis can influence changes in bacterial genome composition via horizontal gene transfer, this being an important factor driving evolution in microbial communities by generating genomic heterogeneity (Cheetham and Katz, 1995; Jiang and Paul, 1995; Jover et al., 2014; Weinbauer, 2004; Weinbauer and Rassoulzadegan, 2004).

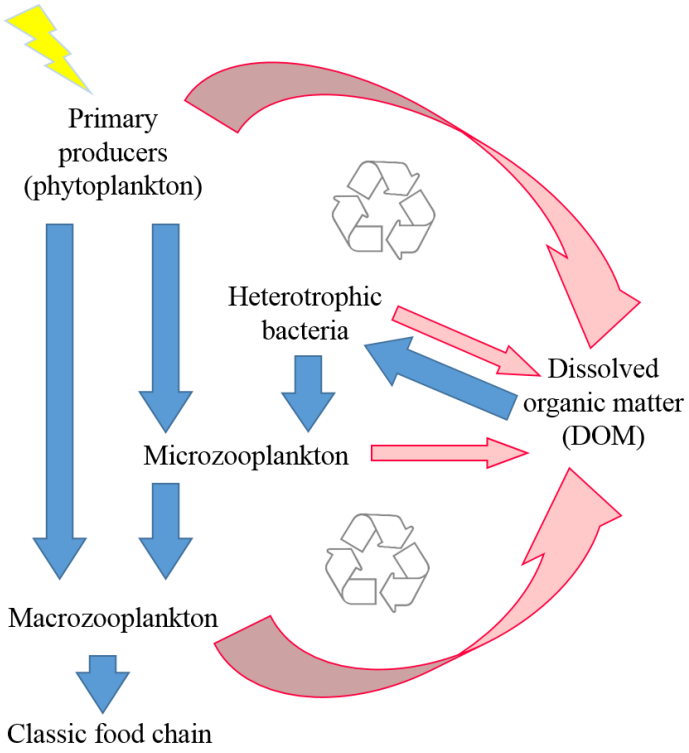


Figure 1.4. Schematic representation of the viral shunt (modified from Breitbart, 2012; Wilhelm and Suttle, 1999)

Viral infection (red arrows) produces dissolved organic matter that can recycle and be used by heterotrophic prokaryotes, whereas grazing brings carbon (blue arrows) to higher trophic levels.

1.4. Marine bacteriophages

Viruses are thought to be the most abundant biological entities found in all habitats. Bacteriophages, or phages, are viruses that infect bacteria. The term is derived from the Greek "to eat" (phagein) and the word “cyanophage” or “cyanovirus” has been adopted to refer to bacteriophages that specifically infect cyanobacteria. Other derivations of the name can refer to the classification of the bacteriophage, for example a “cyanomyovirus” would be specifically referred to as a bacteriophage from the *Myoviridae* family that infects cyanobacteria.

1.4.1 Abundance, enumeration methods and distribution of marine bacteriophages

Marine bacteriophages were first identified in 1955, when a sample of seawater from the North Sea off the coast of Aberdeen, Scotland, was used to lyse a bioluminescent bacteria, *Photobacterium phosphoreum* (Spencer, 1955). Although not a marine bacteriophage, the first bacteriophages isolated from seawater lysed enterobacteria (Zobell, 1946), thus it was believed that the levels of bacteriophages in unpolluted seawater were low and hence not ecologically important.

Nonetheless, viral particles are indeed much more abundant than was thought at that time. Thus, in 1979 the use of transmission electron microscopy (TEM) showed 10^4 viruses mL^{-1} were present in a concentrated seawater sample filtered through a 0.2 µm pore size filter (Torrella and Morita, 1979). Subsequently, Bergh et al. (1989) found up to 2.5×10^8 viruses mL^{-1} using TEM of ultracentrifuged samples from the Atlantic and Chesapeake Bay, revealing the abundance and potential importance of viruses (and bacteriophages) in marine environments.

Subsequently, epifluorescence microscopy was used as a more precise method for counting virus-like particles (VLPs). Use of this technique gave similar results, but provided a simpler and quicker approach than TEM (Wommack and Colwell, 2000). Use of epifluorescence microscopy requires staining of VLPs attached to membrane filters, commonly a dsDNA-binding fluorochrome excitable with UV light, such as 4',6-diamidino-2-phenylindole (DAPI), YO-PRO, SYBR Green I and SYBR Gold (Noble and Fuhrman, 1998; Shibata et al., 2006), although small bacteria and large VLPs can give rise to counting errors depending on the pore size selected. Currently, this method has been improved by the addition of a direct "wet mounting" step instead of using filters (Cunningham et al., 2015).

Flow cytometry is now also widely used to count VLPs in environmental samples (see e.g. Brussaard et al. 2000). Since VLPs are not naturally fluorescent they need to be stained in the same way as for epifluorescence microscopy, and due to their small size they can be confused with 'background' noise. For this reason, flow cytometric

analysis of viral abundance is regularly being further optimised, developing better instruments and protocols to improve sensitivity (Brussaard, 2004).

Nevertheless, these methods can only estimate the total abundance of VLPs but not infective viral particles. To determine the abundance of infective bacteriophage it is necessary to use a plaque assay, a culture-dependent method that relies on the availability of a suitable host. Using this method it is only possible to establish the abundance of specific bacteriophages infecting specific hosts (Millard, 2009).

Overall estimates of viral abundance are based on their average abundance in different marine ecosystems. Thus, Suttle (2005) calculated an average abundance of VLPs of 3×10^6 mL^{-1} suggesting a total of ca. 4×10^{30} viruses in the entire ocean, which he suggested that "if stretched end to end, would span farther than the nearest 60 galaxies" (Suttle 2005, 2007).

Whilst some investigators have questioned the numbers calculated by Suttle as an overestimation, because the techniques used for detecting VLP abundance rely on staining of dsDNA (Biller et al., 2014; Forterre et al., 2013; McDaniel et al., 2010), recent work specifically targeting ssDNA and RNA viruses suggests that these numbers are not far from reality, since these latter viral types have been largely omitted in previous studies (McDaniel et al., 2010; Steward et al., 2013).

The distribution and abundance of bacteriophages is critically dependent on host abundance. Certainly, seasonal variations in marine bacteriophage abundance have been correlated with concomitant fluctuations in bacterioplankton abundance, with a generally higher abundance of bacteriophages during warmer seasons, and with a difference of nearly two orders of magnitude between spring and winter (Bergh et al., 1989; Bratbak et al., 1996; Millard and Mann, 2006; Weinbauer, 2004; Winget and Wommack, 2009).

Likewise, spatial stratification of bacteriophages in the water column is also linked to the availability of hosts with an obvious higher abundance of viruses infecting phototrophs in the euphotic zone, but also bacteriophage found infecting microbes inhabiting hydrothermal vents at depths down to 2,600 m (Boehme et al., 1993; Hara et al., 1996; Jiang et al., 2003).

1.4.2 Bacteriophage classification

Bacteriophages have been classified using various characteristics, such as the host they infect, their nucleic acid composition, infection type and morphology. Based on their morphology and nucleic acid content, the International Committee on Taxonomy of Viruses (ICTV) has currently classified 19 different viral families that infect bacteria and Archaea (Ackermann, 2011).

The Caudovirales order encompasses 96 % of the known bacteriophages. This order comprises three families of tailed bacteriophages with icosahedral heads: the *Myoviridae, Siphoviridae* and *Podoviridae* (Table 1.1). The *Myoviridae* family have larger genomes with a lower GC % content compared to members of the *Podoviridae* and *Siphoviridae* which have smaller genomes and a variable GC % content (Comeau et al., 2012). The broader host range of bacteriophages of the *Myoviridae* family has been explained by their possession of tRNA genes within their genomes, allowing them to maintain translation efficiency across a wide range of hosts, in comparison to members of the *Podoviridae* and *Siphoviridae* where codon usage and the efficiency of translation derives from the host (Bailly-Bechet et al., 2007; Enav et al., 2012).

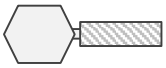
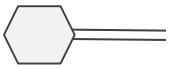
Shape	Family	Nucleic acid	Characteristics	Example
	*Myoviridae*	dsDNA linear	Contractile tail, no envelope, order Caudovirales	T4
	*Siphoviridae*	dsDNA linear	Long tail, non-contractile, no envelope, order Caudovirales	λ
	*Podoviridae*	dsDNA linear	Short tail, no envelope, order Caudovirales	T7
	*Microviridae*	ssDNA circular	12 knob-like capsomers, 27 nm	ϕX174
	*Corticoviridae*	dsDNA circular	Complex capsid, lipids, 63 nm	PM2
	*Tectiviridae*	dsDNA linear	Inner lipid vesicle, pseudo-tail, 60 nm	PRD1
	*Leviviridae*	ssRNA linear	Like poliovirus, 23 nm	MS2
	*Cystoviridae*	dsRNA linear	Segmented, lipidic envelope, 70-80 nm	ϕ6
	*Inoviridae*	ssDNA circular	Filaments or rods, 85-1950 x 7 nm	Fd
	*Plasmaviridae*	dsDNA circular	Lipidic envelope, no capside, 80 nm	MVL2

Table 1.1. Characteristics of bacteriophage families (modified from Ackermann 2011; Ackermann 2007)

1.4.3 The diversity of marine bacteriophages with a focus on marine cyanophages

Molecular biological techniques have been widely used to determine the diversity of bacteriophages. However, there is no universal genetic marker akin to the host 16S rRNA gene, which has limited bacteriophage diversity studies to certain taxa. Nevertheless, structural proteins have been targeted as molecular markers because they can be conserved between different members of a taxonomic group, such as g23 which encodes the major capsid protein in T4-like bacteriophages. The g23 gene, as well as g20 encoding the portal protein located in the neck of the bacteriophage, have been used as a phylogenetic marker to determine the diversity of T4-like bacteriophages that have a wide infection range (Clasen et al., 2013; Filée et al., 2005; Paul et al., 2002; Zhong et al., 2002). Other approaches to determine bacteriophage diversity include pulsed field gel electrophoresis (PFGE), which has helped to understand how marine viral diversity changes over time and in different environments (Steward et al., 2000).

More recently, metagenomic analyses have allowed the incorporation of large datasets obtained from different ocean sampling campaigns, such as Tara and GOS, to determine marine bacteriophage diversity (Brum and Sullivan, 2015). This new approach to studying marine viral diversity has revealed that previous genomic databases did not include 63 – 93 % of marine virus sequences, suggesting a vast diversity yet to be discovered (Edwards and Rohwer, 2005; Hurwitz and Sullivan, 2013). Using metagenomic analysis of the Tara expedition samples, coupled to morphological analyses using quantitative TEM, Brum et al. (2015) found structured viral communities that were transported by ocean currents. These findings support the seed-bank hypothesis in which high local viral diversity and community structure is maintained by environmental conditions and passive spreading of viruses by ocean currents (Breitbart and Rohwer, 2005; Brum et al., 2015).

Currently, all cyanophages investigated so far belong to the three families of the Caudovirales order described in section 1.4.2 above, i.e. *Myoviridae, Podoviridae* and *Siphoviridae*, which are all double-stranded DNA (dsDNA) bacteriophage. Cyanophages of the *Myoviridae* family tend to have a broader host range, including

members of both *Prochlorococcus* and *Synechococcus* as hosts, whereas cyanophages of the *Podoviridae* and *Siphoviridae* families are mostly host-specific (Marston and Sallee, 2003; Millard and Mann, 2006; Waterbury and Valois, 1993).

Marine cyanophages were first isolated and characterised more than two decades ago (Suttle and Chan, 1993; Suttle et al., 1990; Waterbury and Valois, 1993; Wilson et al., 1993) and their importance relies in their role controlling and interacting with marine cyanobacterial populations. In addition to culture-based studies, these interactions between marine cyanobacteria and cyanophages have been observed by metagenomic analysis, where a high diversity of these cyanophages has been detected (Angly et al., 2006).

Cyanophage-host interactions include the presence of bacterial gene homologs in the cyanophage genome involved in bacterial metabolism. These genes have thus been called auxiliary metabolic genes (AMGs) and constitute ca. 30 % of the cyanophage gene pool, contributing significantly to cyanophage genetic diversity (Breitbart et al., 2007; Ignacio-Espinoza et al., 2013; Millard et al., 2009). Currently, most of the AMGs studied are implicated in photosynthesis and carbon metabolism (Mann et al., 2003; Philosof et al., 2011; Puxty et al., 2014; Sharon et al., 2009). However, several genes involved in different cellular processes have also been identified (Bryan et al., 2008; Sullivan et al., 2010; Thompson et al., 2011; Zeng and Chisholm, 2012).

1.4.3.1. The cyanophage S-PM2

S-PM2 (formerly S-PSI) is a T4-like cyanomyovirus, first isolated in 1992 from coastal waters off Plymouth (UK) using *Synechococcus* sp. WH7803 as the cyanobacterial host (Wilson et al. 1993, 1996). In spite of being a T4-like cyanophage, S-PM2 does not share morphological characteristics with T4 apart from the contractile tail, whilst phylogenetic analysis based on g23 showed a divergence from the T4-like bacteriophages, forming a new group, the exoT-evens (Hambly et al., 2001).

The interest in studying this cyanophage started with the discovery of the first viral *psbA* gene (encoding the D1 protein of photosystem II, see Fig. 1.3) containing a group I intron (Millard et al., 2004). Further work found that S-PM2 possesses a homing endonuclease downstream of *psbA*, connected by an antisense RNA that is co-

regulated with them, possibly regulating their expression (Millard et al., 2010). The role of *psbA* in S-PM2 has not yet been confirmed, although it is expressed during the lytic cycle and has been suggested to be involved in protection against photoinhibition, i.e. where a higher rate of photo-induced damage than synthesis of D1 protein occurs (Bailey et al., 2004; Clokie et al., 2006), maintaining photosynthesis and thus energy production during infection.

Moreover, using quantitative real-time PCR it was found that S-PM2 increases the expression of the *cpeAB* and *mpeAB* operons, that encode the subunits of PE I and II in *Synechococcus* sp. WH7803, resulting in an increased light-harvesting capacity possibly helping the cyanophage during the infection cycle by maintaining the photosynthetic machinery (Shan et al., 2008).

Adsorption of S-PM2 to *Synechococcus* sp. WH7803 has been found to be light-dependent, with decreased adsorption under red light. Furthermore, inhibition of photosynthesis (using dichlorophenyl dimethylurea) did not affect adsorption (Jia et al., 2010).

Another environmental factor that affects S-PM2 infectivity is the external pH. Traving et al. (2014) found that a decreased pH negatively affected the S-PM2 infection cycle and infectivity. These authors showed that at pH 7.6 (the pH of seawater is ~ 8) S-PM2 had a shorter latent period and a reduced burst size, and no progeny was detected at pH 7 (Traving et al., 2014).

Cyanophage S-PM2 has been characterised as a lytic cyanophage, although Wilson et al. (1996) proposed the presence of a pseudolysogenic cycle (see Fig. 1.5) when propagated on *Synechococcus* sp. WH7803 grown under phosphate-deplete conditions. This pseudolysogeny was manifest as a decreased burst size of ca. 80 % whilst only 9.3 % of infected cells lysed, compared to 100% in phosphate-replete conditions (Wilson et al., 1996). A decreased burst size of S-PM2 has also been associated with the deletion of a region of ca. 10 kb encompassing various 'ORFan' genes, i.e. genes with no known homologs, in an 'evolved' S-PM2 (Puxty et al., 2015). This deletion was detected by whole genome re-sequencing of S-PM2 after ~10 years of propagation under laboratory conditions compared to its original genome sequence (Mann et al., 2005; Puxty et al., 2015).

1.4.4 The bacteriophage infection cycle

Currently, three main mechanisms of bacteriophage infection and replication in their host have been identified. These are the lytic, lysogenic and pseudolysogenic cycles, although a fourth type, the chronic infection cycle, is also included here for completeness (Fig. 1.5).

During an obligately lytic infection, following injection of bacteriophage nucleic acids into the host, only further bacteriophage replication and host cell lysis, releasing newly assembled progeny bacteriophage, can occur. Whilst the lysogenic cycle commences with the same attachment and nucleic acid injection process, in contrast to the lytic cycle, bacteriophage DNA can either integrate into the host's genome or continue in the same fashion as the lytic cycle ultimately lysing the host cell. In some cases e.g. bacteriophage P1, phage can form a self-replicating plasmid inside their host, in this case *Escherichia coli* (Sternberg and Austin, 1981). Well characterised examples of the lytic and lysogenic infection cycles are bacteriophages lambda and T4 (Ackermann and DuBow, 1987; Campbell, 2003; Miller et al., 2003; Stent, 1963; Wommack and Colwell, 2000).

The pseudolysogenic cycle is similar to the lysogenic cycle with a delayed lysis of the host cell, but importantly without integration of bacteriophage DNA into the host genome (Barksdale and Arden, 1974). A good example of this infection type is that seen with cyanophage S-PM2, one of the cyanophage widely used in this thesis, when infecting phosphate starved *Synechococcus* host cells, where a ca. 90 % decrease in host cells lysis compared to phosphate-replete conditions, and no evidence of S-PM2 integration into the host genome, was observed (see section 1.4.3.1 above and Wilson et al. 1996).

The chronic infection cycle is similar to the lytic cycle but with the difference that the host cell does not lyse and instead constantly produces bacteriophages by budding or extrusion (Fig. 1.5). Such a mechanism has been suggested for a filamentous VLP infecting *E. coli*, although it is not clear how these VLPs are released without lysing the cell due to their large size (Fuhrman, 1999; Hofer and Sommaruga, 2001).

Initiation of lytic infection is highly dependent on the number of hosts present, with for marine cyanophage, a minimum of 10^3 cells mL^{-1} being required for a successful infection, whereas in other bacteriophage/host systems this minimal abundance required can reach 10^4 cells mL^{-1} (Suttle and Chan, 1994; Wiggins and Alexander, 1985). Host morphology may also play a role in infection dynamics, with a study in the northern Adriatic Sea demonstrating that whilst the number of cocci and spirallae bacteria had no influence on infection, a minimum of 10^5 bacterial 'rod-shaped' cells mL^{-1} was required (Weinbauer and Peduzzi, 1994).

Apart from the pseudolysogenic cycle described above, i.e. following infection of a phosphate-depleted *Synechococcus* host by cyanophage S-PM2 (Wilson et al., 1996), all cyanophages isolated on marine *Synechococcus* thus far follow an obligately lytic cycle (Clokie et al., 2011). However, there is some evidence for the presence of temperate phages infecting natural *Synechococcus* populations after prophage induction using mitomycin C, observed as an increase in the number of cyanophage and with a corresponding decrease in *Synechococcus* cell numbers (Long et al., 2008). This induction appears not to be correlated with nutrient availability but appears rather defined by ambient population size and cyanophage abundance (McDaniel and Paul, 2005) with induction generally occurring in the late winter months during times of reduced host abundance (McDaniel et al. 2002).

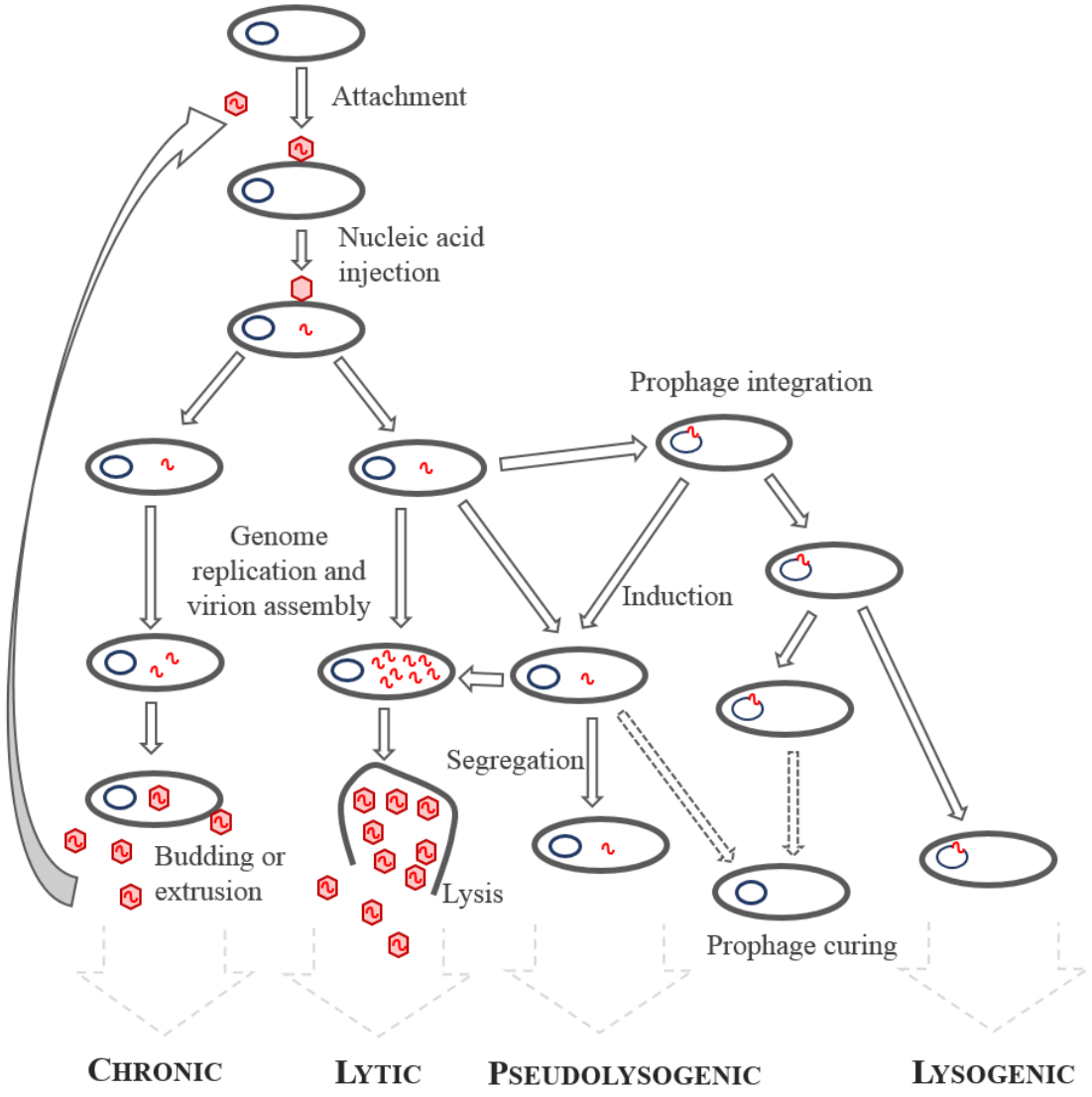


Figure 1.5. Schematic representation of the various potential bacteriophage infection cycles (modified from Weinbauer 2004).

Following injection of the bacteriophage nucleic acids (in red), the bacteriophage can enter into each of the four infection cycles shown. The lysogen can return to the original state by curing of the prophage.

The infection cycle of marine cyanophages is relatively long, with a latent period determined under laboratory conditions of between 5 – 12 hours and up to 36 hours with only a few exceptions (Huang et al., 2012a; Lindell et al., 2005; Raytcheva et al., 2011; Wang and Chen, 2008; Wilson et al., 1996; Zeng and Chisholm, 2012). These lengthy infection times are longer than the turnover of host proteins, some of which

are essential for bacteriophage replication. However, studies in T4 and T7-like cyanophages suggest that they use the canonical process of lytic/tailed bacteriophages of degrading the host's DNA and driving the host transcription machinery towards expression of bacteriophage genes (Clokie et al., 2006; Kovalyova and Kropinski, 2003; Lindell et al., 2005; Mann et al., 2005; Millard et al., 2009; Miller et al., 2003; Sullivan et al., 2010; Weigele et al., 2007).

1.5. Mechanisms of bacteriophage resistance and bacteriophage-bacteria co-evolution

To survive infection by bacteriophages, bacteria have developed several resistance mechanisms that can prevent infection at any stage of the infection cycle (Fig. 1.6; for review see Dy et al. 2014; Labrie et al. 2010).

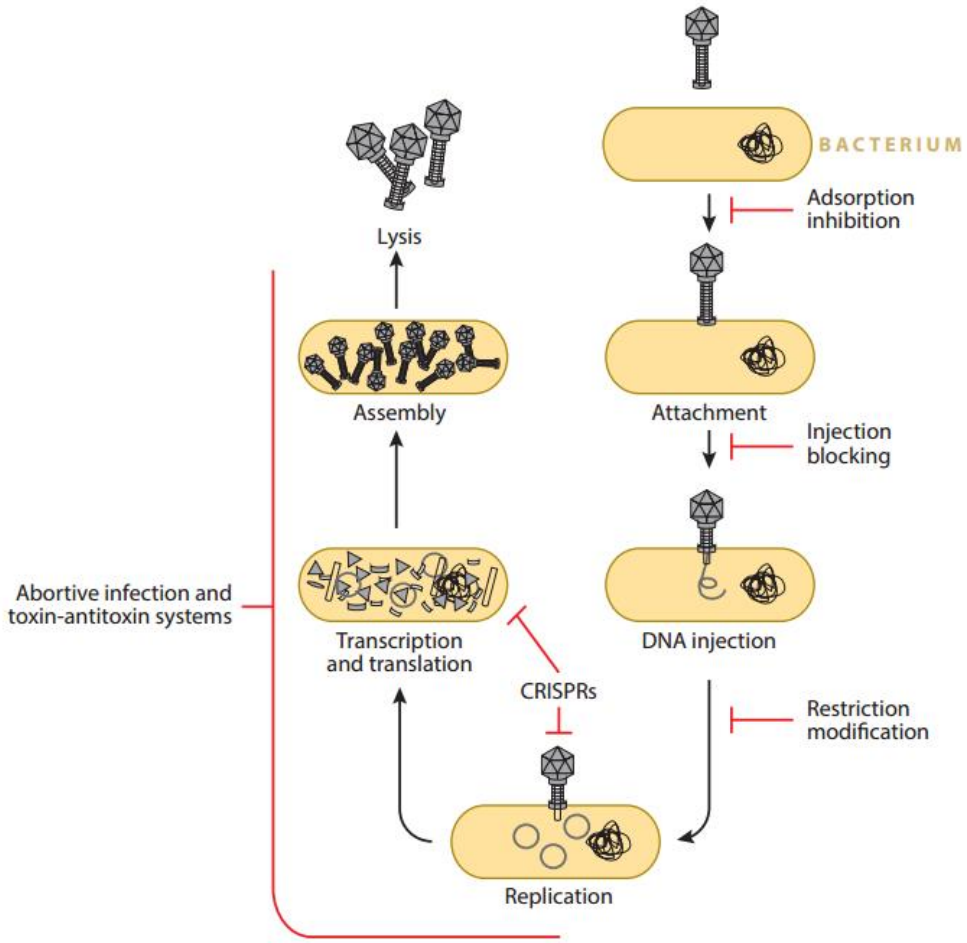


Figure 1.6. Resistance mechanisms associated with the lytic infection cycle of bacteriophage (from Dy et al. 2014).

Bacteriophage infection can be interrupted at any step of the infection cycle. In red are shown the known mechanisms for bacteriophage resistance.

1.5.1 Abortive infection and toxin-antitoxin systems

One mechanism to prevent phage infection is the "Abi" system (abortive infection system), an altruistic mechanism where a single bacterium is "sacrificed" in order to protect remaining members of that bacterial population from future infections. An example of the Abi system is the toxin-antitoxin system, in which the de-regulation of toxin and antitoxin molecules can lead to cell death. These toxins and antitoxins can be either proteins or RNA molecules, the latter like the ToxIN protein-RNA toxin-

antitoxin system described for *Erwinia carotovora* subspecies *atroseptica*, in which the toxin ToxN is inhibited by the formation of a complex with the antitoxin ToxI RNA; when the phage infects, this balance is altered leading to toxicity of ToxN and further cell death (Fineran et al., 2009).

1.5.2 CRISPRs

The clustered regularly interspaced short palindromic repeats (CRISPR) system is an array of 2 to 249 short repeated sequences between 24-47 bp in length with spacers of 26-72 bp that participates in bacteriophage resistance. CRISPRs have been found in 40 % of bacterial genomes and 90 % of archaeal genomes in combination with CRISPR associated (Cas) proteins (Bhaya et al., 2011; Bondy-Denomy and Davidson, 2014). This resistance mechanism relies on the acquisition of bacteriophage sequences called spacers in between CRISPRs by surviving hosts. These short sequences are transcribed and then cleaved by the Cas proteins, generating small crRNAs (CRISPR RNA) containing a specific hairpin structure and a spacer. When the cell gets infected, these crRNAs and Cas proteins recognise the bacteriophage DNA due to base pairing, degrading the foreign DNA before it can be replicated. This system was first proposed as an "adaptive immunity system" nearly two decades after its discovery by Ishino et al. (1987), and has led to the development of targeted genome editing tools using Cas nucleases (such as Cas9) to modify genes in a rapid, easy and efficient fashion (Bhaya et al., 2011; Díez-Villaseñor et al., 2010; Mojica et al., 2009; Sander and Joung, 2014; Sorek et al., 2008; Westra and Brouns, 2012).

1.5.3 Restriction modification systems

In this resistance mechanism host restriction enzymes recognise and degrade foreign DNA based on their lack of methylation of specific DNA sequences (in contrast to host sequences which are methylated) thus preventing bacteriophage replication. Restriction-modification systems encompass a restriction endonuclease enzyme that recognises and cleaves a specific DNA sequence and a methyltransferase that methylates adenine or cytosine residues to prevent host DNA cleavage by the restriction endonuclease. However, if the bacteriophage DNA is methylated inside the host, the replication process and thus the lytic cycle proceeds. This is particularly

important when new viruses are released from the host cell, as they can keep infecting bacteria with similar restriction-modification systems due to their DNA being methylated and therefore immune to cleavage by endonucleases of similar hosts (Tock and Dryden, 2005).

1.5.4 Injection blocking

Super-infection exclusion systems are bacteriophage mechanisms for resistance to infection by other bacteriophages, and thus have been included in this section. T4-like cyanophages, such as S-PM2, inject DNA into the host cell through the needle when the tail is contracted (Fig. 1.7). Super-infection exclusion systems (Sie) are bacteriophage encoded genes that block the entry of DNA from other bacteriophages. In T4 and other T-even bacteriophages, i.e. T2 and T6, the Sie membrane protein Imm (immunity) directly blocks DNA injection whilst the Sie protein Sp blocks the translocation of DNA by inhibiting bacteriophage lysozymes that degrade the peptidoglycan. This results in the degradation of foreign DNA by nucleases in the periplasm for about four minutes after infection. Other genes also appear to play a role in this process although the exact mechanism is not yet fully understood (Lu and Henning, 1989, 1994; Lu et al., 1993; Miller et al., 2003). Bacteriophages of Gram positive bacteria e.g. *Lactococcus lactis* and *Streptococcus thermophilus* have a similar system that protects them from secondary bacteriophage infections (Dy et al., 2014).

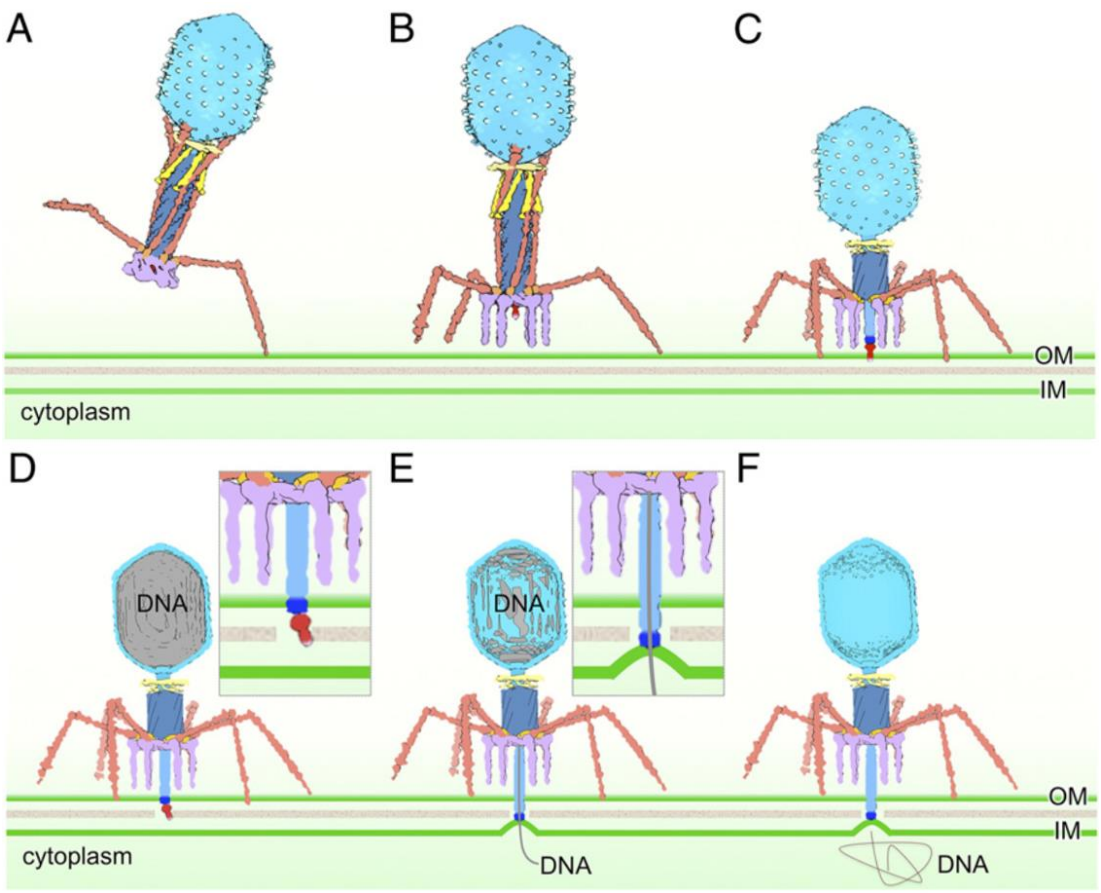


Figure 1.7. Schematic representation of bacteriophage T4 adsorption and DNA injection (from Hu et al. 2015)

(A) Long tail fibres (brown) bind to the host receptor. (B) The attachment of long tail fibres triggers a conformational change in the baseplate proteins (purple) that releases a subset of short tail fibres. (C) A conformational change in the baseplate triggers the contraction of the tail (dark blue) and a subsequent conformational change in the whiskers or fibritin collar (yellow) releases the remaining long tail fibres. (D) Contraction of the tail pushes the needle (light blue) through the outer membrane (OM). (E) Dissociation of the protein complex (red-blue end) allows degradation of the peptidoglycan. (D) The 'tape measure' protein (blue) fuses with the inner membrane (IM) and the DNA is injected into the cytoplasm. (F) The bacteriophage DNA is fully released into the cytoplasm.

1.5.5 Adsorption inhibition

Bacteriophage adsorption to a certain host requires the reversible attachment of the long tail fibres, that 'walk' until finding the optimal site for irreversible attachment of the short tail fibres to a specific receptor on the host cell surface (Heller and Braun, 1979; Hu et al., 2015). This attachment depends on the encounter rate between bacteriophage and host bacterium, and hence a bacteriophage with a broad host range may increase the likelihood of a successful adsorption event. Nonetheless, possessing a broad host range can also affect adsorption to a specific host whereas a narrow host range normally increases the chance of adsorption to a specific hosts thus resulting in a successful infection (Rakhuba et al., 2010). The abundance and diffusion rates of the bacteriophage and bacterial host are also clearly important in determining adsorption efficiency, on top of the actual binding affinity of the bacteriophage to the specific host receptor that all play a role in defining a successful infection (Hyman and Abedon, 2009).

It is worth pointing out that a phenomenon called virion-mediated "lysis from without" can also occur, which has been studied mechanistically in T-even bacteriophage, and causes bacterial cell lysis due to the high concentration of bacteriophage present, but importantly does not cause a productive infection. This occurs as a result of the high number of bacteriophages attaching to the host receptor and further lysis of the peptidoglycan (Fig. 1.7.D) e.g. by the gp5 protein in bacteriophage T4, which causes excess damage to the cell envelope and direct lysis of the bacterial cell. Such a phenomenon was first observed by Delbrück (1940) when a "saturating" amount of bacteriophage lysed *E. coli* 5 min earlier than with a lower concentration, but without producing any progeny phage. To overcome this problem, bacteriophages have developed a "resistance to lysis from without" that involves the expression of similar genes necessary for the super-infection exclusion system (see section 1.4.4). This system can be observed by a delay between primary and secondary high-multiplicity adsorptions (Abedon, 2011).

The attachment of the bacteriophage to the host receptor is the key step defining adsorption and is specific for each bacteriophage. Bacteriophage receptors are cell surface molecules that are generally in high abundance due to them being essential for

a particular biological function. In Gram negative bacteria, bacteriophage receptors are usually part of the outer membrane structure that faces the outside of the cell. Examples of these molecules include outer membrane proteins, lipopolysaccharide (LPS), capsular polysaccharide and lipoproteins. Consequently, development of resistance to bacteriophage adsorption can either be by the modification or the blockage of these cell surface receptors (Dy et al., 2014; Labrie et al., 2010).

Several ways of blocking bacteriophage adsorption to the host cell receptor are known, including producing exopolysaccharide (extracellular matrix), competitive inhibitors that can bind to the bacteriophage receptor, or the synthesis of proteins that can mask the receptor (Labrie et al., 2010). For example, masking of a bacteriophage receptor to avoid super-infection has been identified in bacteriophage T5. This bacteriophage encodes a lipoprotein that blocks the outer membrane protein FhuA, preventing the irreversible attachment of new bacteriophages to *E. coli* thus avoiding super-infection (Pedruzzi et al., 1998).

Other examples of adsorption inhibition by masking the bacteriophage receptor are known for the T-even-like bacteriophages. Thus, in *E. coli*, a common receptor for these bacteriophages is OmpA, a protease located in the outer membrane of the cell and one of the major outer membrane proteins. Protein TraT, an F-factor lipoprotein, interacts with OmpA blocking the attachment of T-even-like bacteriophages due to its sequence similarity with gp38 from the long tail fibres. However, because TraT only blocks part of OmpA, the adsorption is only partially inhibited (Riede and Eschbach, 1986).

Extracellular polymeric substances (EPS), mainly composed of polysaccharide are produced by a wide range of Gram-negative and Gram-positive bacteria. In cyanobacteria, EPS can be found associated with the outer membrane or surrounding the cell as released polysaccharides. EPS can directly protect cyanobacteria from phagocytosis, dehydration and bacteriophage infection by physically blocking adsorption to the bacteriophage receptor (Mota et al., 2013; Pereira et al., 2009). There has been increased recent interest in cyanobacterial EPS components, mainly due to their biotechnological and industrial relevance due to their similarity to plant

polysaccharides, with *Cyanothece* sp. CCY0110 one of the most efficient cyanobacterial producing EPS strains (Mota et al., 2013).

Cell surface polysaccharides can also block bacteriophage receptors. *E. coli* produces several types of polysaccharides, including LPS (O antigen), capsular polysaccharide (K antigen), enterobacterial common antigen and colonic acid (M antigen). The structure of the O and K antigens determines the different *E. coli* serotypes, since the enterobacterial common antigen and M antigen are only expressed under certain growth conditions (Orskov et al., 1977; Whitfield, 2006). LPS is often implicated with bacteriophage adsorption by either masking outer membrane proteins that are bacteriophage receptors or by being the bacteriophage receptor itself, and thus changes in its structure can mediate bacteriophage resistance (Hancock and Reeves, 1976; van der Ley et al., 1986). The role of LPS in cyanophage resistance is not yet understood and only suggested to be linked due to the observed inhibition of cyanophage adsorption in mutants of *Anabaena* sp. PCC7120, a freshwater cyanobacterium, possessing a modified LPS (Xu et al. 1997; see section 3.1).

Noteworthy, is that in bacteriophage T4 the conformational change associated with the baseplate proteins (Fig. 1.6.A and B) is only after a minimum of three of the long tail fibres have bound to a glucose residue of the LPS, which positions the short tail fibres (Fig. 1.6.C). These fibres then irreversibly bind to the heptose residue of the LPS which is referred to as the second bacteriophage receptor (Crawford and Goldberg, 1980; Montag et al., 1987; Riede et al., 1985).

1.5.5.1. LPS structure and biosynthesis

LPS consists of three main structures linked together (Fig. 1.7), comprising a lipid component included in the cell outer membrane (lipid A), a core oligosaccharide with an inner and outer section, and the O-polysaccharide or O-antigen, with repeated polysaccharide motifs. Currently, knowledge of LPS biosynthesis is mainly via studies performed in *E. coli*, with only a limited number of LPS biosynthesis genes predicted in marine cyanobacteria (Snyder et al., 2009).

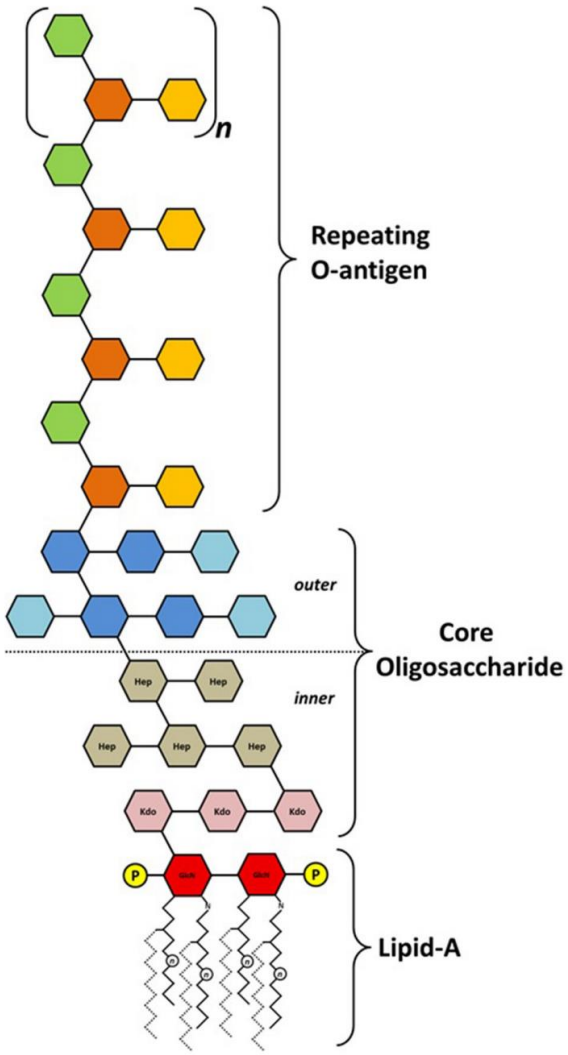


Figure 1.8. General structure of LPS as modified from Serrato (2014).

Fatty acid chains are indicated as lines, glucose (red), phosphate (yellow), Kdo (pink) and heptose (grey).

In *E. coli*, lipid A is synthesised by nine enzymes located in the cytoplasm and inner membrane. These enzymes are encoded by genes of the *lpx* operon, that is conserved in Gram negative bacteria and is constitutively expressed, although the expression of modification enzymes LpxK, KdtA, LpxL and LpxM generally depends on external conditions (Raetz et al., 2007; Trent, 2004).

This biosynthetic pathway is summarised in Fig. 1.9. Briefly, LpxA catalyses the acylation of uridine-diphosphate-N-acetylglucosamine (UDP-GlcNac) with a β-

hydroxymyristoyl chain, followed by deacetylation by LpxC and then LpxD adds a second β-hydroxymyristoyl chain. LpxH cleaves a uridine-monophosphate (UMP) forming lipid X (2,3-diacyl-GlcN-1-phosphate) which is condensed with 2,3-diacyl-GlcN-1-phosphate producing a β,1'-6 linked disaccharide. This molecule is phosphorylated at the 4' position generating lipid IVA and two Kdo (3-deoxy-D-manno-oct-2-ulosonic acid) residues are added by KdtA. Finally, laurate and myristate chains are added by LpxL and LpxM to form the complete lipid A (Emiola et al., 2015; Putker et al., 2015; Raetz et al., 2007, 2009; Trent, 2004).

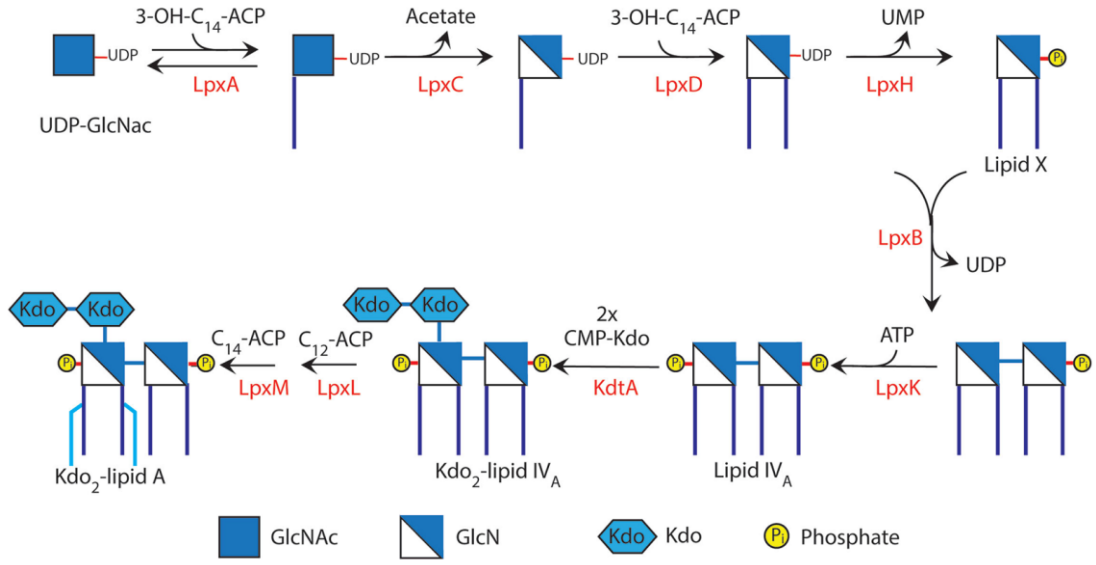


Figure 1.9. Schematic representation of the conserved lipid A biosynthetic pathway (from Putker et al. 2015)

The enzymes involved in each step are indicated in red. GlcNAc, N-acetylglucosamine; GlcN, glucosamine; Kdo, 3-deoxy-D-manno-oct-2-ulosonic acid; ACP, acyl carrier protein.

In marine *Synechococcus* the genes encoding LpxH and LpxK are absent, suggesting an alternative condensation reaction without lipid X. The lack of LpxK might be an adaptation to environments with low phosphate availability (Snyder et al., 2009).

After synthesis of lipid A in the inner membrane, glycosyltransferases attach the LPS core to the Kdo residue of the lipid A facing the cytoplasm. The biosynthesis of the core component of the LPS is encoded by genes in the *gmhD, waaQ* and *kdtA* operons, defined by the first gene of each transcriptional unit (David et al., 1998; Regué et al., 2001; Roncero and Casadaban, 1992)

In enterobacteria, the core LPS contains heptoses in the inner section mostly glucose, galactose and hexose, although in marine *Synechococcus* Kdo and heptose are absent (Holst, 2007; Raetz and Whitfield, 2002; Snyder et al., 2009; Sperandeo et al., 2009).

In *Salmonella enterica, E. coli* and *P. aeruginosa,* genes involved in the biosynthesis of the O-polysaccharide are clustered between *galF* (glucose-1-phosphate uridylyltransferase) and *gnd* (6-phosphogluconate dehydrogenase) and variation in the O-polysaccharide depends on the specific genes within this region. These genes encode enzymes required for nucleotide sugar precursor biosynthesis, glycosyltransferases that link saccharide units and the formation of the polysaccharide on a membrane bound undecaprenyl carrier lipid by polymerisation and translocation of the O-polysaccharide (Raymond et al., 2002; Samuel and Reeves, 2003).

Translocation of the core-lipid A and the O-antigen to the periplasmic side of the inner membrane is by an ABC transporter comprising MsbA and Wzx, respectively (Fig. 1.9). The proteins Wzy and Wzz continue to polymerise the O-antigen that is then transferred to the core-lipid A by WaaL (Abeyrathne et al., 2005; Wang and Quinn, 2010). In marine *Synechococcus* no homologues of the WaaL have been identified, suggesting alternative mechanisms for O-polysaccharide ligation.

Translocation to the outer membrane (Fig. 1.10) is performed by the lipopolysaccharide transport (lpt) machinery comprising the ABC transporter complex LptBCFG. These proteins comprise LptB located in the cytoplasm, LptA in the periplasm, with LptC and LptG in the inner membrane and LptD and LptE in the outer membrane. The LptBFG complex, together with LptC and LptA, translocates the LPS to the inner side of the outer membrane, where LptD and LptE export it to the outer side (Putker et al., 2015; Wu et al., 2006). LptF and LptG are conserved potential permeases whilst LptD is an essential outer membrane porin (Braun and Silhavy, 2002; Ruiz et al., 2008). Marine *Synechococcus* possess homologues of MsbA, LptA, LptB

and LptE suggesting that this same mechanism might be translocating LPS to the outer membrane (Snyder et al., 2009).

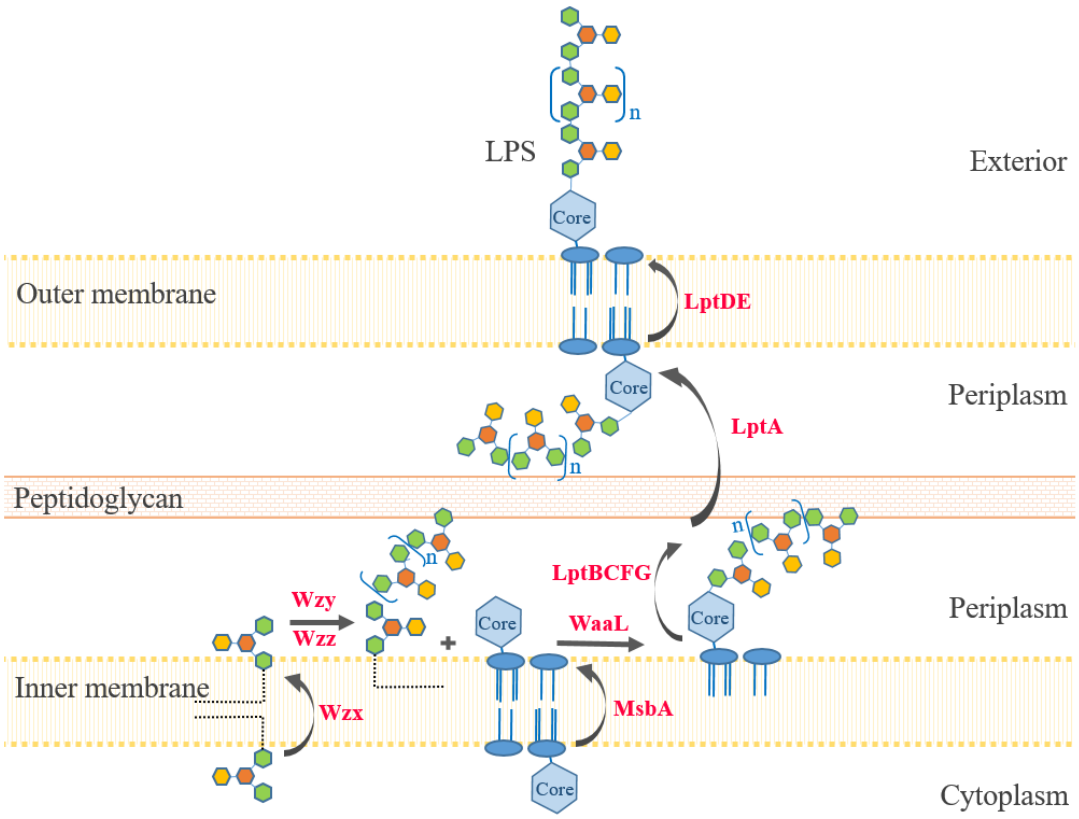


Figure 1.10. Export of LPS and its precursors in *E. coli* (modified from Wang & Quinn 2010)

1.6. Bacteria-bacteriophage co-evolution

Due to the high abundance of both bacteria and bacteriophages, they co-exist in all environments. Interactions between bacteriophages and their bacterial hosts are dynamic, with the high mortality of bacteria due to bacteriophage lysis giving rise to the evolution of many mechanisms to resist bacteriophage infection (see section 1.4; Fuhrman, 1999; Suttle, 2007; Weinbauer, 2004). Marine bacteriophages promote bacterial diversity by transferring genetic material (horizontal gene transfer) but also by forcing the acquisition and development of resistance mechanisms. This dynamic

leads to the evolution of the bacteriophage and host in a reciprocal manner, i.e. antagonistic co-evolution (Fig. 1.11; Brüssow et al. 2004; Buckling & Rainey 2002; Martiny et al. 2014).

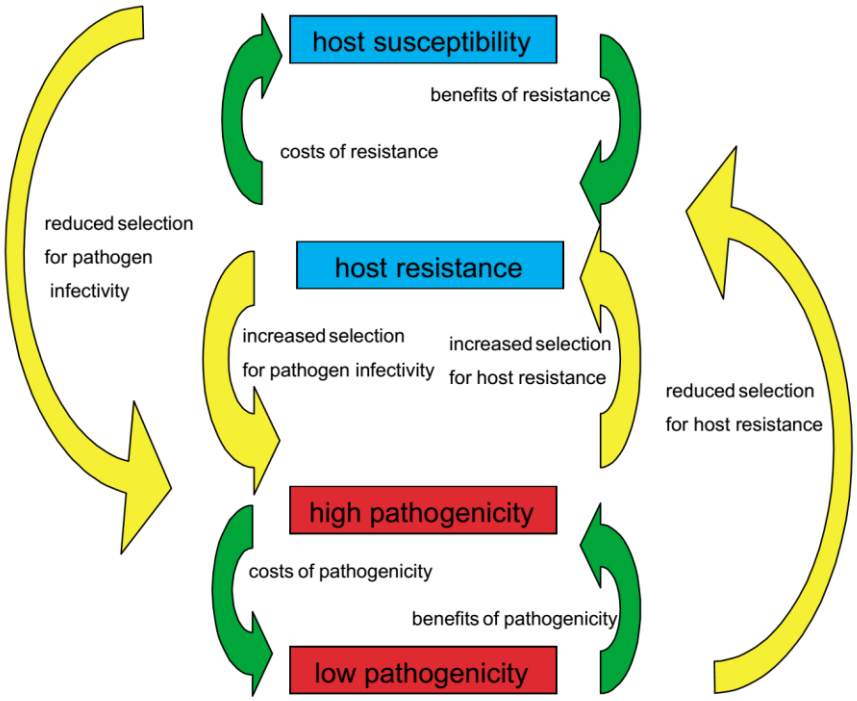


Figure 1.11. Schematic representation of host-parasite reciprocal co-evolution (from Woolhouse et al. 2002)

Green arrows indicate intra-specific selection; yellow arrows indicate inter-specific selection.

These co-evolutionary host-bacteriophage mechanisms have been best studied in γ-proteobacterial model organisms such as *E. coli* and *Pseudomonas fluorescens* largely due to their fast growth rates that allows the study of long term evolution experiments in a short time frame and with greater reproducibility (Brockhurst et al., 2007; Lenski and Levin, 1985; Mizoguchi et al., 2003).

The co-evolution of bacteriophage-bacteria (predator-prey or parasite-host) interactions belong to what is now accepted as the Red Queen hypothesis, which states

that species must constantly adapt, evolve and proliferate to pass on genes to the next generation but also to keep from going extinct when other 'species' are evolving and hence "maintain its fitness relative to the system it is co-evolving with" (van Valen, 1973). This hypothesis, first suggested by the evolutionary biologist Leigh van Valen in 1973, is based on the Alice in Wonderland 'through the looking-glass' book (Carroll, 1871) as a metaphor of what the Queen of Hearts (Red Queen) says to Alice: "it takes all the running you can do to keep in the same place", i.e. the continuous co-evolutionary "arms race" of ecological antagonists like bacteriophage and bacteria.

The selection of different genotypes can be either positive or negative depending on frequency, i.e. how common a certain genotype is. In other words, fitness will increase on the genotype becoming common in a positive frequency-dependent selection, whereas in a negative frequency-dependent selection fitness is the inverse of the genotype's representation in a population (Woolhouse et al., 2002).

Bacteria-bacteriophage systems tend to show a negative-dependent selection interaction, although other types of selection are also possible as a consequence of the cost of resistance (for review see Duffy & Forde 2009). Briefly, directional selection occurs when the bacteriophage reduces host diversity by increasing resistance or susceptibility in the host population. In contrast, disruptive selection increases host diversity by selecting for highly resistant and highly susceptible genotypes, whilst stabilising selection maintains a level of resistance.

The Red Queen hypothesis states that in order to remain competitive in the environment, often with limiting resources, organisms have to make trade-offs to acquire characteristics that will make them advantageous. Trade-offs allow the co-existence of organisms that can exploit different niches, which can incur a fitness cost under different circumstances (Bohannan et al., 2002; Brockhurst et al., 2006; Fuhrman and Schwalbach, 2003; Rodriguez-Valera et al., 2009; Winter et al., 2010). The study of trade-offs in environmental systems is complicated by the high number of varying environmental parameters organisms encounter. However, trade-offs have been investigated under laboratory conditions in several different bacteria-bacteriophage systems (see Perry et al. 2015; Lenski & Levin 1985; Brockhurst et al. 2007; Bohannan & Lenski 1999; Lenski 1988).

The cost of resistance or trade-off hypothesis suggests that the investment to acquire resistance comes with the loss of other competitive features, explaining polymorphisms in resistant strains. This and the Red Queen hypothesis are not mutually exclusive, but the main difference with the Red Queen hypothesis is that it does not necessarily acknowledge a co-evolutionary step between parasite and host (Woolhouse and Webster, 2000).

The Kill the Winner hypothesis proposes that bacteriophages control the abundance of the most active bacterial host (Breitbart, 2012; Rodriguez-Valera et al., 2009; Thingstad, 2000). In marine systems, high viral abundance has been suggested to control microbial diversity, acting as a selective pressure (Fuhrman and Suttle, 1993). In this respect, not the most abundant but the most active (winning) organism, i.e. the competition specialist, is under the control of viral selection, incorporating negative frequency-dependent selection in the killing the winner hypothesis (Thingstad, 2000).

Furthermore, if two populations are competing for a limiting nutrient/resource, i.e. competition specialist and defence specialist, because acquisition of resistance results in fitness costs, oligotrophic environments select for competition specialists compared to defence specialists selected in eutrophic environments where the fitness cost is mitigated with high nutrient availability (Thingstad and Lignell, 1997; Winter et al., 2010).

Such mechanisms may explain the high diversity of marine picocyanobacteria (section 1.1) which, as a result of the 'killing the winner' arms race requires continuous adaptation to viral resistance, a prerequisite feature for the maintenance of abundant populations of these ecologically important marine primary producers.

1.7. Aims and objectives

In this thesis, the major aim was to investigate mechanisms of cyanophage resistance in marine *Synechococcus*. The hypothesis tested was that host cell surface properties, specifically LPS, are implicated in cyanophage resistance in marine *Synechococcus*. Specific objectives that arose from this over-arching aim included the physiological, molecular and genetic characterisation of cyanophage resistant *Synechococcus* sp.

WH7803 mutants in comparison with the wild type in order to identify specific genes conferring cyanophage resistance.

Thus, during the course of this work the following objectives arose:

1. To determine the role that the *Synechococcus* sp. WH7803 cell surface plays in mediating cyanophage resistance in spontaneous cyanophage resistant *Synechococcus* sp. WH7803 mutants previously isolated by Jia (2009) and Spence (2010). Particular attention was paid to assessing changes in LPS profiles of these spontaneous mutants, as well as by targeted interposon mutagenesis in which specific components of the LPS biosynthetic pathway and outer membrane porin proteins were targeted (Chapter 3).

2. To identify the genetic differences that conferred cyanophage resistance in these spontaneous *Synechococcus* sp. WH7803 cyanophage resistance mutants using whole genome sequencing (WGS) of these and wild type *Synechococcus* sp. WH7803, and comparing these with the previously published genome sequence (accession number NC_009481.1) of *Synechococcus* sp. WH7803 (Dufresne et al., 2008), which was obtained via traditional clone library construction and Sanger sequencing technology and hence provides a single 'consensus' sequence (Chapter 4).

3. To assess chromosome copy number in marine *Synechococcus* by optimising a qPCR-based approach and to determine if ploidy levels vary during growth and under P deplete conditions in *Synechococcus* sp. WH7803 (Chapter 5).

4. To obtain and perform a phenotypic and genotypic (WGS) characterisation of cyanophage resistant mutants of the monoploid strain *Synechococcus* sp. WH7805 and compare their analysis with the cyanophage resistant *Synechococcus* sp. WH7803 mutants (Chapter 6).

5. To investigate the relationship between phage infection/resistance and grazing of *Synechococcus* sp. WH7803 by the nanoflagellates *Cafeteria roenbergensis*

and *Pteridomonas danica* and compare with cyanophage infected *Synechococcus* sp. WH7805 as a comparison grazer-prey system (Chapter 7).

Chapter 2

MATERIALS AND METHODS

2.1. Strains and culture conditions

2.1.1 Culture conditions and media

All cyanobacterial and protistan cultures were routinely maintained in Artificial Sea Water (ASW) medium (Table 2.1).

For experiments in P-deplete conditions, 10 mL of an exponentially-growing *Synechococcus* sp. WH7803 culture grown in ASW medium was inoculated into 100 mL P-deplete ASW medium (Table 2.1) and 100 mL ASW medium (i.e. P-replete). Three biological replicates were used for each condition. Cells from the six cultures were counted by flow cytometry (section 2.3) every 4 days and samples for qPCR (section 2.15) were taken every 10 days.

Synechococcus spp. used in this study were routinely grown in ASW liquid medium at 23°C under continuous illumination (10-30 µmol photons $m^{-2}s^{-1}$) and in semi-solid conditions using 0.4 % (w/v) agarose or clean agar in ASW (see Brahamsha 1996). Clean agar was prepared as in Millard (2009).

All *E. coli* strains were grown in lysogeny broth (LB; 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract and 10 g L^{-1} NaCl) with agitation or in solid 1.5 % (w/v) Bacto agar in LB at 37 C.

Component	ASW	P-deplete ASW	SN
NaCl	428 mM	428 mM	-
$MgCl_2 \cdot 6H_2O$	9.8 mM	9.8 mM	-
KCl	6.7 mM	6.7 mM	-
$NaNO_3$	8.8 mM	8.8 mM	8.5 mM
$MgSO_4 \cdot 7H_2O$	14.2 mM	14.2 mM	-
$CaCl_2 \cdot 2H_2O$	3.4 mM	3.4 mM	-
Tris base	9.1 mM	9.1 mM	-
$K_2HPO_4 \cdot 3H_2O$	172 µM	-	90 µM
H_3BO_3	46.3 µM	46.3 µM	-
$MnCl_2 \cdot 4H_2O$	9.2 µM	9.2 µM	7.0 µM
$ZnSO_4 \cdot 7H_2O$	0.95 µM	0.95 µM	0.77 µM
$Na_2MoO_4 \cdot 2H_2O$	1.6 µM	1.6 µM	1.61 µM
$CuSO_4 \cdot 5H_2O$	27.5 nM	27.5 nM	-
$Co(NO_3)_2 \cdot 6H_2O$	0.25 µM	0.25 µM	86 nM
$FeCl_3 \cdot 6H_2O$	11.1 µM	11.1 µM	-
EDTA	1.39 µM	1.39 µM	15 µM
Na_2CO_3	-	-	100 µM
Ferric ammonium citrate	-	-	22.6 µM
Citric acid hydrate	-	-	29.7 µM
Thiamine-HCl	-	-	29.7 nM
Biotin	-	-	2 nM
Vitamin B_{12}	-	-	0.37 nM
Folic acid	-	-	2.3 nM
p-Aminobenzoic acid	-	-	36 nM
Nicotinic acid	-	-	406 nM
Ca-Pantothenate	-	-	210 nM
Pyridoxin-HCl	-	-	243 nM
Seawater (autoclaved)	-	-	750 mL

Table 2.1. Composition of culturing media.

ASW and P-deplete ASW medium were prepared with ultrapure water. The pH of the medium was adjusted to pH 8.0 using HCl prior to autoclaving for 15 min at 121ºC (Wilson et al., 1996). SN was prepared with autoclaved ultrapure water and by filter-sterilising all components apart from EDTA, $NaNO_3$, K_2HPO_4 and $NaCO_3$ that were previously autoclaved (Waterbury et al., 1986).

2.1.2 Marine *Synechococcus* spp.

The marine *Synechococcus* strains used in this study are listed in Tables 2.2, 2.3 and 2.4. All wild type strains (Table 2.2) are axenic. Contamination medium, comprising 1% (w/v) agar and 0.5% (w/v) yeast extract in ASW medium, were routinely used to detect contaminating heterotrophic bacteria.

Strain	WH7803	WH7805	WH8102
Year of isolation	1978	1978	1981
Location	33°44.9'N, 67°29.8'W	33°44.8'N, 67°30'W	22°29.7'N, 65°36'W
Sub-cluster	5.1B	5.1B	5.1A
Clade	V	VIa	IIIa
Pigment type	3a	2	3c
Genome status	Complete	WGS	Complete
Genome size	2,366,981 bp	2,627,147 bp	2,434,429 bp
Contigs	1	2	1
CDS	2,587	2,900	2,803
GC %	60.2	57.5	59.4

Table 2.2. Axenic *Synechococcus* spp. used in this study.

(Waterbury et al. 1986, Cyanorak: http://abims.sb-roscoff.fr/cyanorak)

Cyanophage-resistant *Synechococcus* strains (Table 2.3) were routinely maintained in 40 mL ASW in 50 mL vented culture flasks (Falcon) with no agitation and with an aliquot of the corresponding cyanophage lysate added every 2-3 weeks. These strains correspond to 10 *Synechococcus* sp. WH7803 mutants previously isolated (Spence 2010) resistant to cyanophage S-RSM42 (A-J), two mutants resistant to cyanophages

S-RSM42 and S-PM2 (R1 and R2) and one mutant (Jia, 2009) resistant to cyanophage S-PM2 (PHR). During the course of this work, 10 *Synechococcus* sp. WH7805 mutants resistant to S-PM2 (JC1 - JC10) were obtained and characterised. Heterotrophic contamination of these cultures was determined using contamination medium and by flow cytometry with SYBR Green I staining (Invitrogen) (see section 2.3).

Name	Strain	Cyanophage	Reference
A	*Synechococcus* sp. WH7803S-RSM42AR	S-RSM42	Spence (2010)
B	*Synechococcus* sp. WH7803S-RSM42BR	S-RSM42	Spence (2010)
C	*Synechococcus* sp. WH7803S-RSM42CR	S-RSM42	Spence (2010)
D	*Synechococcus* sp. WH7803S-RSM42DR	S-RSM42	Spence (2010)
E	*Synechococcus* sp. WH7803S-RSM42ER	S-RSM42	Spence (2010)
F	*Synechococcus* sp. WH7803S-RSM42FR	S-RSM42	Spence (2010)
G	*Synechococcus* sp. WH7803S-RSM42GR	S-RSM42	Spence (2010)
H	*Synechococcus* sp. WH7803S-RSM42HR	S-RSM42	Spence (2010)
I	*Synechococcus* sp. WH7803S-RSM42IR	S-RSM42	Spence (2010)
J	*Synechococcus* sp. WH7803S-RSM42JR	S-RSM42	Spence (2010)
R1	*Synechococcus* sp. WH7803R1	S-RSM42-S-PM2	Spence (2010)
R2	*Synechococcus* sp. WH7803R2	S-RSM42-S-PM2	Spence (2010)
PHR	*Synechococcus* sp. WH7803RS-PM2	S-PM2	Jia (2009); Zwirglmaier (2009)
JC1	*Synechococcus* sp. WH7805RS-PM2JC1	S-PM2	This work
JC2	*Synechococcus* sp. WH7805RS-PM2JC2	S-PM2	This work
JC3	*Synechococcus* sp. WH7805RS-PM2JC3	S-PM2	This work
JC4	*Synechococcus* sp. WH7805RS-PM2JC4	S-PM2	This work
JC5	*Synechococcus* sp. WH7805RS-PM2JC5	S-PM2	This work
JC6	*Synechococcus* sp. WH7805RS-PM2JC6	S-PM2	This work
JC7	*Synechococcus* sp. WH7805RS-PM2JC7	S-PM2	This work
JC8	*Synechococcus* sp. WH7805RS-PM2JC8	S-PM2	This work
JC9	*Synechococcus* sp. WH7805RS-PM2JC9	S-PM2	This work
JC10	*Synechococcus* sp. WH7805RS-PM2JC10	S-PM2	This work

(Previous page)

Table 2.3. *Synechococcus* cyanophage resistant strains used in this study.

Three *Synechococcus* sp. WH7803 partial interposon mutants (Table 2.4) were maintained in ASW medium containing up to 1 mg mL^{-1} apramycin. These strains are neither axenic nor fully segregated.

Strain	Description	Reference
Synechococcus sp. WH7803Δ*0192*::aac(3)-IV	*rmlB* homologue, rhamnose biosynthetic pathway	Spence (2010)
Synechococcus sp. WH7803Δ*1767*::aac(3)-IV	*wbaP* homologue, hexose-1-P transferase	Spence (2010)
Synechococcus sp. WH7803Δ2235-2236::aac(3)-IV	double deletion of the genes encoding for P-stress porins	Spence (2010)

Table 2.2.4. *Synechococcus* sp. WH7803 interposon mutants.

These mutants were constructed by double cross-over replacing the target gene with an apramycin-resistance cassette (aac(3)-IV) (Spence, 2010).

2.1.3 Storage of *Synechococcus* spp. cultures

For long term storage, 15-40 mL of a late exponentially growing *Synechococcus* culture was centrifuged at 3220 *g* for 20 min. The pellet was transferred to a 1.8 mL cryovial and centrifuged at 16060 *g* for 3 min to discard any leftover medium. The pellet was quickly re-suspended in 1 mL 100 % (v/v) DMSO and snap-frozen in liquid nitrogen. All *Synechococcus* spp. were maintained in DMSO at -80ºC.

2.1.4 Cyanophages

The cyanophage isolates used in this work were S-PM2, S-RSM4, S-RSM42, S-RSM57, S-RSM76, S-RIM1, S-RIM11 and S-RIM34 (Table 2.5). These were concentrated and purified by caesium chloride (CsCl) gradient standard techniques (section 2.4) and titrated by plaque assay (section 2.5). A lysate filtered through a 0.2 µm pore size filter was used for most experiments, which was stored at 4ºC in the dark. Infection of *Synechococcus* sp. WH7805 by these cyanophages was determined by a spot assay (section 2.6). The presence of plaques indicated successful infection.

	Location	Year of isolation	Reference
S-PM2	English Channel 50°18'N, 4°12'W	1992	Wilson et al. (1993); Wilson et al. (1996); Mann et al. (2005)
S-RSM4	Gulf of Aqaba 29°28’N, 34°55’E	1999	Millard & Mann (2006)
S-RSM42	Gulf of Aqaba 29°28’N, 34°55’E	1999	Millard & Mann (2006)
S-RSM57	Gulf of Aqaba 29°28’N, 34°55’E	1999	Millard & Mann (2006)
S-RSM76	Gulf of Aqaba 29°28’N, 34°55’E	1999	Millard & Mann (2006)
S-RIM1	Mount Hope Bay 41°39’ N, 71°15’ W	1999	Marston & Sallee (2003)
S-RIM11	Mount Hope Bay 41°39’ N, 71°15’ W	2000	Marston & Sallee (2003)
S-RIM34	Mount Hope Bay 41°39’ N, 71°15’ W	2002	Marston & Sallee (2003)

Table 2.5. Cyanophages used in this study.

2.1.5 Protists

The protist grazers used in this work were *Pteridomonas danica* and *Cafeteria roenbergensis*, two species belonging to the Heterokonta division of the Chromalveolata kingdom. These cultures were obtained from the National Oceanography Centre (NOC), Southampton, and from the Culture Collection of Algae and Protozoa (CCAP). These cultures were maintained in 50 mL vented culture flasks (Falcon®) with no agitation at 13°C and room temperature (23°C) in ASW medium supplemented with *Halomonas* sp. (grown in 1 % (w/v) Bacto agar in Difco marine broth) or *Synechococcus* sp. WH8102 as a nutrient (prey) source.

2.1.6 *Escherichia coli*

Competent *E. coli* DH5α (K-12 derivative) was used for general cloning and *E. coli* MC1061 (host strain for the conjugative plasmid pRK24) was used for conjugation with *Synechococcus* sp. WH7803 (Table 2.6).

Strain	Genotype	Reference
E. coli DH5α	F^- *endA1 gln*V44 *thi*-1 *recA*1 *relA*1 *gyrA*96 *deoR nupG* Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, *hsd*R17(r_K^- m_K^+), λ^-	Meselson & Yuan (1968)
E. coli MC1061	F^- araD139 Δ(ara, leu)7697 ΔlacX74 $galU^-$ $galK^-$ hsr^- hsm^+ strA	Casadaban & Cohen (1980)

Table 2.6. *E. coli* strains used in this study.

2.2. . Polymerase chain reaction (PCR) conditions and oligonucleotides

Primers used for identification of segregation in *Synechococcus* sp. WH7803 interposon mutants (section 3.3.2) are listed in Table 2.7, and are located either flanking or internal to the targeted gene. PCR reactions with these primers were performed using MyTaq Red Mix (Bioline) following the manufacturer's guidelines unless otherwise stated, with 35 cycles and an annealing temperature of 52ºC.

Name	Sequence 5'-3'	Reference
F_Porin_Flanking	GGTGGGTAATGCAACAGATT	Spence (2010)
R_Porin_Flanking	TGAGGACAAGAATCTCGAGG	Spence (2010)
F_Porin_Internal	GGCAGAAGACCGTTTCACCC	Spence (2010)
R_Porin_Internal	GAGCTCGTCGGTCACTTCGG	Spence (2010)
F_1767_Flanking	GGGATTGCTCAGCGCATCAG	Spence (2010)
R_1767_Flanking	TGGATGTGGCTGGCGGCAGC	Spence (2010)
F_1767_Internal	CGCTGTTCGGTGTCTTGGCA	Spence (2010)
R_1767_Internal	TCAAGCCTGGACGGACAGCC	Spence (2010)
F_0192_Flanking	GGCGTGAGAGCGATCTACTGC	Spence (2010)
R_0192_Flanking	GGATCAGCAGCCAGTGCAGCC	Spence (2010)
F_0192_Internal	GGTGCACTACCAGTTCAATCC	Spence (2010)
R_0192_Internal	GATCGCGCACCAGTTCAGCC	Spence (2010)
ApraSeq	CGGTTGATCCGCTCCCGCGAC	Spence (2010)

Table 2.7. PCR primers used to determine segregation of *Synechococcus* sp. WH7803 interposon mutants.

Primers used for targeted mutagenesis are listed in Table 2.8. The regions selected correspond to a fragment of 346 bp within the SynWH7803_1386 gene encoding for a two-component system sensor histidine kinase and 395 bp within the SynWH7803_0102 gene encoding for a glucose-1-phosphate thymidylyltransferase. The genes were selected based on the findings of Marston et al. (2012). The PCR was performed using the proofreading Platinum Pfx DNA Polymerase (Invitrogen) following the manufacturer's instructions, for 35 cycles with an annealing temperature of 55ºC.

Name	Sequence 5'-3'	Target
rmlAF	CACGACCAGGCGGCGTTTCA	SynWH7803_0102
rmlAR	ACCTGACGGGCCCTCTCCAC	
baeSF	ACCTCCGCACTCGTGAAGCA	SynWH7803_1386
baeSR	CAAGACACCATGCGTTGAAGAA	

Table 2.8. Primers used for targeted mutagenesis in *Synechococcus* sp. WH7803

To confirm the mutations identified in *Synechococcus* sp. WH7805 cyanophage-resistant mutants (see Chapter Six), Sanger sequencing of PCR products was performed (GATC Biotech) using the primer sets listed in Table 2.9. The PCR was performed using KAPA HiFi PCR Kit (Kapa Biosciences) following the manufacturer's instructions, for 35 cycles with an annealing temperature of 52ºC.

Name	Sequence 5'-3'
08977_F	CGATGGATCACACGTCAAGAA
08977_R	CTACGCGACCGTTATTTCTTCT

Table 2.9. Sequencing primers used to confirm the mutations in *Synechococcus* sp. WH7805 cyanophage-resistant mutants by Sanger sequencing.

The primers target a region of 1,158 bp of the WH7805_08977 gene that encodes a putative alpha-glycosyltransferase.

To determine chromosome copy number (see section 2.15) qPCR was performed using the primers listed in Table 2.10. These primers target a region of 63 nt of the SynWH7803_0934 gene in *Synechococcus* sp. WH7803 and 78 nt of the WH7805_00035 gene in *Synechococcus* sp. WH7805.

Name	Sequence 5'-3'	Organism - Gene ID
0934_F	TTTGTCCCTCGGCAGGC	*Synechococcus* sp. WH7803 SynWH7803_0934
0943_R	CGGAACGATGCAGTGGAGT	
00035_F	CCAGTTCACATCAGAAGGAGTA	*Synechococcus* sp. WH7805 WH7805_00035
00035_R	TACCACAGTGTGTTCAGATAGG	

Table 2.10. qPCR primers for determination of chromosome copy number in *Synechococcus* sp. WH7803 and *Synechococcus* sp. WH7805.

2.3. Cell enumeration by flow cytometry

Flow cytometry was performed using a BD FACScan. Cells (*Synechococcus* or grazers) for flow cytometry were fixed with 1 % (v/v) paraformaldehyde (PFA) or 0.5 % (v/v) glutaraldehyde for 1 h at 4ºC in the dark. For nucleic acid staining, cells were stained with SYBR Green I (Invitrogen) following the manufacturer's specifications for flow cytometry analysis, using a 10,000 fold dilution. Cells were counted using 0.5 µm multifluorescent (Polysciences) beads at a known concentration as an internal standard.

2.4. Cyanophage purification

Purification of cyanophages was carried out using the protocol of Boulanger (2009). Briefly, cyanophage lysates were obtained by infecting a *Synechococcus* sp. WH7803 culture during exponential growth ($OD_{750\,nm}$ 0.35 – 0.4) until complete lysis. The cyanophage lysate was then centrifuged at 3,220 *g* for 15 min to discard cell debris and the supernatant incubated overnight at 4ºC in the dark with 10% (w/v) PEG 6000 final concentration. The concentrated cyanophages were centrifuged at 12,000 *g* for 30 min at 4ºC and the supernatant discarded. The pellet was carefully re-suspended with not more than 3 mL ASW medium (section 2.1.1) and layered on top of a CsCl gradient previously prepared in ultracentrifuge tubes (2 mL of d = 1.7, 3 mL of d = 1.5 and 3 mL of d = 1.45). CsCl gradients were subjected to ultracentrifugation at 217,500 *g* for 2 h at 4ºC. Cyanophages were extracted from the gradient (Fig. 2.1) using a 21 gauge needle and dialysed overnight at 4ºC in ASW medium using a size 2 (18/32'') 12-14 kDa molecular weight cut-off membrane (Visking-Medicell) prepared by incubating for 30 min in 2 % (w/v) $NaHCO_3$ and 1 mM EDTA at 80ºC and rinsed thoroughly with distilled water.

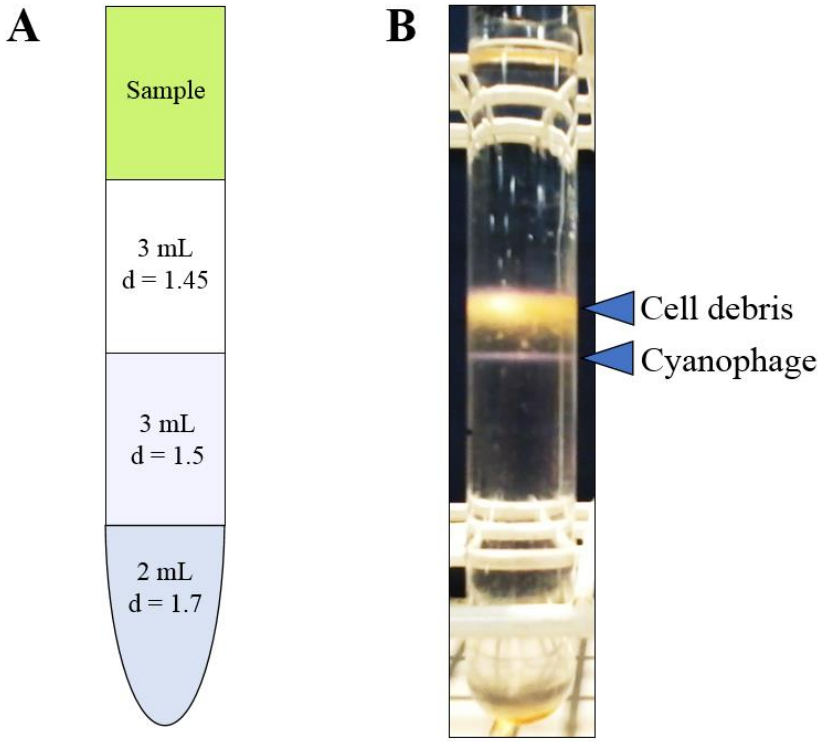


Figure 2.1. CsCl gradient for cyanophage purification.

(A) Schematic representation of a CsCl density gradient on an ultracentrifuge tube. (B) Photography of the gradient after centrifugation. The arrows indicate cell debris located on top of the band of concentrated cyanophages between d = 1.45 – 1.5.

2.5. Cyanophage enumeration by plaque assay

Enumeration of cyanophages was carried out using plaque assays, based on the protocol of Millard (2009). The enumeration was performed in triplicate using a dilution series of up to 10^{-9}. The final titre was the mean of 3 replicates. For the preparation of 40 plates, 1 L of exponentially growing *Synechococcus* sp. WH7803 was concentrated by centrifugation at 6,000 *g* for 20 min and re-suspended in 20 mL ASW medium. Aliquots of 0.5 mL were incubated for 1 h with a serial dilution of the cyanophage at room temperature, then mixed with 2.5 mL molten 0.4 % (w/v) clean agar or agarose in ASW medium (see section 2.1.2) at a temperature of 30-40ºC and poured into previously prepared 1 % (w/v) clean agar or agarose in ASW plates. Agar plates were incubated at 23ºC under continuous illumination of 10-

30 µmol photons m^{-2} s^{-1}. After 3 to 10 days plaques appeared and the titre was calculated.

2.6. Cyanophage enumeration by spot assay

Enumeration of cyanophages by spot assay was performed by modifying the protocol for plaque assay (section 2.5) of Millard (2009). This method was used for experiments where several repetitions were required as ca. 30 "spots" can be tested per plate hence reducing the amount of material needed per experiment.

An exponentially growing *Synechococcus* sp. WH7803 culture ($OD_{750\ nm}$ = 0.35–0.4) was concentrated by centrifugation at 6000 *g* for 20 min and re-suspended in ASW medium. For the preparation of 10 plates, 300 mL culture was used, re-suspending in 6 mL ASW medium. Aliquots of 0.5 mL were mixed with 2.5 mL molten 0.4 % (w/v) clean agar or agarose in ASW medium (see section 2.1.2) at a temperature of 30-40ºC and poured into previously prepared 1 % (w/v) clean agar or agarose in ASW plates. Agar plates were incubated at 23ºC under continuous illumination of 10-30 µmol photons m^{-2} s^{-1} for 24 h.

A serial dilution of the cyanophage with sterile ASW medium was prepared and 10 µL aliquots were "spotted" in the agar plates in an organised manner. For adsorption assays (section 2.8), the serial dilution was prepared with the supernatant of the assayed samples. After 3 to 10 days plaques appeared and the titre was calculated. The experiment was performed in triplicate and the titre was the mean of the three replicates.

2.7. Isolation of cyanophage-resistant *Synechococcus* sp. WH7805 mutants

For isolation of cyanophage-resistant mutants, 40 mL of an exponentially growing *Synechococcus* sp. WH7805 culture in triplicate was infected with cyanophage S-PM2 previously filtered using a 0.2 µm pore size filter and incubated at 23ºC under continuous illumination of 20 µmol photons m^{-2} s^{-1} with agitation until the culture lysed completely. After ca. 4 weeks, when re-growth of survivors was observed, the cultures were grown in semi-solid ASW medium (section 2.1.1) and incubated at 23ºC

under continuous illumination of 10 µmol photons m^{-2} s^{-1}. Colonies were grown in 2 mL ASW medium with 100 µL S-PM2 lysate and transferred to 40 mL ASW medium with 200 µL S-PM2 lysate.

2.8. Cyanophage adsorption assay

Adsorption of cyanophages was determined for the 13 cyanophage-resistant *Synechococcus* sp. WH7803 and 6 cyanophage-resistant *Synechococcus* sp. WH7805 mutants. *Synechococcus* cells were counted by flow cytometry (see section 2.3) and 1 x 10^6-1 x 10^7 cells were mixed with cyanophage (either S-PM2 or S-RSM42) at a multiplicity of infection (MOI) of 1-10 in 1 mL final volume using 1.5 mL clear microcentrifuge tubes (Eppendorf). An aliquot of 100 µL was taken at times 0 and 60 min, centrifuged at 16060 *g* for 5 min and the supernatant titrated by spot assay (see section 2.6). A positive control for cyanophage adsorption was used, comprising an exponentially-growing *Synechococcus* sp. WH7803 culture ($OD_{750\ nm}$ = 0.35–0.4). The experiment was repeated twice in triplicate for each cyanophage.

2.9. Lipopolysaccharide (LPS) extraction and visualisation

2.9.1 LPS micro-extraction and silver staining

LPS from 12 *Synechococcus* sp. WH7803 mutants resistant to S-RSM42 was extracted based on the method of Spence (2010). Briefly, 1 mL of culture in stationary phase ($OD_{750\ nm}$ > 0.5) was centrifuged at 16060 *g* for 5 min, the pellet re-suspended in 100 µL of LPS extraction buffer (0.25 mM Tris-HCl pH 7.5, 4 % (v/v) SDS and 0.5 g L^{-1} proteinase K) and incubated overnight at 60ºC. 100 µL loading buffer (0.125 mM Tris-HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 10 % (v/v) β-mercaptoethanol, 0.004 % (w/v) bromophenol blue) with 0.5 g L^{-1} proteinase K were added to the sample and incubated for 2 h at 60ºC.

LPS samples (15 µL) were boiled for 10 min, centrifuged at 16,060 *g* for 10 min and separated using a 12% Bis-Tris protein gels (NuPAGE Novex, 1.0 mm, 12 well) ran at 100 V for 120 min with 1X running buffer (NuPAGE® MOPS SDS Running Buffer) and visualised by silver staining following Tsai & Frasch (1982).

2.9.2 LPS extraction and visualisation by commercial kits

Extraction of LPS was carried out with 15 mL of a late exponentially-growing *Synechococcus* sp. WH7803 and WH7805 culture ($OD_{750\ nm}$ = 0.5) and cyanophage-resistant *Synechococcus* spp. mutants in stationary phase, including 13 cyanophage-resistant *Synechococcus* sp. WH7803 and 10 cyanophage-resistant *Synechococcus* sp. WH7805 mutants. The cells were centrifuged at 3,220 *g* for 20 min at 4°C and the extraction performed using a LPS extraction kit (Intron biotechnology) following the manufacturer's instructions.

LPS samples (15 µL) were boiled for 5 min in 5 µL loading buffer (0.125 mM Tris-HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 10 % (v/v) β-mercaptoethanol, 0.004 % (w/v) bromophenol blue) and then separated using a 12% Bis-Tris protein gels (NuPAGE Novex, 1.0 mm, 12 well) ran at 150 V for 80 min with 1X running buffer (NuPAGE® MOPS SDS Running Buffer). LPS was visualised using Pro-Q Emerald 300 LPS gel stain kit (Molecular Probes) following the manufacturer's instructions.

2.10. Preparation of chemocompetent cells

E. coli cells were grown overnight in 5 mL LB medium at 37ºC. 1 mL was used to inoculate 100 mL of LB and grown at 37ºC with agitation until exponential growth was reached ($OD_{600\ nm}$ = 0.4 - 0.6). The cells were incubated on ice for 5 min and centrifuged at 2,013 *g* for 5 min at 4ºC. The pellet was re-suspended in 50 mL cold sterile 0.1 M $CaCl_2$, incubated in ice for 30 min and centrifuged at 2013 *g* for 15 min at 4ºC. The pellet was re-suspended in 1 – 5 mL of cold sterile 70 % (v/v) $CaCl_2$ in glycerol. 30 µL aliquots were placed in 1.5 mL microcentrifuge tubes, snap frozen with liquid nitrogen and stored at -80ºC.

2.11. Chemical transformation

Chemocompetent *E. coli* cells were thawed in ice. 270 µL of cold sterile 0.1 M $CaCl_2$ were added to 30 µL chemocompetent cells. 100 µL aliquots were mixed with 10 µL of DNA and incubated in ice for 30 min, following a heat shock of 45 s at 42 ºC and put back in ice for 5 min. 1 mL of LB was added to the sample following incubation

at 37 ºC for 1 - 1.5 h. The sample was centrifuged at 16,060 *g* for 3 min. The pellet was re-suspended carefully with 100 µL LB and incubated overnight at 37 ºC in solid LB previously prepared with the corresponding antibiotic.

2.12. Targeted mutagenesis

Construction of *Synechococcus* sp. WH7803 mutants was carried out by single cross over mutagenesis as described by Brahamsha (1996). Briefly, a 300-400 bp fragment of the gene of interest (PCR conditions in section 2.2 and primer sets listed in Table 2.8) was cloned into pGEM-T Easy vector (Promega) following the manufacturer's instructions and transformed into *E. coli* DH5α (see sections 2.10 and 2.11). The vector was subsequently extracted and purified using a Wizard Plus SV Minipreps DNA Purification System (Promega) and the insert released by digesting with *Eco*RI (New England BioLabs) following the manufacturer's instructions.

The suicide vector pMUT100 (plasmid based on pBR322 containing a Kan^R cassette) was digested with *Eco*RI (New England BioLabs) following the manufacturer's instructions and ligated with the target fragment previously obtained with a T4 DNA Ligase (Promega). The resulting construct (Fig. 2.2) was Sanger sequenced (GATC Biotech) to confirm the correct insertion.

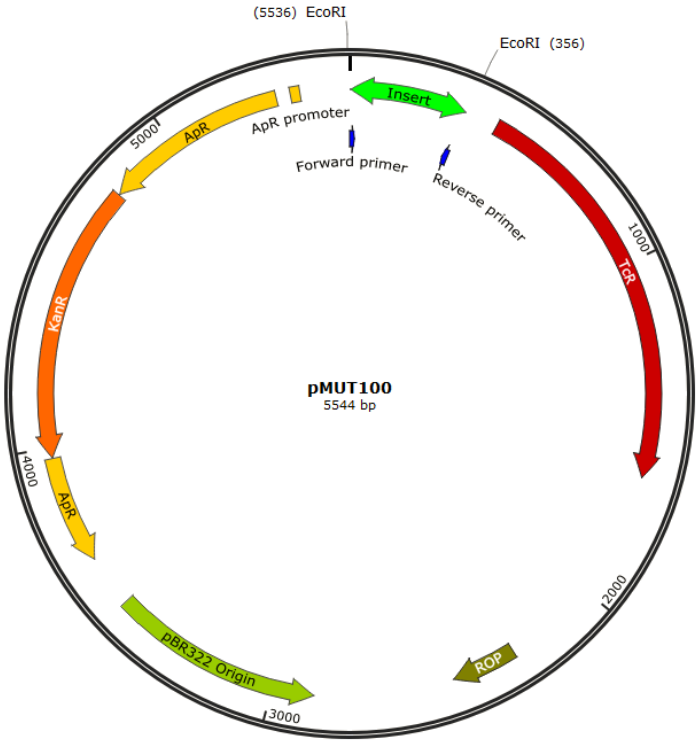


Figure 2.2. Schematic representation of the construct used for targeted mutagenesis in *Synechococcus* sp. WH7803.

The suicide vector pMUT100 was ligated with a 300-400 bp fragment (Insert) described in section 2.2. *Eco*RI restriction sites are labelled. TcR = tetracycline resistance, ROP = regulatory protein rop, ApR = ampicillin resistance, KanR = kanamycin resistance.

The construct was transformed into *E. coli* MC1061 (see section 2.11) containing the conjugal plasmid pRK24 and the pRL528 helper plasmid. *E. coli* MC1061 was then conjugated with *Synechococcus* sp. WH7803 using different donor:recipient ratios (1:1, 5:1, 1:5, 2:1 and 1:2) and plated on 0.3 % (w/v) clean agar-ASW (Millard 2009) or SN solid medium with 50 µg mL^{-1} kanamycin, incubated in low light (10 µmol photons m^{-2} s^{-1}) for 48 h at 23ºC and then at 20 µmol photons m^{-2} s^{-1} until colonies appeared. Colonies were subsequently grown in 2 mL ASW medium with 50 µg mL^{-1} kanamycin and analysed by PCR to confirm the mutation (PCR conditions and primers are described in section 2.2 and Table 2.8).

2.13. Extraction of genomic DNA from marine *Synechococcus*

15 mL of *Synechococcus* spp. cells were centrifuged at 3,220 *g* for 20 min and the pellet re-suspended in 567 µL TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0). 30 µL 10 % (w/v) SDS and 3 µL 20 mg mL^{-1} Proteinase K were added and the sample was incubated for 4 h at 60ºC. An equal volume of phenol was added to the sample following centrifugation for 5 min at 16,060 *g*. The aqueous layer was then mixed with an equal volume of chloroform:isoamyl alcohol (24:1 v/v) and centrifuged for 5 min at 16,060 *g*. 0.1 volumes of 3.5 M sodium acetate were added to the aqueous layer. The sample was incubated overnight at -20ºC with an excess of ethanol and centrifuged at 16,060 *g* for 30 min at 4ºC. The supernatant was discarded and the pellet washed with 1 mL of 70 % (v/v) ethanol. The sample was incubated for 15 min at 4ºC and centrifuged at 16,060 *g* for 20 min at 4ºC. The supernatant was removed completely. The pellet was re-suspended in 100 µL of nuclease free water (Ambion) and purified with DNeasy Plant mini spin columns (Qiagen).

After purification, the DNA was quantified by Nanodrop and QuantiFluor dsDNA System (Promega) and the integrity was visualised by agarose gel electrophoresis and ethidium bromide staining. A minimum of 1 µg of DNA was sent for Illumina sequencing at the Centre for Genomic Research, University of Liverpool.

2.14. Bioinformatic analysis

DNA libraries were prepared using the TruSeq (Illumina) protocol and sequenced with the MiSeq (Illumina) platform at the Centre for Genomic Research, University of Liverpool, generating reads of either 2 x 150 bp or 2 x 250 bp (Table 2.11).

Wild type *Synechococcus* sp. WH7805 re-sequencing included 2 DNA samples (Table 2.11): one taken before obtaining the cyanophage resistant mutants (May 2014; see section 2.7) and one 7 months after, taken at the same time as the cyanophage resistant mutants (Dec. 2014). Wild type *Synechococcus* sp. WH7803 re-sequencing also included 2 DNA samples (Table 2.11): Oct. 2013 and Dec. 2014.

Raw sequences were trimmed for Illumina adaptors using Cutadapt 1.1 option -O 3 and further trimmed using Sickle version 1.2 with a minimum window quality score of 20 by the Centre for Genomic Research, University of Liverpool.

Strain(s)	Paired-end	DNA sample date
Synechococcus sp. WH7803 WT	250 bp	Oct. 2013
Synechococcus sp. WH7803 WT	150 bp	Dec. 2014
cyanophage resistant *Synechococcus* sp. WH7803 mutants	250 bp	Oct. 2013
Synechococcus sp. WH7805 WT	150 bp	May 2014
Synechococcus sp. WH7805 WT	150 bp	Dec. 2014
cyanophage resistant *Synechococcus* sp. WH7805 mutants	150 bp	Dec. 2014

Table 2.11. Whole genome sequencing of marine *Synechococcus*.

WT: wild type *Synechococcus* spp. (Table 2.2), cyanophage resistant mutants (Table 2.3).

2.14.1 Read alignment

Whole genome sequencing reads from marine *Synechococcus* were obtained in fastq format and the quality of the reads was checked using FastQC v.0.10.1 (www.bioinformatics.babraham.ac.uk/projects/fastqc).

The aligners used include Bowtie2 (Langmead and Salzberg, 2012), BWA mem v.0.7.10 (Li, 2013) and NextGenMap v.0.4.11 (Sedlazeck et al., 2013).

For Bowtie2 alignment, a DNA index was constructed using bowtie2-build with the reference genome in fasta format. The paired-end sequencing reads in fastq format were aligned to that index (option –x) with Bowtie2, producing a SAM output file

(option –S). The options --very-sensitive (end-to-end alignment) or --very-sensitive-local (soft clip of the reads) were used for the different samples.

For BWA mem alignment, a DNA index was constructed using bwa index with the reference genome in fasta format. The paired-end sequencing reads in fastq format were aligned to that index, producing a SAM output file.

For NextGenMap alignment, the paired-end (option –p) sequencing reads in fastq format (option –q) were aligned to the reference genome in fasta format (option –r), producing a SAM output file.

Alignment of *Synechococcus* sp. WH7803 wild type reads to the published genome sequence of *Synechococcus* sp. WH7803 (accession number NC_009481.1) was done using Bowtie2 with the --very-sensitive option, BWA mem and NextGenMap.

Alignment of cyanophage resistant *Synechococcus* sp. WH7803 mutants reads to the published genome sequence of *Synechococcus* sp. WH7803 (NC_009481.1) and S-PM2 (NC_006820.1) was done using Bowtie2 with the --very-sensitive and --very-sensitive-local option and BWA mem.

Alignment of *Synechococcus* sp. WH7805 wild type and cyanophage resistant mutants reads to *Synechococcus* sp. WH7805 (NZ_AAOK00000000.1) was done using BWA mem.

Reads from re-sequenced wild type *Synechococcus* sp. WH7803 (Marston et al. 2012) were obtained from CAMERA (Community cyberinfrastructure for Advanced Microbial Ecology Research and Analysis, (Sun et al., 2011)) in fasta format and were aligned to *Synechococcus* sp. WH7803 (NC_009481.1) genome with Bowtie2 option --very-sensitive-local.

Reads in fastq format from cyanophage resistant *Prochlorococcus marinus* MED4 mutant R28 (Avrani et al., 2011) were kindly sent by Dr. D. Lindell and were aligned to the *Prochlorococcus marinus* MED4 genome (NC_005072.1) using BWA mem.

The output files produced by three aligners in SAM format that were processed using Samtools v.0.1.18 (Li et al., 2009). For conversion to binary format Samtools view

options –bS (input in SAM format and output in BAM) -F4 (skip alignments when the query sequence is unmapped) was used.

Alignment statistics and figures were produced using Qualimap v.1.0 (García-Alcalde et al., 2012) with BAM files.

2.14.2 Mutation detection

Alignment files in BAM format were sorted using Samtools sort and indexed using Samtools index. The sorted and indexed alignment was used with Samtools mpileup for producing an mpileup file using the reference genomes in fasta format (option –f) stated in section 2.14.1.

Mutations were detected using VarScan v.2.3.3 (Koboldt et al., 2012) pileup2snp for SNPs and pileup2indel for Indels, with option --p-value 0.01. All mutations were visualised using Artemis (Rutherford et al. 2000) and some of them randomly selected for manual inspection.

2.14.3 Simulation of sequencing reads

ART Illumina v.2.1.8 (Huang et al., 2012b) was used for simulating Illumina sequencing reads for *Synechococcus* sp. WH7803. The options used to generate reads were –p (paired-end reads), -l s. The output in fastq format was aligned to *Synechococcus* sp. WH7803 with BWA mem (see section 2.14.1) and mutations detected with VarScan v.2.3.3 (section 2.14.2).

2.15. Quantitative PCR (qPCR) assay for determination of chromosome copy number in marine *Synechococcus*

The qPCR was adapted from Pecoraro et al. (2011). To optimise the assay, 10^0 - 10^7 cells of a late exponentially-growing axenic *Synechococcus* spp. culture were sorted in triplicate with a BD Influx cell sorter (by Dr. F. Pitt). Cells had been previously stained using SYBR Green I nucleic acid stain (Invitrogen) following the manufacturer's instructions in order to count all cells.

Subsequently, sorted cells were centrifuged at 16,060 *g* and the supernatant removed since the salt concentration in ASW inhibited PCR (see Fig. 5.2.B). The pellet was re-suspended in 1 mL nuclease free water and the cells disrupted with a TissueLyser (Qiagen) in 2 mL Lysing Matrix E tubes (MP Biomedicals). *Synechococcus* sp. WH7805 chromosome copy number was determined using both *Synechococcus* sp. WH7805 alone or mixed with an equimolar amount of *Synechococcus* sp. WH7803 to control the DNA extraction method described.

A standard curve was constructed using *Synechococcus* sp. WH7803 or WH7805 genomic DNA obtained as described previously (section 2.13) and quantified using the QuantiFluor dsDNA System (Promega).

The qPCR reaction was performed in triplicate, including a dilution series of *Synechococcus* cells of 10^0 - 10^{-4}, a standard curve constructed with genomic DNA and a negative control of nuclease free water. An aliquot of 1 µL of sample was used in a 20 µL final qPCR reaction volume following the manufacturer's instructions (SYBR Select Master Mix, Applied Biosystems), with a primer concentration of 100 nM (primers listed in Table 2.10), and run with the fast settings (20 s at 95ºC and 40 cycles of 3 s at 95ºC and 30 s at 60ºC) in a 7500 Fast Real-Time PCR thermocycler (Applied Biosystems).

All experiments used a standard curve of DNA and were performed in triplicate. Results are indicated as the average of 3 replicates ± standard deviation.

2.16. Assessment of chromosome copy number using a fluorometric based approach

Quantification of genomic DNA also used a fluorometric assay facilitated by the QuantiFluor dsDNA System (Promega) kit. Briefly, late-exponentially growing *Synechococcus* sp. WH7803 cells (OD_{750} = 0.45) were serially diluted (10^0 – 10^7) and counted by flow cytometry (section 2.3). Genomic DNA was extracted as described in section 2.9. Calculation of chromosome copy number from genomic DNA content per cell (Equation 1) was performed based on Doležel et al. (2003) Assuming that at pH 8 (ASW) phosphate protons of nucleotides are dissociated and that 1 pg of DNA equals

0.978×10^9 bp, the chromosome of *Synechococcus* sp. WH7803 (genome size of 2,366,980 bp) should weigh 2.42×10^{-3} pg.

$$\text{DNA content [pg]} = \frac{\text{Genome size [bp]}}{0.978 \times 10^9 \text{ [bp pg}^{-1}\text{]}}$$

Equation 1. DNA content calculated from genome size (Doležel et al., 2003).

2.17. Grazing experiments

Grazing experiments were carried out by measuring predator and prey cell concentrations using flow cytometry. An exponentially growing *Synechococcus* culture (4 mL) at a concentration of 10^7 cells mL^{-1}, was mixed with 4 mL of a 1 week old nanoflagellate culture at a concentration of 10^4 or 10^5 cells mL^{-1} and 1 mL of cyanophage at a concentration of 10^8 PFU mL^{-1}. 80 µL samples were fixed and stained for flow cytometry (see section 2.3). Results were analysed by plotting the number of cells mL^{-1} or percentage of change in cell concentration against time.

Chapter 3

MOLECULAR BASIS OF CYANOPHAGE RESISTANCE IN *SYNECHOCOCCUS* SP. WH7803

3.1. Introduction

Cyanophage populations and marine *Synechococcus* co-exist in the natural environment. Recently, Weitz et al. (2015) modelled a marine system with and without viruses and determined that the presence of viruses increases turnover which leads to a higher primary productivity. The exact percentage of *Synechococcus* mortality due to viral lysis is unknown, but different studies have calculated up to 50 % of cyanobacterial infection rates at any given point, with an estimated absolute mortality rate of 2.8 - 48 % for *Synechococcus* (Clokie et al., 2011; Mann, 2003; Murray and Eldridge, 1994; Proctor and Fuhrman, 1990; Waterbury and Valois, 1993) with only a minority of *Synechococcus* spp. sensitive to co-isolated lytic cyanophages. This indicates that despite the high diversity and titre of cyanophages found in marine systems (Lu et al., 2001; Marston and Sallee, 2003; Millard and Mann, 2006; Suttle and Chan, 1994; Zhong et al., 2002), viruses appear not to be alone in regulating *Synechococcus* populations. Such results are in contrast though to other studies that have shown low levels of cyanophage resistance amongst host populations, probably due to the fitness cost associated with phage resistance, which has been estimated to be ~20 % in isolated strains based on growth rates (Lennon et al., 2007; Stoddard et al., 2007; Suttle and Chan, 1994). Even so, overall, the importance of phage resistance in *Synechococcus* population dynamics is well recognised.

Attempts to understand mechanisms of cyanophage resistance have suggested that they are largely related to preventing cyanophage attachment to the host, possibly by modifying and/or blocking the cell surface and hence perhaps the cyanophage receptor (Avrani et al., 2011; Marston et al., 2012). Thus, studies performed in *Anabaena* sp. PCC7120, a freshwater cyanobacterium, showed that a modified lipopolysaccharide (LPS) layer prevents cyanophage infection (Xu et al., 1997). Two genes involved in LPS biosynthesis, were identified by random transposon (Tn5) mutagenesis and shown to be responsible for the resistant phenotype. These genes are *rfbP* (undecaprenyl-phosphate galactosephosphotransferase) and *rfbZ* (first mannosyl transferase) which insertional inactivation caused a modified O-polysaccharide profile.

Here at Warwick, recent work has focused towards determining the molecular basis of cyanophage resistance in marine cyanobacteria. Thus, both Jia (2009) and Spence

(2010) isolated various spontaneous cyanophage-resistant mutants (Table 2.3). These include the *Synechococcus* sp. WH7803 mutant PHR, resistant to cyanophage S-PM2. This mutant possesses a modified LPS profile compared to wild-type *Synechococcus* sp. WH7803, lacking an important portion of the O-polysaccharide chain. This feature may be the reason for the preferential grazing of this PHR strain by heterotrophic flagellates compared with the wild-type (Zwirglmaier et al., 2009), and suggesting that LPS plays an important role in mediating both cyanophage infection and grazing by protists. Moreover, the 13 cyanophage-resistant *Synechococcus* sp. WH7803 mutants used in this study (Table 2.3; Fig. 3.1) also show differences in pigmentation, that might be related to phycoerythrin content, and an aggregation or clumping phenotype, that might be related to a modified LPS layer, or to production of exopolysaccharide.

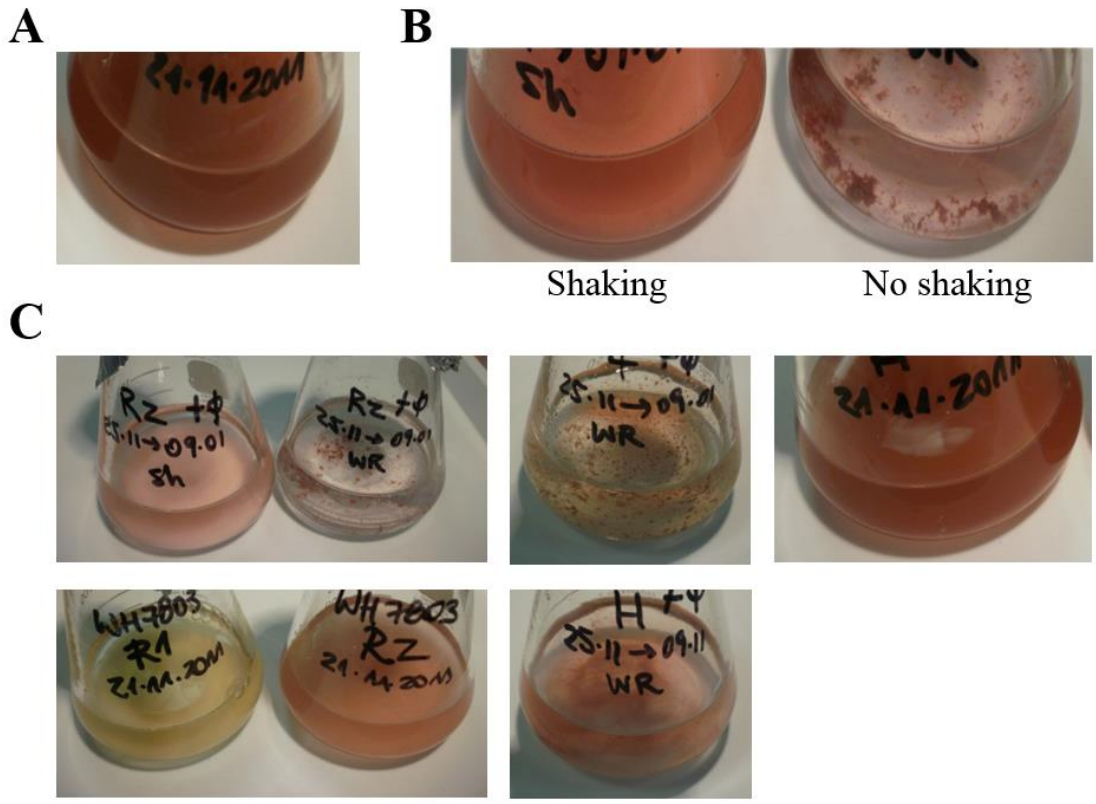


Figure 3.1. Clumping and pigment phenotypes of some of the cyanophage-resistant *Synechococcus* sp. WH7803 mutants isolated by Spence et al., (2010).

(A) Wild-type *Synechococcus* sp. WH7803. (B) Representative culture showing the clumped phenotype when cultured without shaking. (C) Representative cyanophage-resistant *Synechococcus* sp. WH7803 mutant cultures showing different phenotypes, including clumping behaviour and differences in pigmentation.

LPS is a conserved molecule present in the outer membrane of Gram negative bacteria, comprising a basic structure of i) lipid A, ii) a core oligosaccharide and iii) the O-polysaccharide chain or O-antigen (Fig 1.7). Only the O-polysaccharide chain varies, conferring a signature for different strains and species. The LPS structure of two marine *Synechococcus* strains (CC9311 and WH8102) was recently described (Fig 3.2; Snyder et al. 2009). Compared to LPS structures from enteric bacteria, marine *Synechococcus* possess a simpler molecule (Fig. 3.2) in which the main saccharide is 4-linked glucose, instead of heptose or 3-deoxy-D-manno-octulosonic acid (Kdo). In addition, differences in LPS biochemistry were seen between the two *Synechococcus* strains, with rhamnose present in the LPS of *Synechococcus* sp. WH8102 (consistent

with the presence of two gene clusters encoding a rhamnose biosynthetic pathway in its genome) but absent in *Synechococcus* sp. CC9311, although it is unclear whether this rhamnose is present in the core or O-polysaccharide chain.

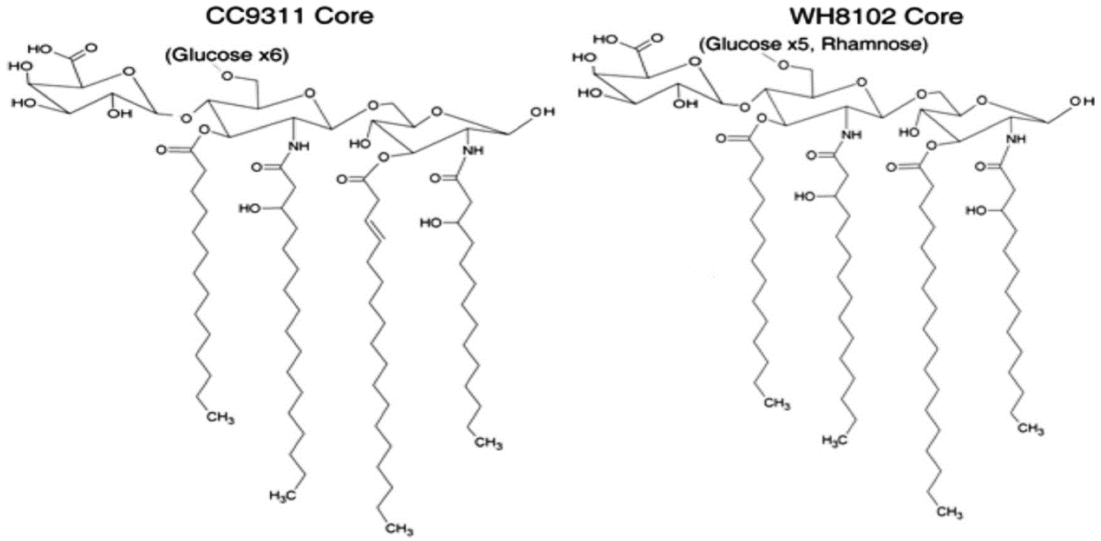


Figure 3.2. Putative structure of *Synechococcus* sp. CC9311 and WH8102 minimal LPS lipid A and core (from Snyder et al. 2009).

These structures represent one of a variety of the acylated lipid A structures found in these organisms.

To specifically investigate the role of the cell surface in cyanophage-resistance, Spence (2010) constructed three interposon mutants targeting different components of the membrane. These mutants comprised i) a deletion mutant of two adjacent genes encoding potential P-stress porins (Δ*synWH7803_2235-2236::aac(3)-IV*), thought to be the most abundant proteins in the *Synechococcus* sp. WH7803 cell surface ii) and iii) involved in LPS biosynthesis: ii) disrupting the *rmlB* homologue (Δ*synWH7803_0192::aac(3)-IV*), that encodes the second enzyme in the rhamnose biosynthetic pathway involved in the synthesis of the O-polysaccharide, and iii) disrupting the *wbaP* homologue (Δ*synWH7803_1767::aac(3)-IV*), encoding an hexose-1-P transferase that catalyses the initial glycosylation of the undecaprenyl phosphate lipid carrier, contributing to maintaining the length of the O-polysaccharide

(Saldías et al., 2008; Whitfield and Larue, 2008). However, none of these mutants have been biochemically characterised, nor assessed to determine if they confer cyanophage resistance.

During the course of my PhD studies two manuscripts were published (Avrani et al. 2011; Marston et al. 2012) that also focused on cyanophage resistance mechanisms in marine picocyanobacteria. These studies provided evidence to link resistance with an inability of the cyanophage to adsorb to the host cell surface. Thus, Avrani et al., (2011) investigated the mechanisms of cyanophage resistance in several *Prochlorococcus* strains, using ten different cyanophage, and isolated 77 resistant sub-strains. Using whole genome sequencing, mutations were found in genes located largely within genomic islands, including genes associated with membrane components or cell-wall biosynthesis. Avrani et al (2011) found that all their cyanophage resistant mutants prevented cyanophage attachment, probably by modification of the cell surface, as demonstrated by adsorption assays. Unfortunately, since no reliable genetic system has been developed for *Prochlorococcus* it is not yet possible to unequivocally prove that the mutations found in genes identified by whole genome sequencing actually confer resistance to cyanophage infection.

Similarly, Marston et al. (2012) investigating the evolution of *Synechococcus* sp. WH7803 grown in chemostat culture with cyanophage RIM8, showed the evolution of resistant populations, and by performing whole genome sequencing identified four mutations, two of them in a genomic island (ISL1). The mutations were identified as a SNP in *SynWH7803*_0102 encoding a putative glucose-1-phosphate thymidylyltransferase (*rmlA*/*rfbA*), a SNP in a glycosyltransferase (*SynWH7803_0140*), a SNP in a two-component system sensor histidine kinase (*baeS*, *SynWH7803_1386*) and a deletion that generates an early stop codon in an aminopeptidase N (PepN, *SynWH7803_1555*). However, none of these mutations have yet been subsequently proven by molecular genetic techniques, i.e. using gene knock-out technology, to actually confer cyanophage resistance.

3.2. Objectives

Here then, the role that the *Synechococcus* sp. WH7803 cell surface plays in mediating cyanophage resistance was investigated in all of the previously isolated spontaneous *Synechococcus* sp. WH7803 mutants isolated by Jia (2009) and Spence (2010), as well as those interposon mutants in which specific genes were targeted (see above), particularly assessing any differences in the LPS profiles of these strains.

3.3. Results

3.3.1 Phenotypic characterisation of spontaneous cyanophage-resistant *Synechococcus* sp. WH7803 mutants

To study the phenotype of the S-RSM42-resistant *Synechococcus* sp. WH7803 mutants A-J, and mutants R1 and R2 resistant to both S-RSM42 and S-PM2, LPS profiles and adsorption assays were carried out.

LPS profiles from cyanophage-resistant *Synechococcus* sp. WH7803 mutants were initially obtained by performing LPS "micro-extractions" (see section 2.9.1), comprising whole cell digestion with proteinase K. The LPS banding patterns observed (Fig. 3.3) correspond to three regions of the LPS basic structure: the lipid A core and O-polysaccharide chain (Region 1), LPS biosynthetic intermediates with incomplete core or O-polysaccharide chains (Region 2) and free lipid A with attached core polysaccharide intermediates (Region 3).

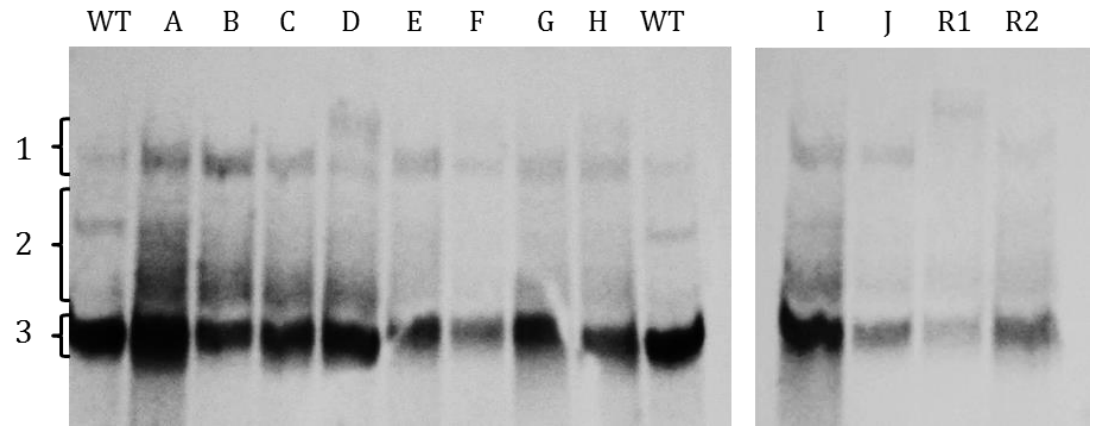


Figure 3.3. LPS profiles of cyanophage-resistant mutants obtained by LPS micro-extraction.

WT: *Synechococcus* sp. WH7803; A-J: *Synechococcus* sp. WH7803 mutants resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 mutants resistant to cyanophage S-RSM42 and S-PM2. 1: Mature LPS, 2: LPS biosynthetic intermediates, 3: Free lipid A.

However, using this type of extraction it was not possible to distinguish between O-polysaccharide profiles. This was important since Zwirglmaier et al. (2009) showed that the cyanophage resistant *Synechococcus* sp. WH7803 mutant PHR had lost a significant portion of its O-polysaccharide side chain units. The Zwirglmaier et al. (2009) study used the hot phenol-water method (see Westphal and Jann 1965) to specifically extract polysaccharides. However, this method requires a large volume of culture (1 L) to obtain sufficient LPS material, a facet not conducive with the S-RSM42-resistant *Synechococcus* sp. WH7803 mutants since these cultures only grow stably to a maximum volume of 40 mL. As a result a commercial LPS extraction kit (Intron Biotechnology) was used (see section 2.9.2).

The banding pattern obtained (Fig. 3.4) represents the number of O-polysaccharide substitutions in the LPS molecule of each strain. Mutants B, F, J and R1 show a similar pattern in comparison to wild type *Synechococcus* sp. WH7803, whereas mutants A, C, D, E, G, H and R2 possess more O-polysaccharide substitutions. The cyanophage S-PM2 resistant *Synechococcus* sp. WH7803 mutant PHR, was used as a negative control since as mentioned above this essentially lacks O-polysaccharide side chains.

The LPS profile of mutant I could not be determined due to its poor growth, with insufficient material available to perform LPS extraction.

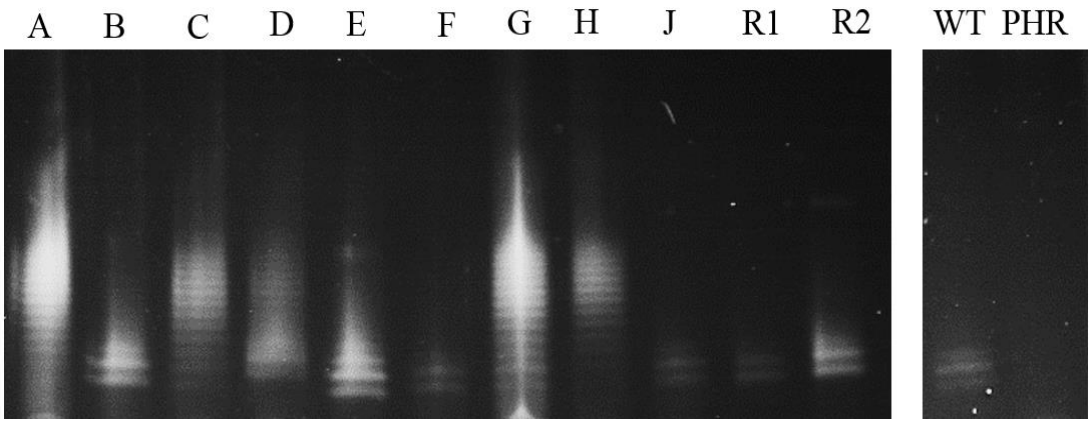


Figure 3.4. LPS profiles of cyanophage-resistant *Synechococcus* sp. WH7803 mutants obtained using a commercial LPS extraction kit (Intron Biotechnology).

A-J: S-RSM42-resistant *Synechococcus* sp. WH7803 mutants; R1 and R2: *Synechococcus* sp. WH7803 mutants resistant to both S-RSM42 and S-PM2; WT: *Synechococcus* sp. WH7803; PHR: S-PM2-resistant *Synechococcus* sp. WH7803 mutant, LPS 'negative' control.

To prevent phage infection, potentially any step of the infection cycle (see Fig 1.5) can be interrupted. To directly assess whether attachment of cyanophages to the host cell, the first step in the infection process, was interrupted, thus conferring resistance to *Synechococcus* sp. WH7803 mutants, adsorption assays were performed (see section 2.8 and Fig. 3.5). Cyanophage-resistant *Synechococcus* sp. WH7803 mutants C and R1 proved to be sensitive to adsorption by cyanophage S-RSM42 (Fig. 3.5.A) showing a 94 % reduction (mutant C) or 96 % reduction (mutant R1) in cyanophage titre after 60 minutes, respectively. Similarly, mutant R2 was sensitive to cyanophage S-PM2 (Fig. 3.5.B), showing a 96 % reduction in phage titre after 60 minutes. In contrast, the remaining mutants were all resistant to phage attachment (Fig. 3.5A, B) showing no difference in phage titre following the 60 minute incubation period. It is worth noting that mutants R1 and R2 (which are resistant to both cyanophages S-PM2 and S-RSM42) show inverse adsorption profiles, i.e. mutant R1 adsorbs cyanophage

S-PM2 but not cyanophage S-RSM42, whilst mutant R2 shows the opposite profile, suggesting these mutants possess different phage resistance mechanisms.

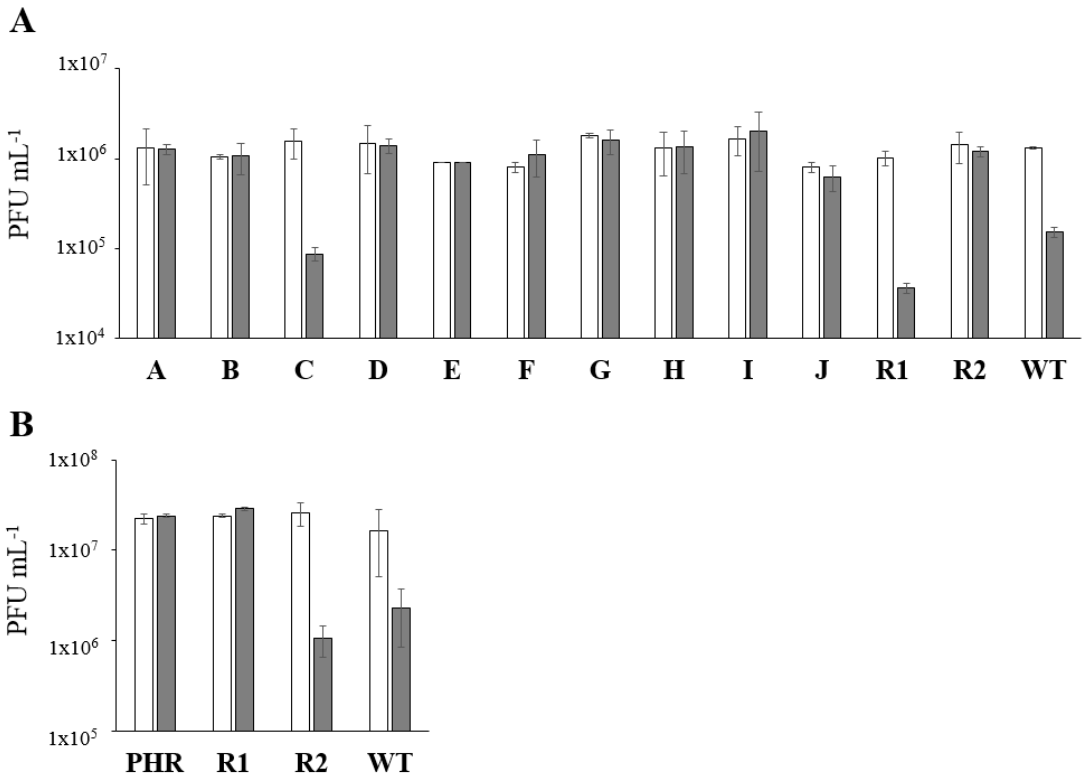


Figure 3.5. Adsorption assays of cyanophage-resistant *Synechococcus* sp. WH7803 mutants

(A) Mutants resistant to cyanophage S-RSM42 and (B) mutants resistant to S-PM2. White and grey bars represent phage titres after 0 minutes or 60 minutes absorption with each *Synechococcus* strain, respectively. A-J: S-RSM42-resistant *Synechococcus* sp. WH7803 mutants; R1 and R2: *Synechococcus* sp. WH7803 mutants resistant to S-RSM42 and S-PM2; PHR: S-PM2-resistant *Synechococcus* sp. WH7803 mutant; WT: *Synechococcus* sp. WH7803, positive control for adsorption assays. Bars are the average of three replicates with error bars as standard deviation.

3.3.2 Assessment of cyanophage resistance in partially segregated *Synechococcus* sp. WH7803 interposon mutants in the *synWH7803_2235-2236, synWH7803_0192* and *synWH7803_1767* genes

In order to study the relationship between cell surface properties of *Synechococcus* sp. WH7803 and cyanophage resistance, three partially segregated interposon mutants constructed by Spence (2010) were investigated (Table 2.2.4).

Cyanophage resistance in these *synWH7803_2235-2236, synWH7803_0192* and *synWH7803_1767* gene interposon mutants was evaluated using a qualitative spot assay (see section 2.6). However, all mutants were susceptible to cyanophages S-RSM42, S-PM2 and S-RSM4 (Figure 3.6A) in a similar manner to wild type. LPS profiles (using a commercial extraction kit, see section 2.9.2) of *synWH7803_0192* and *synWH7803_1767* interposon mutants also showed no differences to the wild type (Fig. 3.6B). The wild type LPS profile of the *synWH7803_2235-2236* interposon mutant is consistent with the function of these genes as outer membrane porins, and hence unrelated to LPS biosynthesis.

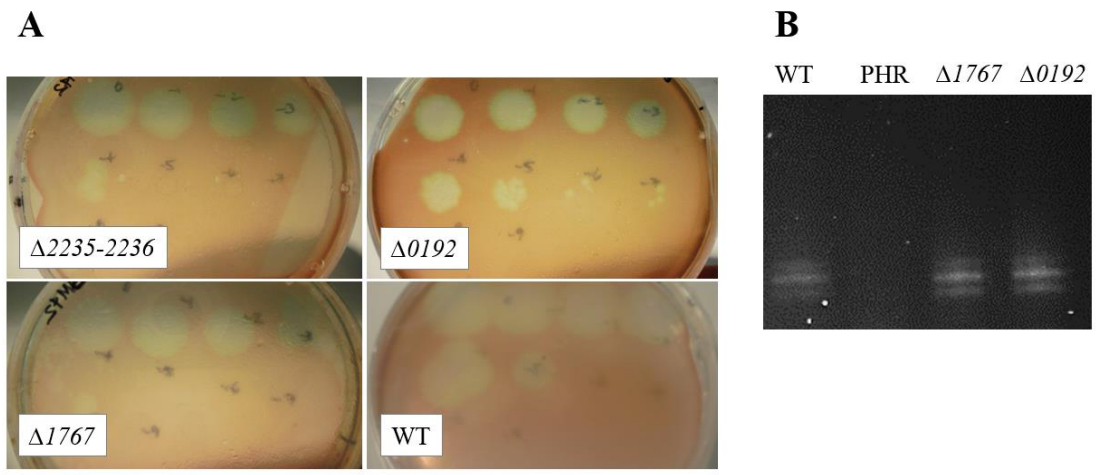
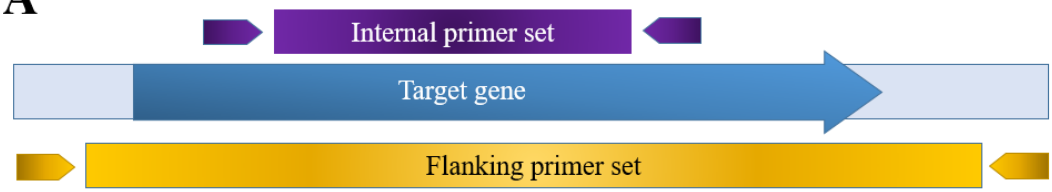


Figure 3.6. Assessment of cyanophage resistance (A) and LPS profiles (B) in interposon mutants of *Synechococcus* sp. WH7803.

(A) Representative images of the qualitative spot assay for determining cyanophage resistance and (B) LPS profiles of two *Synechococcus* interposon mutants. Spot assays were performed in triplicate using *Synechococcus* sp. WH7803 wild type as a control and cyanophages S-RSM42, S-PM2 and S-RSM4. For each mutant, sensitivity to all three cyanophage was observed. Δ*2235-2236*: Δ*synWH7803_2235-2236::aac(3)-IV*; Δ*0192*: Δ*synWH7803_0192::aac(3)-IV*; Δ*1767*: Δ*synWH7803_1767::aac(3)-IV*; WT: *Synechococcus* sp. WH7803 wild type; PHR: S-PM2-resistant *Synechococcus* sp. WH7803 mutant (LPS negative control).

Since *Synechococcus* sp. WH7803 is an oligoploid strain, i.e. it possess multiple copies of the chromosome (see chapter 5), it was thought possible that incomplete segregation of these interposon mutants had occurred, which might explain the lack of phenotype observed above. In order to try to force segregation, each mutant was subjected to a gradual increase in antibiotic concentration. PCR assays using specific primers (see Table 2.8) flanking and internal to the disrupted gene (Fig. 3.7.A and B) were used to detect the wild type gene and directly assess if complete segregation had occurred. However, even in those mutants cultured in apramycin at 1 mg mL^{-1} (the 'normal' working concentration is 50 µg mL^{-1}), which was clearly lethal to wild type *Synechococcus* sp. WH7803 (Fig. 3.7.D), did not result in full segregation and the wild type gene was detected in all three interposon mutants (Fig. 3.7.C).

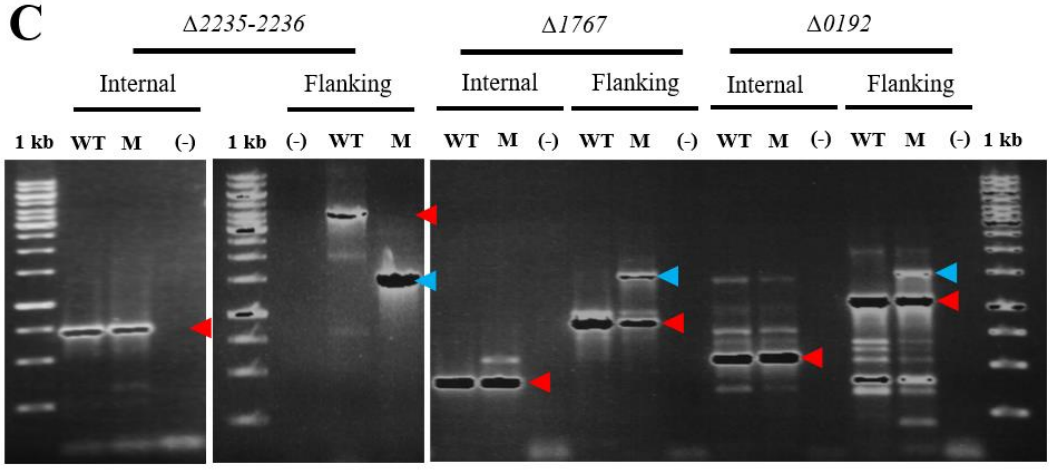
A



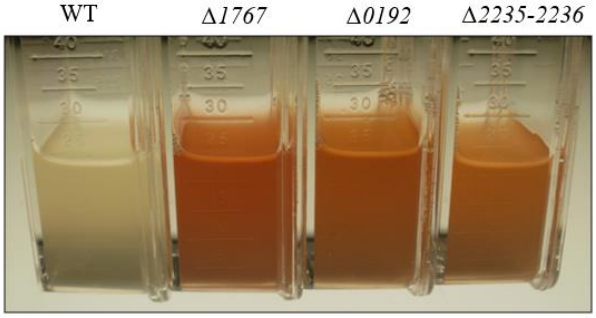
B

Interposon mutant	PCR fragment	Expected PCR product [bp] ▶ Wild type	▶ Mutant
Δ2235-2236	Internal	696	-
	Flanking	3959	1454
Δ1767	Internal	390	-
	Flanking	830	1423
Δ0192	Internal	527	-
	Flanking	1064	1451

C



D



(Previous page)

Figure 3.7. PCR assessment of segregation and growth in apramycin in the three *Synechococcus* sp. WH7803 interposon mutants, i.e. in genes *synWH7803_2235-2236*, *synWH7803_0192* and *synWH7803_1767*.

(A) Schematic representation of the PCR primer sets internal and flanking the disrupted gene designed to detect segregation of the interposon mutation. (B) Expected PCR fragment sizes expected in *Synechococcus* sp. WH7803 (wild type, red arrow) and Interposon mutant (Mutant, blue arrow) amplified with the PCR primer in (A) (see Table 2.8). (C) PCR assessment of segregation in each interposon mutant using genomic DNA extracted from the interposon mutants cultures in (D): WT: *Synechococcus* sp. WH7803; Δ*1767*: Δ*synWH7803_1767::aac(3)-IV;* Δ*0192*: Δ*synWH7803_0192::aac(3)-IV;* Δ*2235-2236*: Δ*synWH7803_2235-2236::aac(3)-IV*. In each case PCR primers are internal and flanking the disrupted gene. Red arrows: wild type gene; blue arrows: apramycin-resistance cassette; WT: *Synechococcus* sp. WH7803 genomic DNA, M: interposon mutant genomic DNA, (-) PCR negative control, 1 kb ladder (Invitrogen). (D) Interposon mutants after three weeks growth in ASW medium supplemented with 1 mg mL^{-1} apramycin.

A further approach to try to segregate these interposon mutants was by growth in P-deplete ASW medium (section 2.1.1), whilst also maintaining the 1 mg mL^{-1} apramycin selection, with the idea that P stress might force a reduction in chromosome copy number due to the large P requirement of DNA. However, even after 3 months of culturing in P-deplete medium, with cultures being transferred every ~3 weeks, PCR assays with the same internal and flanking primers used above (and see Table 2.8), showed no evidence of complete segregation (Fig. 3.8).

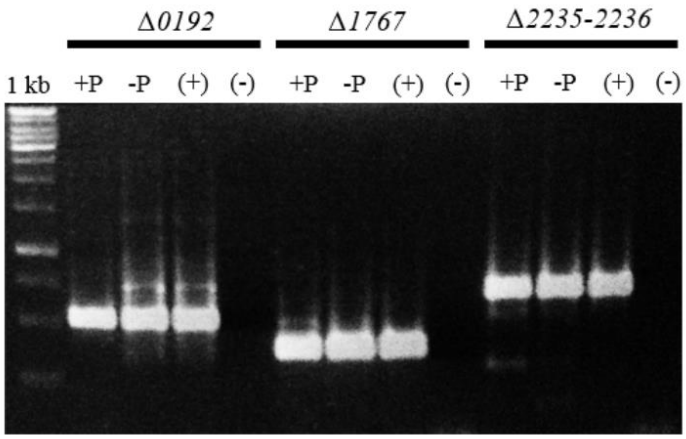


Figure 3.8. PCR assessment of segregation in the three *Synechococcus* sp. WH7803 interposon mutants following culturing in P-deplete ASW medium for three months.

This PCR was performed using the internal primer set (Table 2.8).Δ*0192*: Δ*synWH7803_0192::aac(3)-IV*; Δ*1767*: Δ*synWH7803_1767::aac(3)-IV;* Δ*2235-2236*: Δ*synWH7803_2235-2236::aac(3)-IV*; (+) *Synechococcus* sp. WH7803 genomic DNA, (-) PCR negative control, 1 kb ladder (Invitrogen).

3.3.3 Targeted mutagenesis of genes identified by whole genome sequencing to be responsible for cyanophage resistance in *Synechococcus* sp. WH7803

Marston et al. (2012) identified specific genes thought to be responsible for cyanophage resistance in *Synechococcus* sp. WH7803 through whole genome sequencing of strains resistant to virus RIM-8. Two of these genes were selected for directed interposon mutagenesis, both genes being present in single copy in the genome and with each gene possessing a non-synonymous mutation in the original cyanophage-resistant mutant strain. These genes were *SynWH7803_0102* (*rmlA*/*rfbA*) encoding the first enzyme in the rhamnose biosynthetic pathway, a glucose-1-phosphate thymidyltransferase, and *SynWH7803_1386* (*baeS*), a histidine kinase from a two-component system possessing a PAS domain, known to be a signalling component that can detect changes in different intracellular energy conditions, such as small ligands (Taylor and Zhulin, 1999). The mutations observed by Marston et al. (2012) are in regions of the genes encoding specific protein domains,

in particular the transferase domain for SynWH7803_0102 and in the PAS domain for SynWH7803_1386, predicted to be a signal sensor domain (Fig. 3.9) (see Henry & Crosson 2011 for a recent review).

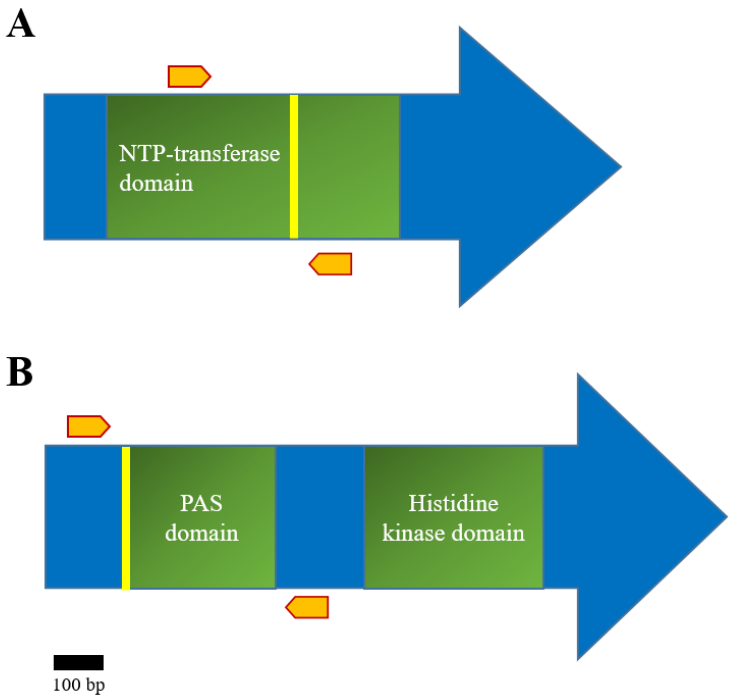


Figure 3.9. Schematic diagram of *Synechococcus* sp. WH7803 genes targeted for interposon mutagenesis.

(A) SynWH7803_0102 (*rmlA*/*rfbA*), 1156 bp; (B) SynWH7803_1386 (*baeS*), 1366 bp. Blue arrows represent the genes; green blocks indicate those regions of the protein encoding specific domains; yellow vertical lines indicate the position of mutations found by Marston et al. (2012) and orange arrows indicate the position of PCR primers used to make the constructs.

The interposon mutant constructs were made by cloning a 395 bp region of *SynW7803_0102* and a 347 bp region of *SynWH7803_1836* (PCR primers are listed in Table 2.8) into the *Eco*RI restriction site of pMUT100 (Fig. 2.2). Correct clone construction was verified by PCR and sequencing of the insert fragment (data not shown). Interposon mutant plasmid constructs were subsequently transformed into *E. coli* (contained the helper plasmids pRK24, pRL528). Conjugation of the mutant

constructs into *Synechococcus* sp. WH7803 was carried out using both SN and ASW semi-solid medium and different ratios of the *E. coli* donor: *Synechococcus* receiver strains (see section 2.12). However, whilst kanamycin resistant transconjugants were obtained, PCR did not confirm recombination of the kanamycin cassette into the correct position in the genome in either mutant construct (data not shown).

3.4. Discussion

Thus far, cyanophage resistance in marine cyanobacteria has been found to be exclusively due to modification of the host cell surface and hence possible inhibition of cyanophage attachment (Avrani et al. 2011; Marston et al. 2012). In this study, three cyanophage resistant *Synechococcus* mutants were found to be able to adsorb the cyanophage (Fig. 3.5) whilst still resistant to infection, suggesting they somehow prevent cyanophage replication and, at least, demonstrating that more mechanisms for resistance occur in these cyanophage-host systems than previously thought. Indeed, adsorption patterns of cyanophages S-PM2 and S-RSM42 to *Synechococcus* mutants R1 and R2 (Fig. 3.5), which are resistant to both cyanophages, are clearly different, suggesting that *Synechococcus* sp. WH7803 possesses several mechanisms for cyanophage resistance and that the host cell receptor for these cyanophages is distinct.

Bacterial resistance to viral infection can be acquired by interrupting any of the steps of the infection cycle (see Section 1.4), although in cyanobacteria only the adsorption of the cyanophage has been found to be interrupted in resistant mutants and suggested to be related to modifications in the LPS O-polysaccharide. In *Pseudomonas fluorescence* SBW25 co-evolution experiments with and without bacteriophages for ~400 generations also led to changes in LPS as a result of host-bacteriophage co-evolution, which was linked with a reduce fitness in the mutant (co-evolved) host strains (Scanlan et al., 2015). In *E. coli* modification of LPS has also been found to be implicated in bacteriophage resistance. Furthermore, mutations in the biosynthesis of this molecule are related to cross-resistance to bacteriophage T7 and T4. However, the cost of resistance to bacteriophage T7 is higher because its binding site is deeper in the LPS molecule (core oligosaccharide rather than O-polysaccharide) compared to T4 (Bohannan and Lenski, 2000).

In this chapter, analysis of the LPS profiles of the spontaneous cyanophage-resistant *Synechococcus* sp. WH7803 mutants (Fig. 3.4) showed that they do not correlate with the adsorption phenotypes (Fig. 3.5). For instance, cyanophage S-PM2 is able to attach only to mutant R2, whilst cyanophage S-RSM42 can attach only to mutants C and R1. However, mutants B, F, J and R1 have the same LPS profile as the wild type (Fig. 3.4), whereas mutants C and R2 possess an altered LPS compared to wild type, but this altered LPS profile does not inhibit cyanophage attachment. Moreover, the LPS profiles of mutants A, C, D, G. H and R2 show a pattern indicating the presence of larger O-polysaccharide molecules compared to the wild type. This evidence is inconclusive regarding the role of LPS in cyanophage resistance.

The adsorption assays and LPS profiles carried out with the spontaneous cyanophage resistant mutants, indicates that either LPS is not the receptor for cyanophages S-PM2 and S-RSM42 or that cyanophage attachment also need the presence of other molecules/proteins. For example outer membrane porin proteins, have previously been shown to be important for attachment in other phage-host systems in addition to LPS, with both components required for efficient attachment (Rakhuba et al., 2010; Silverman and Benson, 1987; Traurig and Misra, 1999). Certainly it is well known in enteric bacteria e.g. *E. coli*, that the outer membrane porin proteins OmpA, OmpC, OmpF and LamB, (the latter a maltodextrin channel), act as bacteriophage receptors in their own right (Clement et al., 1983; Mizoguchi et al., 2003; Montag et al., 1987; Morita et al., 2002; Riede and Eschbach, 1986; Riede et al., 1985; Silverman and Benson, 1987; Traurig and Misra, 1999).

It is worth noting that heterotrophic bacterial contamination is present in all the cyanophage resistant *Synechococcus* sp. WH7803 mutant cultures which could interfere with some of the observed alterations in LPS profiles (Fig. 3.5). However, this is unlikely given the relatively low levels of contamination generally detected in these cultures (< 30 %; see Table 4.1).

The role of LPS in cyanophage adhesion has thus far only been inferred from whole genome sequencing and adsorption data (Avrani et al. 2011; Marston et al. 2012) or using random transposon mutagenesis, in the latter where interruption of the *rfbP* (undecaprenyl-phosphate galactosephosphotransferase) and *rfbZ* (first mannosyl

transferase) genes occurred, these both being involved in LPS biosynthesis (Xu et al., 1997). To directly determine the role of LPS in cyanophage resistance it is necessary to understand the molecular mechanisms involved. With this aim, a number of genes directly involved in LPS biosynthesis, as well as two genes encoding predicted outer membrane porins, were specifically targeted for mutagenesis. However, none of these mutants completely segregated, potentially explaining the lack of a phenotype i.e. mutants were still susceptible to cyanophage infection, and no evidence of LPS modification was observed (Fig. 3.6).

To more directly assess the molecular basis of resistance in the spontaneous cyanophage resistant *Synechococcus* mutants obtained by Spence (2010) whole genome sequencing was undertaken (Chapter 4).

Chapter 4

WHOLE GENOME SEQUENCING OF *SYNECHOCOCCUS* SP. WH7803 CYANOPHAGE RESISTANT MUTANTS

4.1. Introduction

The molecular mechanisms underlying bacteriophage-bacteria interactions and co-evolution remain largely unknown. By far the best understood systems of bacteriophage-bacteria co-evolution are those performed in heterotrophic γ-proteobacteria including *E. coli* and *P. fluorescens* (Mizoguchi et al., 2003; Paterson et al., 2010; Scanlan et al., 2015), but for cyanobacteria and their co-occurring cyanophages such mechanisms are still poorly understood.

In chapter 3 phenotypic variation amongst several cyanophage resistant *Synechococcus* sp. WH7803 mutants was presented. These mutants possessed differences in their LPS profiles and also in their adsorption to cyanophages, suggesting different mechanisms for cyanophage resistance may exist.

Recent work in *Prochlorococcus* and marine *Synechococcus* (Avrani and Lindell, 2015; Avrani et al., 2011; Lennon et al., 2007; Marston et al., 2012) suggests that mutations in genes involved in the biogenesis of LPS and other membrane components confers resistance to cyanophages, but these authors found only cyanophage resistant mutants that were adsorption deficient, i.e. incapable of cyanophage attachment, which is not the case here (see Section 3.3.1).

To identify the molecular basis of cyanophage–resistance in the cyanophage resistant *Synechococcus* sp. WH7803 mutants characterised here (chapter 3) a whole genome sequencing (WGS) approach was taken, with the idea to subsequently undertake Sanger sequencing to independently confirm these mutations and relate these back to the observed phenotype. WGS was performed on all 13 of the cyanophage resistant mutants described in the previous chapter (chapter 3), as well as wild type *Synechococcus* sp. WH7803 cultures from October 2013 and December 2014, approximately 14 months apart. The latter provided an independent 'sequencing pipeline comparison' with the cyanophage-resistant mutants obtained for *Synechococcus* sp. WH7805 (see chapter 6).

At the time of sequencing, different next generation sequencing technologies were available. The work published on *Prochlorococcus* was performed using Illumina

HiSeq (Avrani et al., 2011; Avrani & Lindell 2015) whilst the *Synechococcus* sp. WH7803 study (Marston et al. 2012) used 454FLX technology. In this work, the platform chosen for WGS was Illumina MiSeq with Truseq library preparation, because it allowed high sequence yield at a relatively low cost per Gb of data obtained (Quail et al., 2012; see section 2.13).

4.2. Objective

To identify the genetic differences that confer cyanophage resistance in *Synechococcus* sp. WH7803 using WGS of the 13 cyanophage resistant mutants and wild type *Synechococcus* sp. WH7803, and comparing these with the previously published genome sequence (accession number NC_009481.1) for *Synechococcus* sp. WH7803 (Dufresne et al., 2008), which was obtained via traditional clone library construction and Sanger sequencing technology and hence provides a single 'consensus' sequence.

4.3. Results

4.3.1 Reducing contamination by heterotrophic bacteria of the cyanophage resistant *Synechococcus* sp. WH7803 mutants

Since the cyanophage resistant *Synechococcus* sp. WH7803 mutants are not axenic, attempts were made to maximise the amount of genomic DNA extracted from *Synechococcus* rather than contaminating heterotrophic bacteria. Thus, contamination was determined at two different time points during growth, early and late-exponential, with the purpose of minimising the relative contribution of the heterotrophic bacteria to the *Synechococcus* culture.

Contamination was assessed by analysing cultures using flow cytometry (see section 2.3) and the relative composition of the samples, i.e. percentage contribution of heterotrophs and *Synechococcus*, calculated as the percentage of total bacteria. Analysis of cultures one week (early exponential phase) and one month (stationary phase) after transfer (Fig. 4.1 A and B, respectively) showed a much lower level of

heterotrophic contamination during early exponential phase, suggesting that genomic DNA extracted from these rapidly growing cultures would be sufficient for WGS.

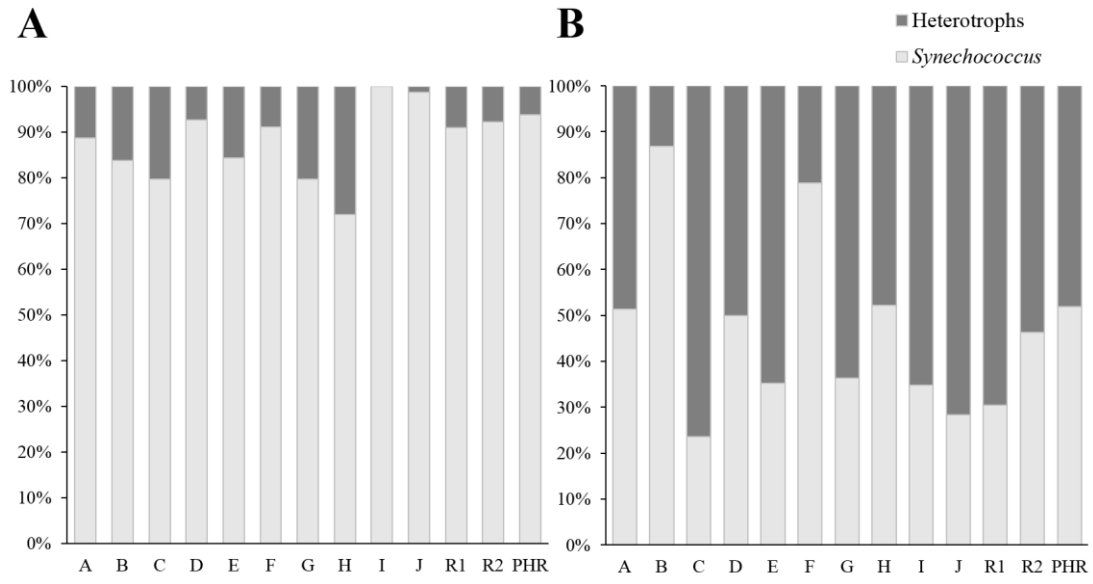


Figure 4.1. Percentage variation in the proportion of heterotrophic bacteria relative to *Synechococcus* sp. WH7803 cyanophage-resistant mutants in each culture at early exponential (A) or stationary (B) phase.

Cyanophage resistant *Synechococcus* sp. WH7803 mutant cultures screened by flow cytometry after one week (A) and one month (B) of inoculation. A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2. The data was obtained by flow cytometry of cells stained with SYBR Green I (Invitrogen) detected by green (DNA) and orange (phycoerythrin) fluorescence.

Subsequently, genomic DNA for WGS was extracted from each of the 13 cyanophage resistant *Synechococcus* sp. WH7803 mutants (see section 2.13), and where contamination with heterotrophic bacteria was not > 30%, except for mutant R2 where the contamination was 32%. DNA was also extracted from the axenic wild type strain (Table 4.1).

Strain	Contaminating heterotrophic bacteria [%]
A	16
B	27
C	26
D	25
E	15
F	N/D
G	21
H	23
I	12
J	16
R1	12
R2	32
PHR	6
WT	0

Table 4.1. Percentage contamination of heterotrophic bacteria in each *Synechococcus* cultures used for WGS.

Genomic DNA from these cultures was extracted for Illumina WGS. A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2; WT: wild type *Synechococcus* sp. WH7803. N/D: not determined. The data was obtained by flow cytometry of cells stained with SYBR Green I (Invitrogen) detected by green (DNA) and orange (phycoerythrin) fluorescence.

4.3.2 Whole genome sequencing of wild type *Synechococcus* sp. WH7803 and cyanophage resistant mutants

The integrity of genomic DNA from wild type and mutant *Synechococcus* sp. WH7803 isolates was assessed by agarose gel electrophoresis (Fig. 4.2). DNA (1 µg) was subsequently sent for sequencing at the Centre for Genomic Research, University of Liverpool using the Illumina platform (see Section 2.13 and Table 2.11). Note that all but one of the wild type and mutant *Synechococcus* sp. WH7803 strains were sequenced using 250 bp paired-end libraries. The exception was wild type *Synechococcus* sp. WH7803 extracted during December 2014 and sent for sequencing at the same time as DNA extracted from wild type and mutant *Synechococcus* sp. WH7805 (see chapter 6). All these latter strains used the Illumina sequencing platform and 150 bp paired-end libraries (see chapter 6 and Table 2.11).

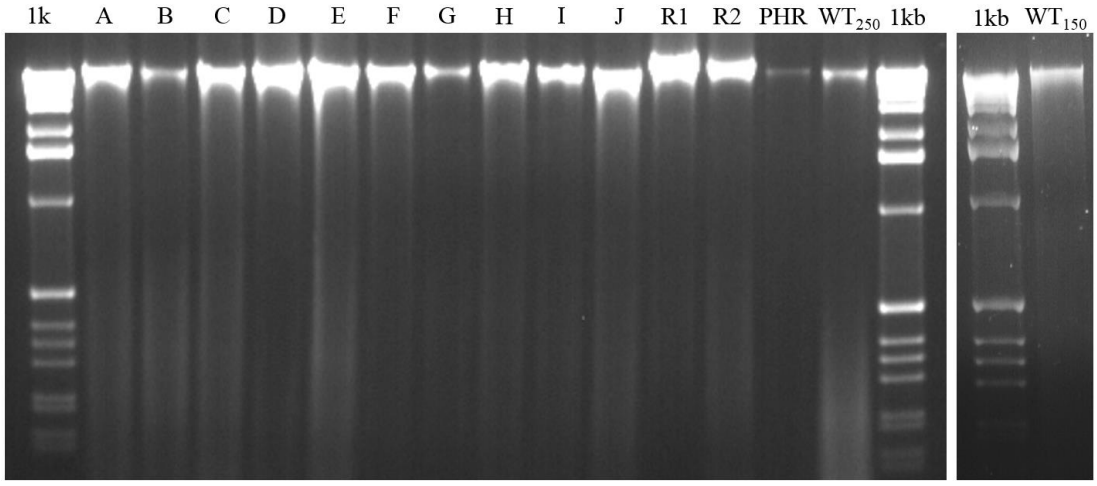


Figure 4.2. Integrity of *Synechococcus* sp. WH7803 genomic DNA visualised by agarose gel electrophoresis following ethidium bromide staining.

A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2; WT_{250}: wild type *Synechococcus* sp. WH7803 250 bp paired-end library; WT_{150}: wild type *Synechococcus* sp. WH7803 150 bp paired-end library. 1 kb: 1 kb ladder (Invitrogen).

Sequence reads were obtained in fastq format and the quality was controlled using FastQC v.0.10.1 (see section 2.14). After trimming, the average number of reads for all samples was $1.7 \times 10^6 \pm 4.1 \times 10^5$ (Fig. 4.3).

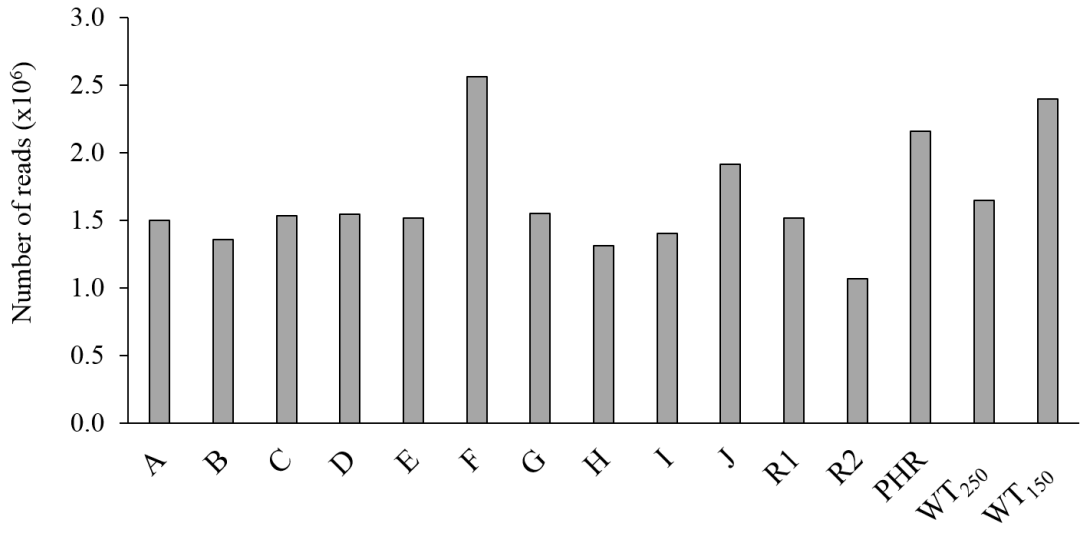


Figure 4.3. Total number of reads obtained from WGS of each *Synechococcus* sp. WH7803 strain.

A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2; WT_{250}: wild type *Synechococcus* sp. WH7803 250 bp paired-end library;, WT_{150}: wild type *Synechococcus* sp. WH7803 150 bp paired-end library.

4.3.3 Alignment of reads against the published *Synechococcus* sp. WH7803 genome sequence

Alignment of the wild type *Synechococcus* sp. WH7803 250 bp paired end library was performed using three different genome alignment algorithms available at the time of analysis: Bowtie2, BWA mem and NextGenMap (see section 2.14.1) and mapping to the published sequence of *Synechococcus* sp. WH7803 (accession number NC_009481.1). Based on this comparison (Table 4.2), BWA mem was chosen for

further alignments since this algorithm produced a higher mean mapping quality and mapped number of reads.

	NextGenMap	BWA mem	Bowtie2
Number of mapped reads	1,630,282	1,631,902	1,628,990
Mapped reads (%)	99.1	99.2	99.0
Coverage mean	151.67	151.76	151.06
Coverage SD	26.98	26.98	26.90
Mean mapping quality	54.43	59.54	40.97
General error rate (%)	0.09	0.09	0.51
Homopolymer Indels (%)	44.82	60.49	68.47

Table 4.2. Comparison of three different alignment algorithms with *Synechococcus* sp. WH7803 wild type (250 bp paired end library) sequence reads.

The data was obtained using Qualimap from the Bam files of the alignment of wild type *Synechococcus* sp. WH7803 250 bp paired-end reads to the published sequence (see Section 2.14 for details). The mapped reads are the percentage of mapped reads as a proportion of the total number of reads obtained for each sample [%].

In other words BWA mem (and NextGenMap in a similar way) was better at reducing errors in the alignment, without affecting coverage, by reducing the number of misaligned reads and hence the number of variants called, making it easier to handle the subsequent output (Table 4.3).

	Bowtie2		BWA mem	
Strain	Mapped reads (%)	Total number of mutations called	Mapped reads (%)	Total number of mutations called
A	10.5	946	10.5	317
B	34.8	1,028	34.8	130
C	13.3	924	13.3	146
D	10.3	1,022	10.3	234
E	12.7	828	12.7	161
F	2.3	1,342	2.3	423
G	11.1	1,031	11.1	200
H	9.5	737	9.5	134
I	10.3	655	10.3	140
J	11.1	923	11.1	219
PHR	37.5	1,303	37.5	273
R1	15.2	675	15.2	152
R2	10.9	1,031	10.9	196
WT_{150}	98.3	2,333	98.7	2,149
WT_{250}	99.0	829	99.2	74

Table 4.3. Comparison of Bowtie2 and BWA mem alignment algorithms using the wild type and mutant *Synechococcus* sp. WH7803 sequence reads.

The mapped reads are the percentage of mapped reads as a proportion of the total number of reads obtained for each sample [%].

The sequence reads obtained in fastq format from WGS of all the cyanophage resistant *Synechococcus* sp. WH7803 mutants and the two wild type libraries (WT_{250} and

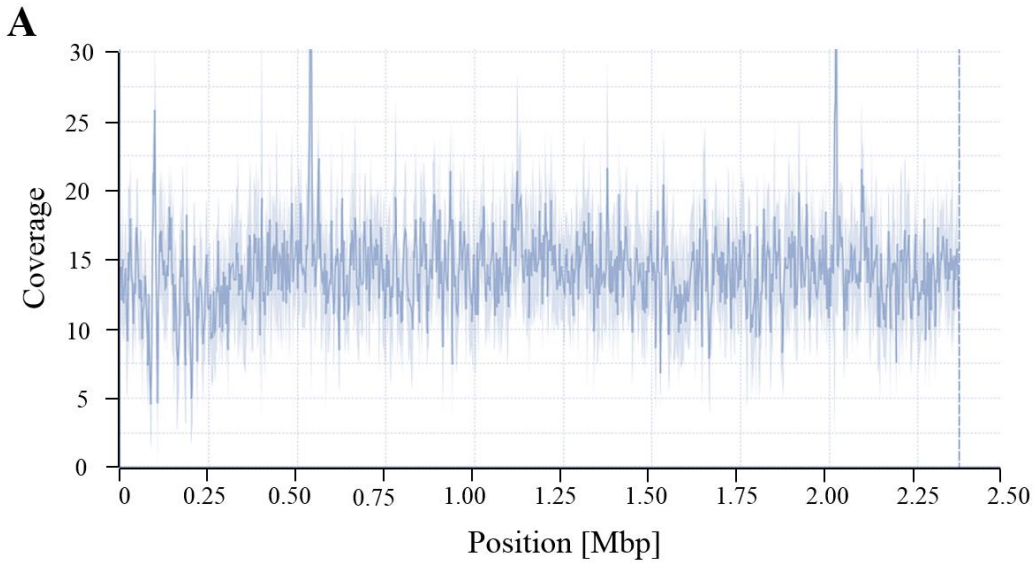
WT_{150}) were subsequently mapped to the *Synechococcus* sp. WH7803 published genome sequence using BWA mem (see section 2.14.1).

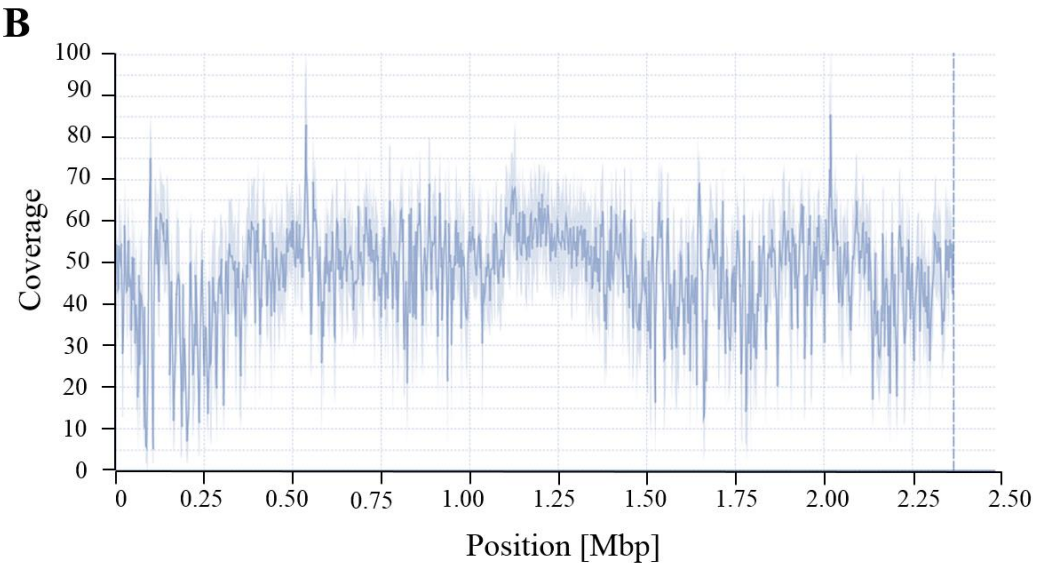
Strain	Mapped reads	Mapped reads [%]	Coverage mean
A	159,266	10.5	14.17±5.34
B	480,564	34.8	47.42±17.14
C	207,302	13.3	18.39±5.53
D	161,624	10.3	14.48±5.00
E	196,612	12.7	17.63±5.20
F	59,522	2.3	4.97±7.47
G	173,980	11.1	15.95±8.24
H	127,774	9.5	11.21±4.43
I	148,448	10.3	13.1±4.67
J	213,458	11.1	19.54±5.96
R1	165,046	15.2	14.62±4.77
R2	239,558	10.9	21.65±6.54
PHR	578,871	37.5	52.00±13.14
WT_{250}	1,631,902	99.2	151.76±26.98
WT_{150}	2,358,026	98.7	144.87±22.92

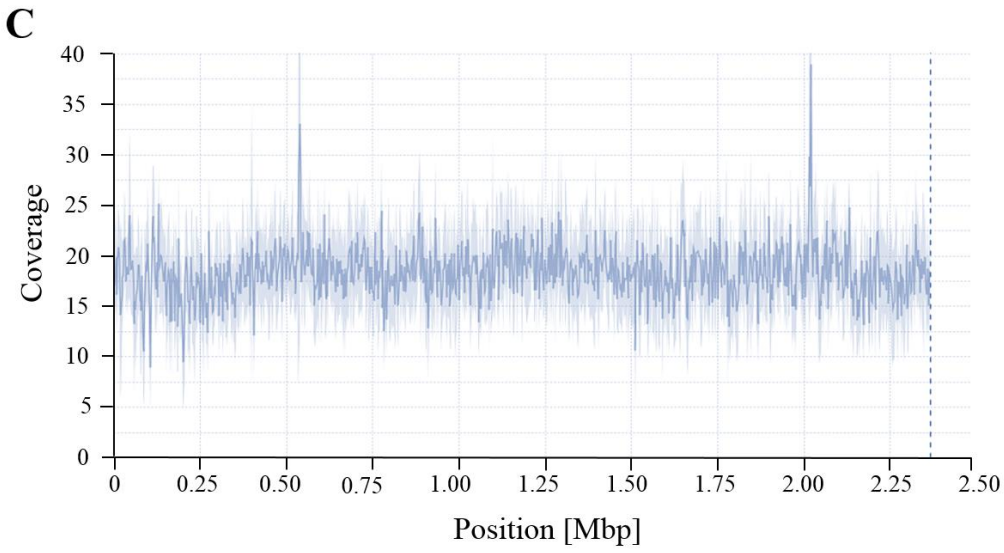
Table 4.4. Alignment of *Synechococcus* sp. WH7803 to the published reference sequence (NC_009481.1) using BWA mem.

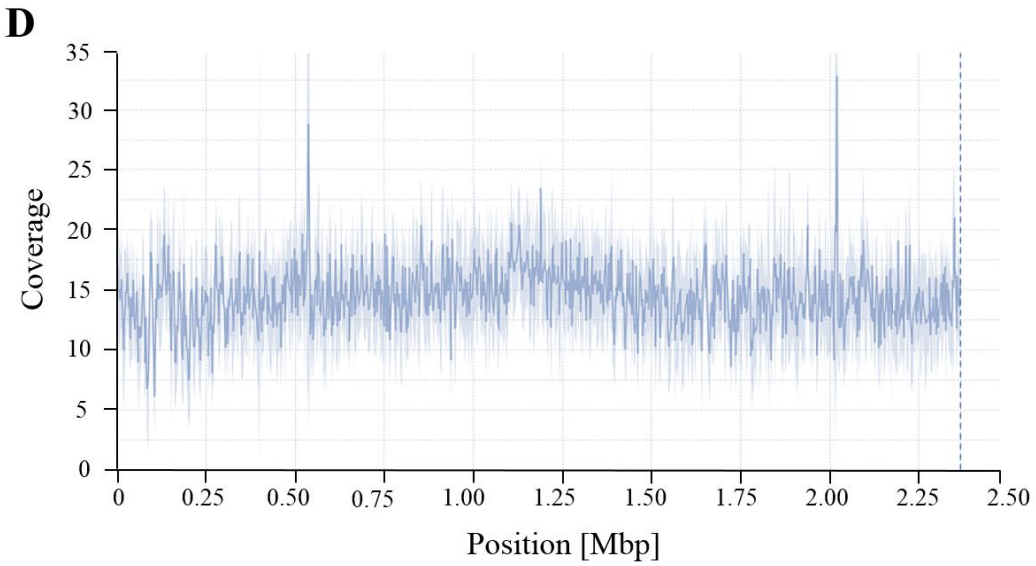
A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2; WT_{250}: wild type *Synechococcus* sp. WH7803 250 bp paired-end library; WT_{150}: wild type *Synechococcus* sp. WH7803 150 bp paired-end library. The coverage was obtained using Qualimap (see section 2.14). The mapped reads are the percentage of mapped reads as a proportion of the total number of reads obtained for each sample [%].

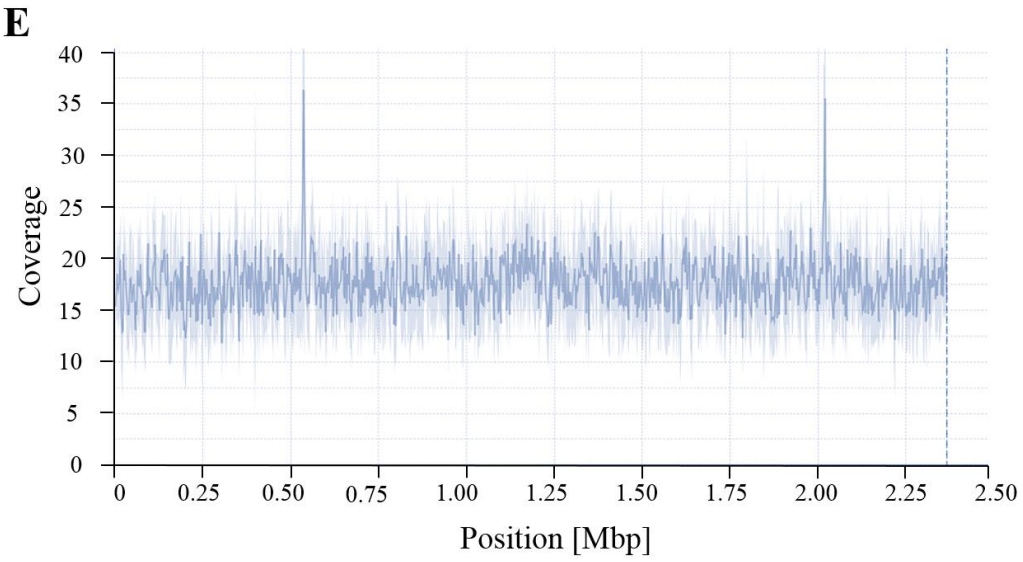
The sequencing coverage of each nucleotide from WT and mutant genomes compared to the published *Synechococcus* sp. WH803 genome sequence is shown in Fig. 4.4. Coverage values range between 9.5 to ~152 fold across all the samples, with the exception of Mutant F where the coverage was 4.97±7.47 fold (Fig. 4.4.F). Due to the low coverage and poor alignment of Mutant F, this sample was excluded from all further analysis.

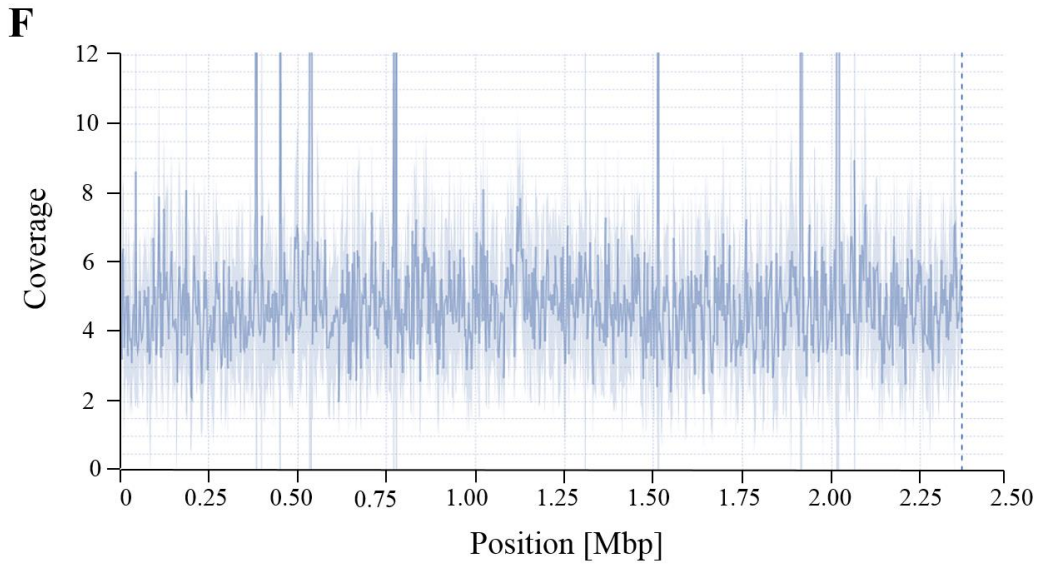

A
Coverage
0
5
10
15
20
25
30
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

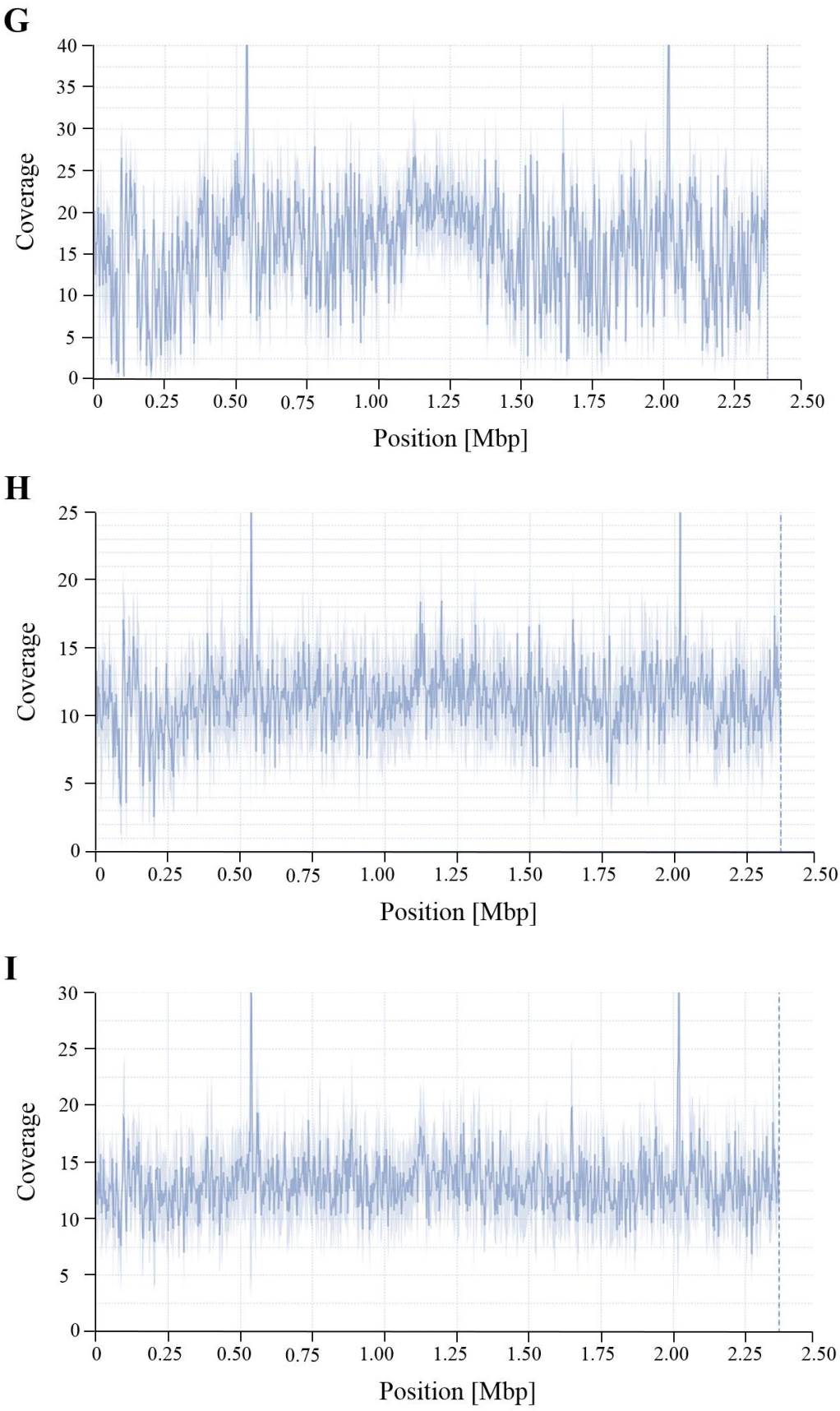

B
Coverage
0
10
20
30
40
50
60
70
80
90
100
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

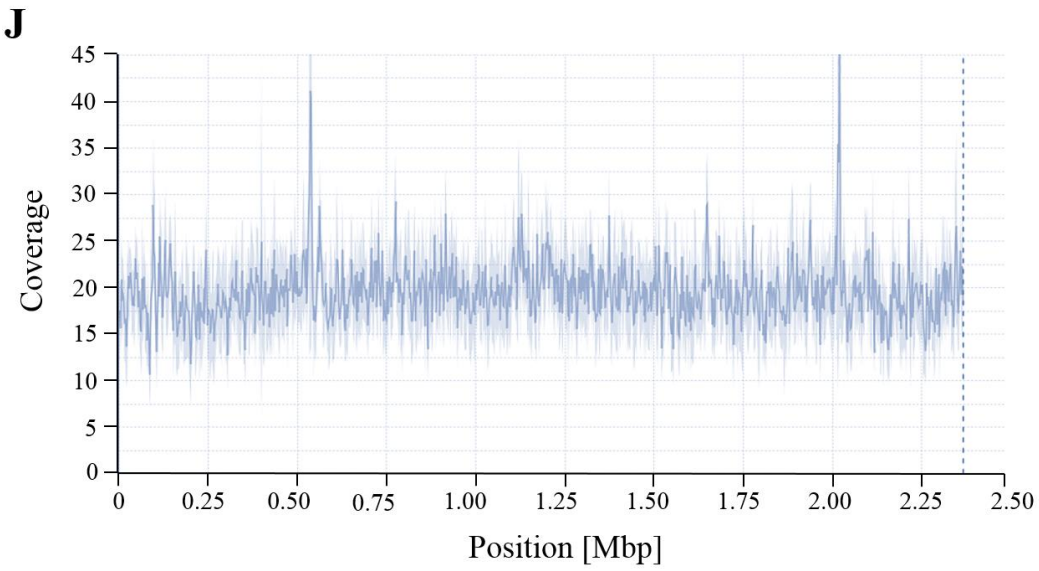

C
Coverage
0
5
10
15
20
25
30
35
40
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

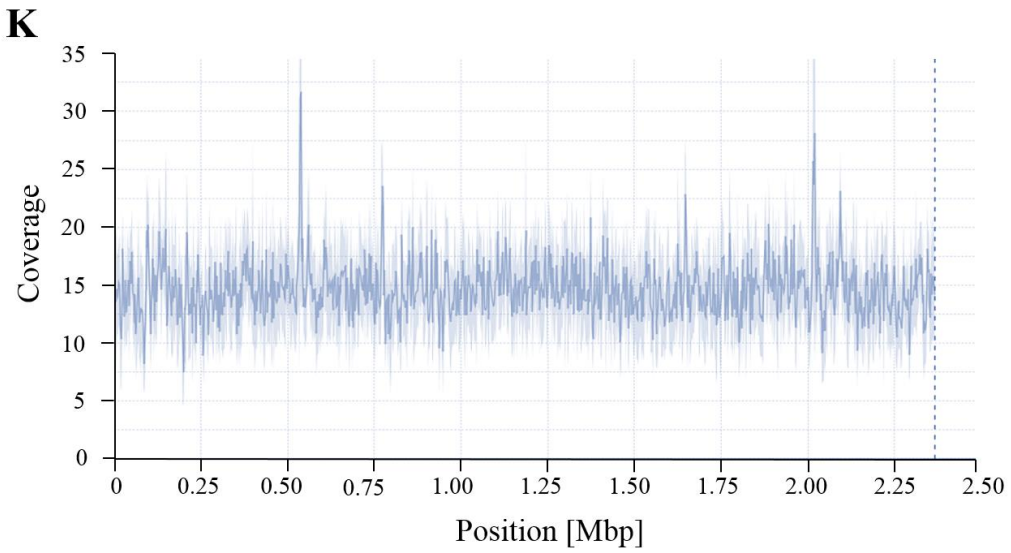

D
Coverage
0
5
10
15
20
25
30
35
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

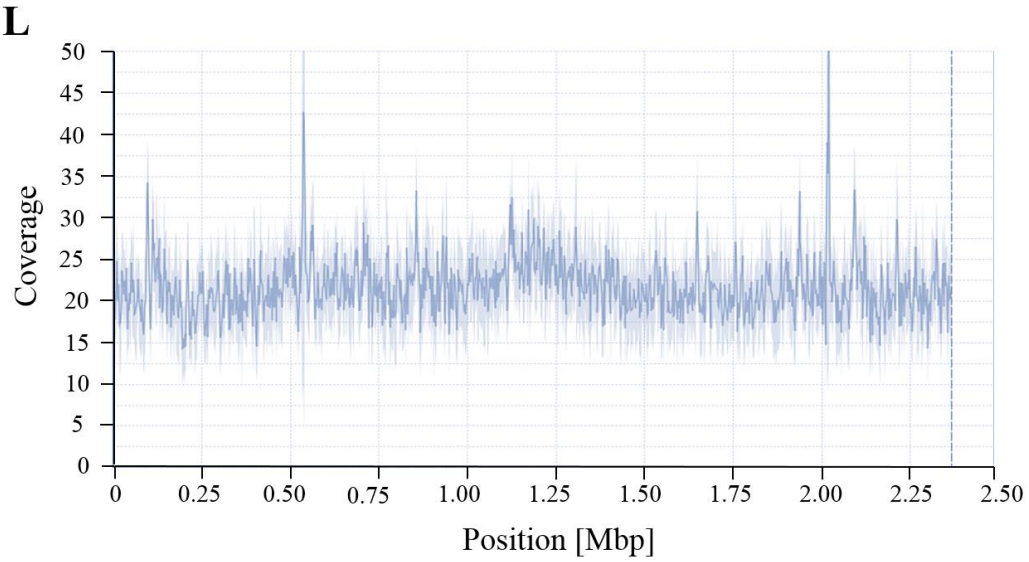

E
Coverage
0
5
10
15
20
25
30
35
40
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

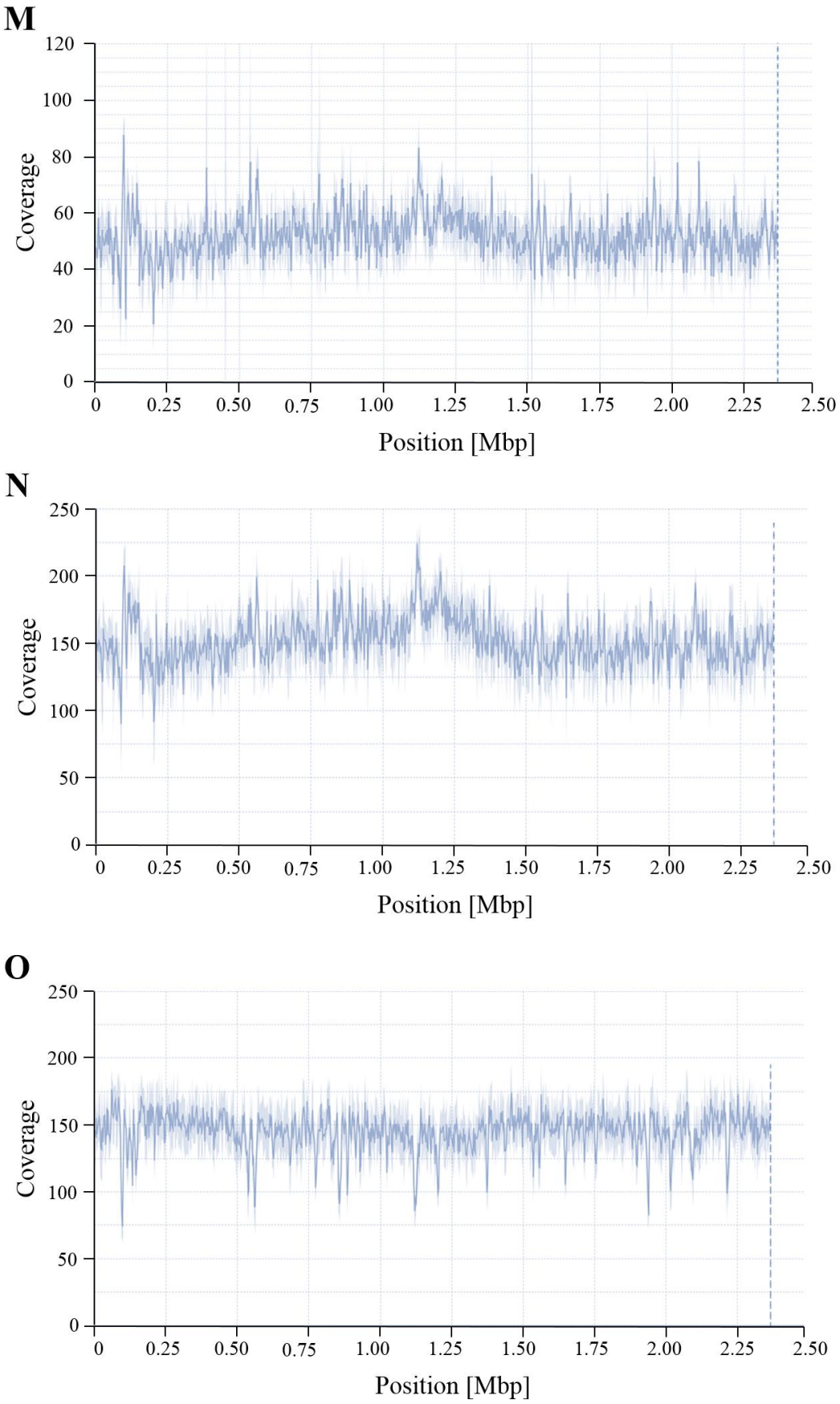

F
Coverage
0
2
4
6
8
10
12
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]


G
Coverage
Position [Mbp]
H
Coverage
Position [Mbp]
I
Coverage
Position [Mbp]


J
Coverage
Position [Mbp]


K
Coverage
Position [Mbp]


L
Coverage
Position [Mbp]


M
Coverage
Position [Mbp]
N
Coverage
Position [Mbp]
O
Coverage
Position [Mbp]

(Previous page)

Figure 4.4. Coverage of sequence reads from each WT and mutant strain aligned with the published *Synechococcus* sp. WH7803 genome.

Fold coverage for each mutant and WT *Synechococcus* sp. WH7803 genome using BWA mem as the alignment algorithm. (A) Mutant A; (B) Mutant B; (C) Mutant C; (D) Mutant D; (E) Mutant E; (F) Mutant F; (G) Mutant G; (H) Mutant H; (I) Mutant I; (J) Mutant J; (K) Mutant R1; (L) Mutant R2; (M) Mutant PHR; (N) Wild type WT_{250}; (O) Wild type WT_{150}. The figures were modified from Qualimap output (see section 2.14).

The two coverage 'spikes' observed in the alignment of all mutants (Fig. 4.4), but not in the wild type, at positions 534,563 bp and 2,019,450 bp of the *Synechococcus* sp. WH7803 genome, correspond to the high number of reads aligning to the 16S rRNA gene (Fig. 4.5), where there are two identical copies in *Synechococcus* sp. WH7803 and most other marine *Synechococcus* (see Dufresne et al., 2008). Here, the alignment algorithm cannot differentiate between reads that are conserved in the 16S rRNA gene of *Synechococcus* sp. WH7803 and heterotrophic bacteria, hence mapping all reads to those regions. To avoid misinterpretation, these regions were excluded from the variant analysis.

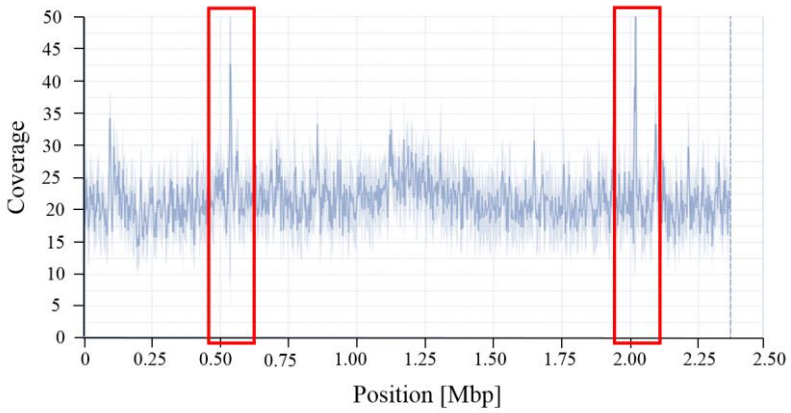


Figure 4.5. Schematic representation of coverage, using the BWA mem alignment algorithm, across the *Synechococcus* sp. WH7803 genome highlighting the two 16S rRNA gene regions in red.

The BAM files (binary alignment map files; see section 2.14.1) were processed and 'variants' were called using VarScan (see section 2.14.2). The number of mutations, including indels and SNPs, found in each of the *Synechococcus* sp. WH7803 cyanophage resistant mutants as well as the WT_{150} and WT_{250} strains, compared to the published *Synechococcus* sp. WH7803 genome sequence is summarised in Table 4.5.

Strain	Number of indels	Number of SNPs	Total number of mutations
A	12	305	317
B	8	122	130
C	4	142	146
D	6	228	234
E	7	154	161
G	10	190	200
H	6	128	134
I	4	136	140
J	7	212	219
PHR	10	263	273
R1	8	144	152
R2	7	189	196
WT_{150}	6	2,143	2,149
WT_{250}	8	66	74

Table 4.5. Number of mutations found in *Synechococcus* sp. WH7803; indels: insertions or deletions; SNPs: single nucleotide polymorphisms.

The total number of mutations found for each *Synechococcus* sp. WH7803 mutant and re-sequenced WT is higher than 100, except WT_{250} with 74 mutations. This greatly exceeds the number reported by (Marston et al. 2012) where genome sequencing of two *Synechococcus* sp. WH7803 cyanophage resistant mutants revealed only four changes (though in four different coding regions) between the WT and mutants: three nucleotide substitutions and a single nucleotide deletion, leading to an early stop codon.

The high number of 'mutations' detected is, to some extent, explained by genes present in multi-copy in the genome. Thus, a total of 435 'mutations' across all *Synechococcus* sp. WH7803 mutant and WT strains were detected in the *psbA* gene, which exists as four virtually identical copies in this strain (see Garczarek et al., 2008) as well as 103 'mutations' in *psbD* (with two near identical copies). This is because the alignment algorithm mis-assigns reads to the four *psbA* and two *psbD* genes. Hence, 'mutations' in these two genes were excluded from further analysis.

Other genes present in more than one copy in the genome e.g. *ftsH* (see section 4.3.3.1) also gave rise to mis-assigned 'mutations', as well as genes present in single copy but with some level of identity to other genes in the genome, e.g. the P-stress porins encoded by the genes *synWH7803_2235* and *synWH7803_2236* which possess 72.7% identity at the nucleotide level. Likewise, genes with homopolymer or repeated regions, like *synWH7803_2113* that encodes a hypothetical protein, also contributed a high number of 'mutations' to the total detected.

In these latter cases (i.e. except the *psbA* and *psbD* genes where the identity between orthologs was extremely high) a manual inspection of reads was performed to sieve out 'mis-aligned' mutations from potentially 'real' ones. As an example an in-depth analysis of the genes encoding FtsH is presented below.

4.3.3.1. Identification of 'mis-aligned' versus potentially 'real' mutations in the multi-copy *ftsH* gene family from *Synechococcus* sp. WH7803

Four copies of the *ftsH* gene, *synWH7803_0355*, *synWH7803_1116*, *synWH7803_1216* and *synWH7803_1689* exist in the *Synechococcus* sp. WH7803 genome. Each of these four genes encode regions with similar protein domains: an

AAA domain that belongs to an ATPase family related to various cellular activities, a peptidase M41 domain and a FtsH extracellular domain (except gene *synWH7803_1116* which only contains the AAA and peptidase M41 domains).

WGS analysis identified SNPs in the *ftsH* gene from cyanophage-resistant mutants B, E, G, J, PHR as well as the re-sequenced wild type (WT_{150}).

In total, 51 SNPs were found in WT_{150}, of which 49 were silent mutations (15 in *synWH7803_1689* and 35 in *synWH7803_0355*) and one was a non-synonymous SNP (in gene *synWH7803_0355*). Moreover, 27 SNPs were identified in the mutants, of which 24 were silent mutations (in mutants B, G and PHR) and 3 were non-synonymous SNPs (in genes *synWH7803_1116* (Mutant E) and *synWH7803_1689* (strains B and J)). No mutations were found in *synWH7803_1216*.

A manual analysis of the reads aligning to these genes confirmed that all of them belonged to *Synechococcus* sp. WH7803, but that some of the reads were misaligned. In Mutant E, the total number of reads supporting the mutant genotype in gene *synWH7803_1116* was 17, not 7 as initially detected, because 10 reads were misaligned with the other three genes.

In the case of Mutants B and J, the same non-synonymous mutation was detected in gene *synWH7803_1689*. Of the 51 reads obtained from mutant B, 50 support the mutant genotype and one read aligned with gene *synWH7803_0355*. For Mutant J, 17 reads support the mutant genotype and no reads were found to support the wild type genotype. This manual re-analysis confirmed that mutants B and J both possess a non-synonymous SNP in gene *synWH7803_1689* with the mutation present in all the sequence reads, i.e. at a 100% frequency, at this position.

In contrast, manual analysis of the WT_{150} sequence reads aligning to the four *ftsH* genes confirmed that none of the four genes had a mutant genotype; potential 'mutations' in genes *synWH7803_0355* and *synWH7803_1689* were all properly mapped to gene *synWH7803_1216*.

As a result of this manual inspection of sequence reads in multi-copy genes or those containing homopolymer or repeat regions, the number of mutations detected in each mutant and re-sequenced wild type WT_{150} decreased (Table 4.6). However, the number

of detected mutations was still extremely high, especially when considering other studies (Avrani et al. 2011; Marston et al. 2012).

Strain	Number of indels	Number of SNPs	Total number of mutations
A	8	192	200
B	8	121	129
C	4	140	144
D	6	212	218
E	5	132	137
G	10	189	199
H	6	127	133
I	4	134	138
J	7	211	218
PHR	7	93	100
R1	7	134	141
R2	7	187	194
WT_{150}	6	2,118	2,124
WT_{250}	8	66	74

Table 4.6. Number of mutations found in *Synechococcus* sp. WH7803 after manual exclusion of mis-aligned reads; indels: insertions/deletions; SNPs: single nucleotide polymorphisms.

4.3.4 Re-analysis of sequence alignments including only those 'mutations' that occur at > 90 % frequency

The high number of mutations identified in all the cyanophage-resistant and re-sequenced WT strains compared to the published *Synechococcus* sp. WH7803 genome sequence, even after manual inspection to exclude read mis-alignment in multi-copy genes (Table 4.6), was still not conducive with identifying those mutations that might confer cyanophage resistance to *Synechococcus* sp. WH7803.

Noteworthy though, a common practice when analysing bacterial WGS is to determine mutations by discarding any 'mutations' that occur in < 90 % of the reads at a specific location (Lewis et al., 2010). Such is the mutation frequency, i.e. the number of reads that are 'mutant', as a percentage of the total number of reads, mapping to a specific location. When a > 90 % mutation frequency threshold was incorporated into the bioinformatics pipeline, a significant reduction in the number of SNPs and indels found in all strains was obtained (see Table 4.7), though the numbers are still high compared to previous literature (Marston et al. 2012).

Strain	Number of indels	Number of SNPs	Total number of mutations
A	4	13	17
B	6	20	26
C	3	12	15
D	3	18	21
E	4	15	19
G	4	15	19
H	4	10	14
I	3	17	20
J	4	13	17
PHR	7	53	60
R1	4	26	30
R2	5	26	31
WT_{150}	4	17	21
WT_{250}	7	45	52

Table 4.7. Number of mutations occurring at a frequency > 90 % in each cyanophage-resistant and re-sequenced WT *Synechococcus* sp. WH7803.

Given the large difference in the total number of mutations detected for each strain following excluding reads that were mis-aligned (Table 4.6) compared to only including 'mutations' occurring at a frequency > 90 % (Table 4.7), the frequency versus number of mutations for all cyanophage-resistant and wild type strains was plotted (Fig. 4.6). Such a plot shows a high number of mutations with frequencies varying between 10-100 %.

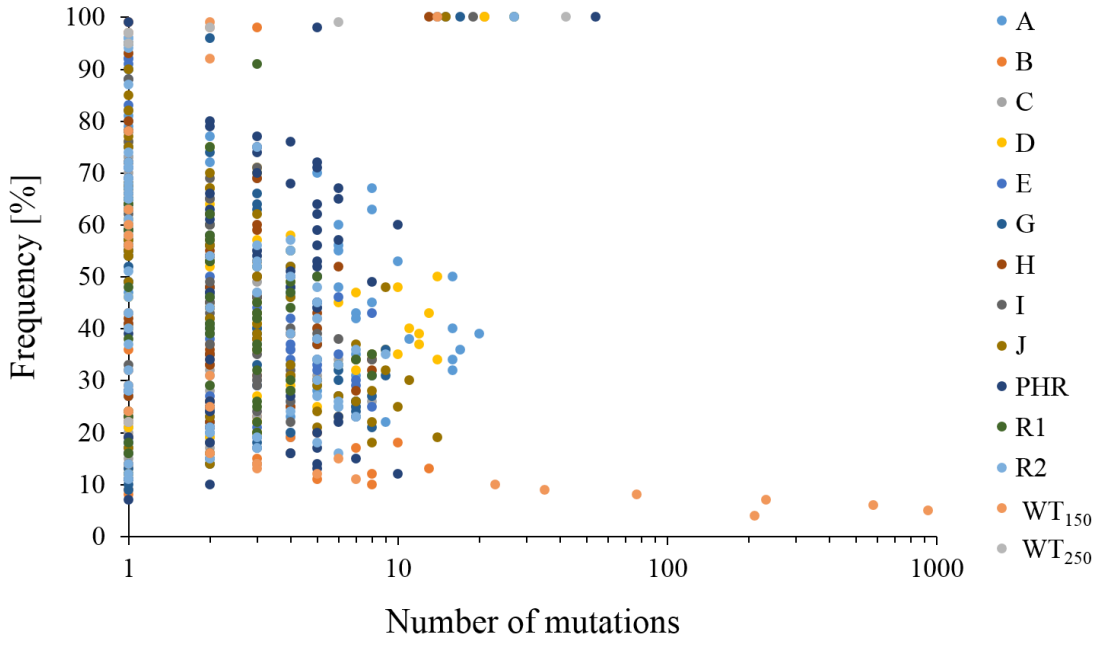


Figure 4.6. Frequency of mutations found in *Synechococcus* sp. WH7803 cyanophage-resistant and WT strains compared to the published genome sequence.

A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2; WT_{250}: wild type *Synechococcus* sp. WH7803 250 bp paired-end library; WT_{150}: wild type *Synechococcus* sp. WH7803 150 bp paired-end library. Frequency is the percentage of 'mutant' reads compared to the total number of reads for that specific location.

In order to compare the results obtained with previously published data and to assess any 'bioinformatic pipeline problems', sequence reads from the ancestral WT *Synechococcus* sp. WH7803 (chemostat A, time 0) reported in (Marston et al. 2012) were obtained from the CAMERA database (see section 2.14.1) and subjected to the bioinformatic pipeline described in section 2.14.1 and 2.14.2. Again, sequence reads were aligned to the previously published *Synechococcus* sp. WH7803 genome sequence (NC_009481.1), though reads were mapped using Bowtie2 due to the fact that the CAMERA database was no longer available when the BWA mem algorithm became available (Table 4.8).

Moreover, to determine whether the widely varying mutation frequency observed in *Synechococcus* sp. WH7803 was an isolated phenomenon or also occurred in closely

related organisms, we took advantage of available WGS reads obtained from a cyanophage resistant *Prochlorococcus marinus* MED4 mutant R28 (Avrani et al., 2011), which were kindly provided by Dr. D. Lindell in fastq format. Alignment of these WGS reads from this cyanophage-resistant mutant compared to the published *Prochlorococcus marinus* MED4 genome sequence (NC_005072.1) was performed using BWA mem as the alignment algorithm, and a similar bioinformatics pipeline as that described previously (Table 4.8).

Strain	Mapped reads	Mapped reads [%]	Coverage mean
AC0	194,152	65.3	18.04±5.36
R28	13,007,100	81.7	438.86±300.71

Table 4.8. Alignment of published WGS reads from ancestral WT *Synechococcus* sp. WH7803 (Marston et al. 2012) and a cyanophage-resistant *Prochlorococcus marinus* MED4 mutant (Avrani et al. 2011).

AC0: ancestral WT *Synechococcus* sp. WH7803 from chemostat A time 0; R28: cyanophage resistant *Prochlorococcus marinus* MED4 mutant. The coverage mean was obtained using Qualimap (see section 2.14). Mapped reads are the percentage of mapped reads as a proportion of the total number of reads for each sample [%].

Following read mapping, a profile of the mutation frequency was derived for reads obtained from the ancestral *Synechococcus* sp. WH7803 WT (see Marston et al. 2012). Such a profile (Fig. 4.7) is very similar to that found in this work (Fig. 4.6), revealing many mutations present at varying frequency.

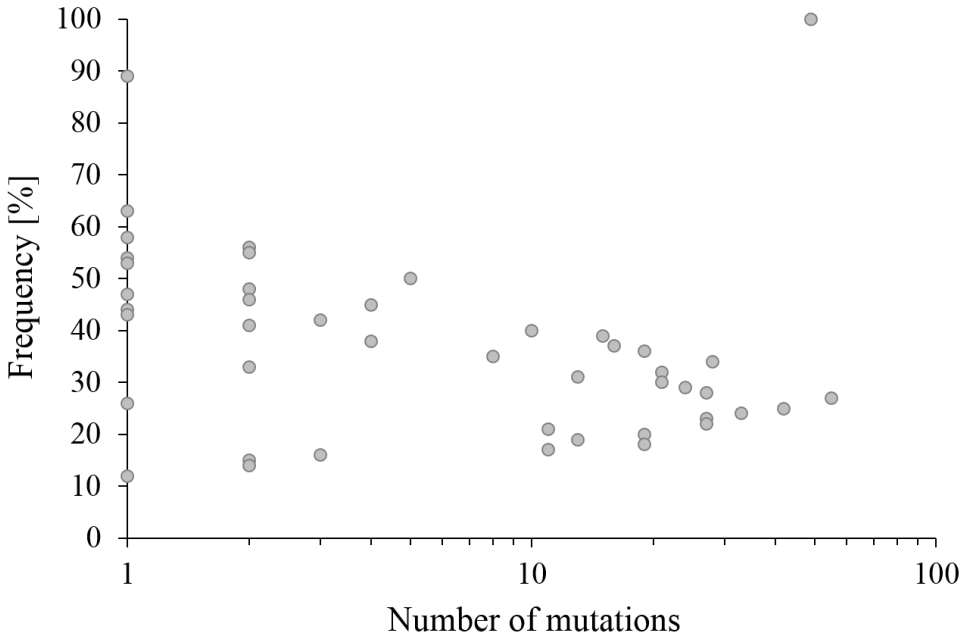


Figure 4.7. Mutation frequency in sequence reads from ancestral WT *Synechococcus* sp. WH7803 (Marston et al. 2012) compared to the previously published *Synechococcus* sp. WH7803 genome sequence.

Reads were obtained from the CAMERA database in fasta format and aligned to the published sequence of *Synechococcus* sp. WH7803 (NC_009481.1). Frequency is the percentage of 'mutant' reads compared to the total number of reads for that specific location.

In contrast, the mutation frequency profile found for the cyanophage-resistant *Prochlorococcus marinus* MED4 mutant R28 (Fig. 4.8) showed all mutations occurring with a frequency > 90 % or < 15 % (Table 4.8). Those mutations occurring at a frequency < 15 % likely relate to sequencing errors associated with the high depth of coverage for this strain (see Table 4.8). The only mutation found within this range is an insertion mutation occurring at a frequency of 62 %, which corresponds to a T-rich homopolymeric intergenic region that could lead to read mis-alignment. With this analysis, 10 mutations occur at a frequency of 100 % which were exactly those detected by Avrani et al. (2011).

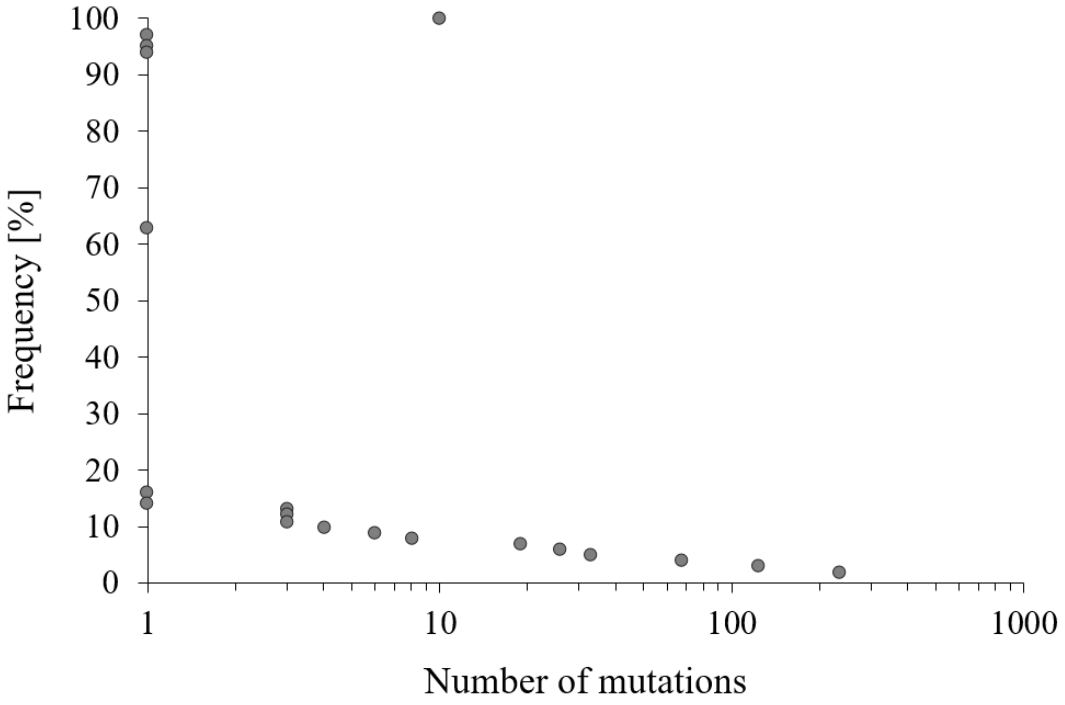


Figure 4.8. Mutation frequency in sequence reads from the cyanophage resistant *Prochlorococcus marinus* MED4 mutant R28 compared to the published *Prochlorococcus marinus* MED4 genome sequence.

Frequency is the percentage of 'mutant' reads compared to the total number of reads for that specific location.

To elucidate whether this variation in mutation frequency found in *Synechococcus* sp. WH7803 was due to errors introduced using the Illumina sequencing platform or if it is an artefact of the pipeline used for the analysis (see section 2.14), the program Art (Huang et al. 2012) was used to simulate Illumina sequencing reads from the published *Synechococcus* sp. WH7803 genome sequence (section 2.14.3). Simulation of 250 bp paired-end reads was repeated several times varying the insert size up to 6,000 bp and the coverage 10, 30 and 100 fold. The output file in fastq format, containing the simulated reads, was put through the same bioinformatic pipeline to identify mutations as described previously (sections 2.14.1 and 2.14.2) but no variation in mutation frequency was observed under any of the conditions, suggesting that the mutation frequency profile is not related to either the Illumina sequencing platform or the bioinformatic pipeline used.

4.3.5 DNA sequencing of cyanophage S-PM2 from cyanophage-resistant *Synechococcus* mutants

The cyanophage resistant *Synechococcus* sp. WH7803 mutants R1, R2 and PHR were routinely maintained with the addition of S-PM2 (see Section 2.1.2). To determine if there were any associated nucleotide changes in the cyanophage genome, the reads from mutants R1, R2 and PHR were mapped to the cyanophage S-PM2 sequence (NC_006820.1). This analysis was contributed as part of the work published by Puxty et al. (2015). Briefly, the genome comparison with the published S-PM2 sequence (NC_006820.1) identified a missing region of ~ 10 kb (13 % of the genome) at a locus containing ORFs not homologous to any protein (ORFans). After determining there was a fitness cost associated with the original S-PM2 (S-PM2^{WT}) possessing this ORFanage region, as well as performing a metagenomic analysis of these ORFan genes in the environment, we hypothesised that this region confers flexibility and a fitness advantage to this cyanophage potentially in freshwater and hypersaline environments (Puxty et al., 2015).

VarScan (see Section 2.14.1) was used to identify any mutations in the re-sequenced S-PM2 cyanophage and these mutations were confirmed by Sanger sequencing. Mutations are reported in Puxty et al., (2015) and the modified S-PM2 genome sequence is deposited in EMBL-EBI (Accession number LN828717; Puxty et al. 2015). The variation in frequency of mutations in cyanophage S-PM2 (Fig. 4.9) showed a similar profile to that observed in *Prochlorococcus* (Fig. 4.8).

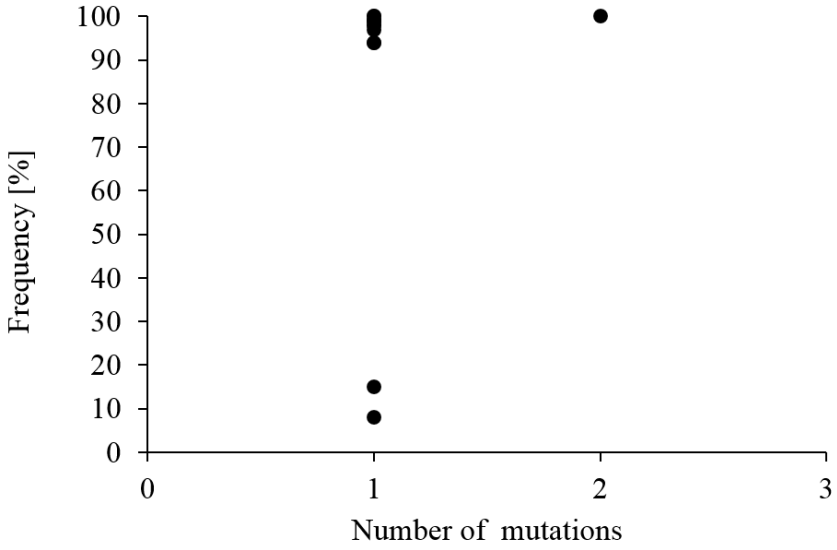


Figure 4.9. Mutation frequency in the re-sequenced cyanophage S-PM2 genome compared to the previously published sequence.

Mutation frequency is the percentage of 'mutant' reads compared to the total number of reads for that specific location.

4.3.6 Identifying the molecular basis of cyanophage-resistance in *Synechococcus* sp. WH7803

In order to identify the genes responsible for cyanophage resistance in *Synechococcus* sp. WH7803, firstly genes with mutations only present in the wild type were discarded. In total, the wild type *Synechococcus* sp. WH7803 contained mutations in 105 genes (compared to the original published sequence), of which nine were also present in all or most of the cyanophage-resistant mutants. Interestingly, of these nine mutations, two had a higher mutation frequency in the mutants compared to the wild type (Table 4.9).

Table 4.10 lists all the mutations that were not present in the re-sequenced wild type but which occurred in cyanophage-resistant mutant(s) with a frequency > 10 %. Due to the high number of mutations in the cyanophage resistant mutants, it is still unclear which genes confer cyanophage resistance (Table 4.10).

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2	WT
1446|1447	Intergenic region	100	100	100	100	100	100	100	100	100	100	100	100	100
0980	ABC-type Mn^{2+} transport system permease component	100	100	100	100	100	100		100	100	100	100	100	99
0996	Hypothetical protein	100	100	100	100	100	100	100	100	100	100	100	100	99
0749	Endopeptidase Clp ATP-binding chain B	100	98	100	100	100	100	100	100	95	100	91	95	97
0371	Fe-S oxidoreductase	100	99	93	100	91	100	93	100	100	98	100	100	92
0791	HIT family hydrolase	100	95			100	96	100	88	100	98	91	94	99
1272	Hypothetical protein	96	98	100	100	100	100	100	100	100	100	100	97	98
0738	Hypothetical protein	100	100	100	100	100	100		100	100		100	100	60
0273	Choline-glycine betaine transporter	90	100	50	100	90	100	38	100	90	100			31

Table 4.9. Mutation frequency in genes where there were mutations in both *Synechococcus* sp. WH7803 cyanophage resistant mutants and re-sequenced wild type.

A-J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42, R1 and R2; *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2; WT: wild type *Synechococcus* sp. WH7803 (including WT_{150} and WT_{250}). Mutation frequency is the percentage of reads possessing the mutation compared to the total number of aligned reads. Frequencies of reads aligned to *psbA* and < 10 % were filtered out. Genes with a higher mutation frequency in the mutants compared to the wild type are highlighted in red.

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2
0048	GMP synthase											32	
0064	ATP-dependent protease ATP-binding subunit ClpX										100		
0092	UDP-glucose-4-epimerase											100	100
0103	dTDP-4-dehydrorhamnose 3,5-epimerase											91	
0105	Sugar transferase		100		100	95			100	85	100	100	100
0122	Hypothetical protein										99		
0124	Glycosyl transferase family protein										100		
0134	Hypothetical protein		100										
0138	Glycosyl transferase family protein										100		
0141	Preprotein translocase subunit SecA	75		45	67	63		80		37			
0169	S-adenosyl-L-homocysteine hydrolase	39		34	48	48		44		56			
0176	Two-component system response regulator										100		
0190	Thiamine biosynthesis protein ThiC	45			40	46		40	38	38			

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2
0199	Glucosamine--fructose-6-phosphate aminotransferase					34							
0200	Mannose-1-phosphate guanylyltransferase	67				92							
0209|0210	Intergenic region	100	100	100	100	100	100	100	100	100		100	100
0217	Isocitrate dehydrogenase			39	59	35		60					
0230	Hypothetical protein											100	100
0379	Elongation factor Tu	55		47	50	44	50	67		37			
0385	Ferredoxin-dependent glutamate synthase									25			
0411	DNA-directed RNA polymerase subunit alpha										100		
0445	Fructosamine kinase						48						
0529	Septum site-determining protein MinD		100										
0558	Hypothetical protein										100		
0586	Hypothetical protein											100	100
0707	Biopolymer transport protein											100	100
0718|0719	Intergenic region		100										

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2
0737	Hypothetical protein											100	100
0748	Chaperone ClpB									52			
0870	Hypothetical protein	92		55		88							
0877	Glutathione-dependent formaldehyde dehydrogenase								42	35			
0885	Thiopurine S-methyltransferase												100
0940	Hypothetical protein											100	100
0948	Putative multicopper oxidase						36						
0978	ABC-type Mn^{2+} transport system, periplasmic component											100	100
1040	Acetyl-coenzyme A synthetase									32			
1118|1119	Intergenic region												87
1305	Major facilitator superfamily permease				40								
1316	D-alanyl-D-alanine carboxypeptidase											100	
1328	Elongation factor Ts		100	65	100		96	43	95	82			
1384	Carbonic anhydrase		100	69	100		100	54	100	78			

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2
1405	Major facilitator superfamily sugar transporter		100		100		100	43	100	77			
1415	ABC-type amino acid transport system permease component											75	
1422	Putative sirohydrochlorin cobaltochelatase										27	100	100
1450	Two-component system sensor histidine kinase				45								
1499	Hypothetical protein					78							
1532	Membrane protein, metal dependent phosphoesterase		33										
1544	Type II alternative sigma-70 family RNA polymerase sigma factor	100	100	100	100	100	100	100	100	100		100	100
1553	Ribose-phosphate pyrophosphokinase	100	100	100	100	100	100	100	100	100		100	100
1554	Hypothetical protein	100	100	100	100	100	100	100	100	100	100	100	100
1715	Riboflavin biosynthesis protein *ribF* (riboflavin kinase / FMN adenylyltransferase)	100	100	100	100	100		100	100	100		100	100
1724	tRNA-(guanine-N1)-methyltransferase										98		
1797	Hypothetical protein											100	100

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2
1821	Glutamate-1-semialdehyde aminotransferase	67											
1832	Methionine aminopeptidase	100	100	100	100	100	100	100	100	100		100	100
1875	Hypothetical protein											55	
1958	UDP-3-O-[3-hydroxymyristoyl] N-acetylglucosamine deacetylase		100										
1994|1995	Intergenic region				100								
2000	F0F1 ATP synthase subunit beta										21		
2040	Superfamily II RNA helicase										100		
2061	DNA-directed RNA polymerase subunit gamma	63	18	53	43	82		64		41			
2062	DNA-directed RNA polymerase subunit beta	50		34	47								96
2069	Carbohydrate-binding protein											67	
2070	Two-component system response regulator											100	100
2152	Fructose-bisphosphate aldolase												
2153	Transcriptional regulator										100		

Gene ID	Gene product	A	B	C	D	E	G	H	I	J	PHR	R1	R2
2213	Shikimate kinase						54						
2232	Hypothetical protein											75	100
2285|2286	Intergenic region										100		
2351	Phosphoglycerate kinase		100	100	100	100	100	100	100	100		100	100
2461	CTP synthetase	44			39	43							
2493	Formate/nitrate transporter										100		

Table 4.10. Mutation frequency in genes that possess mutations in the *Synechococcus* sp. WH7803 cyanophage resistant mutants but which are absent in the re-sequenced WT.

A - J: *Synechococcus* sp. WH7803 resistant to cyanophage S-RSM42; R1 and R2: *Synechococcus* sp. WH7803 resistant to cyanophages S-RSM42 and S-PM2; PHR: *Synechococcus* sp. WH7803 resistant to cyanophage S-PM2. The frequency is the percentage of reads possessing the mutation compared to the total number of aligned reads. Frequencies of reads aligned to *psbA* and < 10% were filtered out.

4.4. Discussion

This chapter described the use of WGS to elucidate the molecular basis of cyanophage resistance in *Synechococcus* sp. WH7803 mutants described and characterised in Chapter 3. Unfortunately, the identification of genes potentially involved in cyanophage resistance was confounded by the complex nature of the data obtained. Certainly, the variation in mutation frequencies observed in the cyanophage-resistant and re-sequenced wild type *Synechococcus* sp. WH7803 have not previously been documented in this strain (e.g. see Marston et al. 2012) and are very different to that observed in other bacteria where mutations essentially appear only at 100 % frequency (see Avrani et al., 2011).

As shown in Table 4.9, mutations in nine genes and in an intergenic region were also present in the wild type, so these were discarded from the final analysis. These mutations are probably either sequence errors in the original genome annotation or were acquired before the isolation of the cyanophage resistant mutants but subsequent to sequencing of the published genome sequence.

Gene *synWH7803_0738*, encoding a hypothetical protein, contained a non-synonymous mutation in all the cyanophage-resistant mutants except mutants G and PHR, and gene *synWH7803_0273*, encoding a predicted choline-glycine betaine transporter also contained a non-synonymous mutation in all the mutants, except mutants R1 and R2. Interestingly, these two genes contained the same mutation in the re-sequenced wild type, but at a lower frequency, tentatively suggesting they may play a role in the development or maintenance of the resistant phenotype.

Whilst the high number of mutations found in all the mutant and re-sequenced wild type *Synechococcus* sp. WH7803 strains made it difficult to determine the genes involved in cyanophage resistance, some inferences can be made from mutations in genes that are present at 100 % frequency only in specific mutants.

Only a non-synonymous mutation in *synWH7803_1554*, encoding a hypothetical protein, was shared by all the cyanophage resistant mutants and absent in the re-sequenced wild type, potentially suggestive of a role in cyanophage resistance.

Noteworthy, five mutations were shared only between the mutants resistant to cyanophage S-RSM42 but absent in the PHR mutant. Of these five mutations, two were non-synonymous mutations in the *rpoD* (*synWH7803_1544*) and *map* (*synWH7803_1832*) genes, encoding a predicted RNA polymerase sigma factor (sigma-70 family) and methionine aminopeptidase, respectively. Although the high number of mutations present in the cyanophage-resistant mutants makes it difficult to make precise conclusions about mechanisms of resistance, it is important to highlight mutations in *rpoD* because it can be related to gene expression and hence be contributing to the resistance phenotype (Davis et al., 1995; Hu and Gross, 1988; Siegele et al., 1988). Curiously, a mutation in s*ynWH7803_1555*, encoding PepN a predicted aminopeptidase N, was also found in a cyanophage resistant *Synechococcus* sp. WH7803 mutant reported by Marston et al. (2012).

In total, there were mutations in 72 genes in the cyanophage resistant mutants compared to the re-sequenced wild type of which 38 where exclusive to a specific mutant. Mutant PHR contained 13 mutations not shared with any of the other mutant, of which mutations in two glycosyltransferases (*synWH7803_0124* and *synWH7803_0138*) and a two-component system response regulator (*synWH7803_0176*) are particularly interesting. Mutations in glycosyltransferases have previously been reported in cyanophage resistant *Prochlorococcus* (Avrani et al., 2011; Avrani & Lindell 2015) and *Synechococcus* sp. WH7803 mutants (Marston et al. 2012), the latter in a different gene (*synWH7803_0140*) to those reported here. *synWH7803_0176* is predicted to be a member of the two-component system family of regulators, specifically of the OmpR family (using annotations based within cyanorak, see http://abims.sb-roscoff.fr/cyanorak/; Scanlan et al. 2009). In *E. coli* OmpR regulates the expression of outer membrane porin proteins and, as discussed previously (see Chapter 3), porins and LPS can work together as phage receptors (Nara et al., 1986; Rakhuba et al., 2010).

Unfortunately, no general conclusions on cyanophage resistance mechanisms can be made from the WGS data obtained due to the high number of mutations found. Ideally, confirmation of specific mutations would be performed using Sanger sequencing. However, the high number of mutations made it difficult to choose which genes to target. Furthermore, the varying mutation frequency observed may be playing a more

direct role as a mechanism of cyanophage resistance, perhaps by allowing differential expression of specific alleles that could confer resistance. Because of the fitness cost associated with cyanophage resistance, keeping different alleles could be beneficial for *Synechococcus* in maintaining flexibility and adaptability to different conditions.

Clearly, further study of the role of mutation frequency variability with respect to cyanophage resistance mechanisms is warranted.

With respect to the latter, bioinformatic analysis of WGS reads is central to understanding this varying mutation frequency phenomena. Noteworthy is that the use of different alignment algorithms had an effect on the output obtained. Initially, Bowtie2 was used to analyse the WGS data, being the best alignment algorithm available at the time. Subsequently, BWA mem and NextGenMap became available and use of these algorithms decreased the number of mutations found compared to Bowtie 2 by one order of magnitude for some samples (Table 4.3). Because NextGenMap and BWA mem only became available late in the analysis, the exceptionally high number of mutations found made it difficult to determine true mutations from misalignments early during this research.

One way to filter out some of the misalignments was by using the soft clipping (–local) option in Bowtie2 that trims the ends of the reads to align them better. After using this option the number of mutations significantly decreased in some of the samples (data not shown) indicating that the end-to-end alignment to *Synechococcus* sp. WH7803 was deterring the detection of real mutations. The new aligners mentioned above, i.e. NextGenMap and BWA mem, have soft clipping options included, making them more precise in the alignment by reducing the general error rate from 0.51 % to 0.09 % (Table 4.2). Between NextGenMap and BWA mem no obvious differences in the number of reads mapped was observed (Table 4.2) and hence BWA mem was used in further analysis.

The only sample that did not significantly decrease the number of mutations found when using BWA mem was WT_{150} (wild type *Synechococcus* sp. WH7803 150 bp paired-end library): 2,333 mutations detected with Bowtie 2 compared with 2,149 mutations detected with BWA mem (Table 4.3). This number plummeted when

considering mutations with a frequency > 10 % only, probably due to the shorter read length (~150 bp compared to ~250 bp) and high number of reads (the compared frequency of mutations for the *Synechococcus* sp. WH7803 samples is shown in Fig. 4.6). The same occurs with the *Prochlorococcus marinus* MED4 sample analysed (Fig. 4.8), which also has a high number of short length reads that could lead to an accumulation of errors. Noteworthy, these *Prochlorococcus* WGS reads are not paired-end and only ~50 bp long, although the high coverage (438.86±300.71; Table 4.7) permits an accurate detection of mutations.

However, even despite the use of different alignment algorithms and sequence read lengths, there is a clear difference in mutation frequency profiles between all the cyanophage-resistant and re-sequenced wild type *Synechococcus* sp. WH7803 obtained here and the previously published *Prochlorococcus marinus* MED4 cyanophage-resistant mutants (see Fig. 4.8 and Avrani et al., 2011). Moreover, no such variation in mutation frequency is seen in sequence reads from cyanophage S-PM2 which was re-sequenced with the mutants (Figs. 4.9). This difference is independent of the sequencing technology and pipeline used, which was demonstrated using the same pipeline with previously published *Synechococcus* sp. WH7803 sequence reads (Marston et al. 2012), an Illumina read simulator (section 2.14.3) and the cyanophage S-PM2 reads as an internal control (section 4.3.6).

One further explanation is that the variation in mutation frequency is due to the co-existence of many different *Synechococcus* cyanophage-resistant populations within a single mutant, though this is unlikely given that all the cyanophage-resistant mutants analysed were isolated from single colonies. The latter could only be assessed by re-sequencing these *Synechococcus* sp. WH7803 mutants using DNA extracted from single cells.

Assuming that each cyanophage-resistant mutant ultimately derived from a single colony then the variation in mutation frequency observed must indicate that there is a difference within the genome of each cell. In this respect, it has been suggested that *Synechococcus* sp. WH7803 can possess more than one copy of its chromosome (Binder and Chisholm, 1995; Griese et al., 2011), and hence is oligoploid, whereas *Prochlorococcus* appears to be monoploid, i.e. possesses only one chromosome copy

(Vaulot et al., 1995). This variation in genome copy number could explain the differences in mutation frequency observed in the WGS data. In *Prochlorococcus* and cyanophage S-PM2 mutations occurred in either all the reads (~ 100 % frequency) or none of them (< 10 % frequency observed), consistent with a single copy genome. In contrast, in *Synechococcus* sp. WH7803 variation in chromosome copy number (which potentially is from 1-10 copies, Binder and Chisholm 1995) could explain the variation in mutation frequency observed.

Interestingly during the course of this thesis, re-sequencing of thermo-tolerant *Synechocystis* sp. PCC 6803 mutants showed a variation in the frequency (or percentage of reads) of mutations. The presence of the mutant and wild type genotype was confirmed by Sanger sequencing, showing no segregation of the mutations (Tillich et al., 2014). Two different genotypes, i.e. the occurrence of a mutant and wild type gene, has also been observed in *Thermus thermophilus* HB8, where both an antibiotic resistance cassette in conjunction with the wild type gene were detected in the presence of antibiotic selection pressure (Ohtani et al., 2010). Likewise, the *Synechococcus* sp. WH7803 interposon mutants investigated (see section 3.3.2) did not completely segregate and retained a copy of the mutant and wild type gene (see Fig. 3.7 and 3.8). Both *Thermus thermophilus* HB8 and *Synechocystis* sp. PCC 6803, as well as *Synechococcus* sp. WH7803, possess more than one copy of their chromosome, suggesting that ploidy status may play an important role in genomic flexibility for the maintenance of mutants and, potentially be of importance in mechanisms of cyanophage resistance.

The next chapter describes work to confirm the ploidy status of *Synechococcus* sp. WH7803 and begins to assess the role ploidy status may play in conferring cyanophage resistance in marine *Synechococcus*.

Chapter 5

PLOIDY STATUS OF MARINE *SYNECHOCOCCUS* SPP.: IMPLICATIONS FOR THE GENOTYPIC BASIS OF CYANOPHAGE RESISTANCE

5.1 Introduction

It is well known that eukaryotic genomes are dynamic, varying in content and structure between and within lineages but also throughout life cycles. This dynamism also encompasses the distinction of somatic and inherited DNA by epigenetic mechanisms which help maintain genome integrity between generations. Eukaryotes have a complex genome architecture that includes features such as the unusual arrangement of rDNA genes in extrachromosomal and subtelomeric regions, the hypervariable karyotype, i.e. variation in the number and length of chromosomes, genome duality, i.e. two distinct types of genome within an organism. For example, studies in ciliates have determined that they possess two different types of nuclei, one with most of the genes (a macronuclear genome) and another involved in conjugation (a micronuclear genome) (McGrath and Katz, 2004; Parfrey et al., 2008, 2010).

The term "haploid" refers to a single copy of the chromosome, "diploid" in cases where the chromosome has been duplicated or "polyploid" where there are more than two copies of the same chromosome in the cell.

Extensive research in plants and microbial eukaryotes has shown that high ploidy levels are common for different organisms, with ~70 % of angiosperms being polyploid, and its evolutionary importance has been widely discussed (Masterson, 1994; McGrath and Katz, 2004; Otto and Whitton, 2000). However, whilst cell size and membrane properties have been attributed to ploidy levels (Tsukaya, 2013), little is known about how ploidy levels are regulated.

In prokaryotes the story is completely different, as it is widely accepted that most organisms possess a single chromosome, or chromosomes in the case of vibrios (Okada et al., 2005), and are thus "monoploid" since the term haploid only refers to organisms with the capacity for chromosome duplication (in this work, the terms ploidy and chromosome copy number are used indistinctively). Even so, almost four decades ago Hansen (1978) discovered that the radiation-resistant bacterium *Deinococcus radiodurans* (formerly *Micrococcus radiodurans*) possessed more than one chromosome copy and subsequently several other prokaryotes have been found to have multiples copies of the chromosome (Table 5.1), classifying them as either

monoploid (1 copy), oligoploid (between 2 and 9 copies) or polyploid (10 or more copies of the chromosome) (Griese et al., 2011).

Organism	Number of chromosomes	Ploidy level	References
Anabaena cylindrica	25	Polyploid	Simon (1977)
Anabaena sp. PCC7120	8	Oligoploid	Hu et al. (2007)
Anabaena variabilis	5-8	Oligoploid	Simon (1980)
Azotobacter vinelandii	1-40	Mono-polyploid	Maldonado et al. (1994)
Bacillus subtilis	1	Monoploid	Webb et al. (1998)
Blatta orientalis	323-353	Polyploid	López-Sánchez et al. (2008)
Borrelia hermsii	4-16	Oligo-polyploid	Kitten & Barbour (1992)
Buchnera aphidicola	50-200	Polyploid	Komaki & Ishikawa (2000)
Caulobacter crescentus	2	Monoploid	Pecoraro et al. (2011)
Deinococcus radiodurans	4-8	Oligoploid	Hansen (1978)
Desulfovibrio gigas	9, 17	Polyploid	Postgate et al. (1984)
Desulfovibrio vulgaris	4	Oligoploid	Postgate et al. (1984)
Epulopiscium sp. type B	50,000-120,000	Polyploid	Mendell et al. (2008)
Escherichia coli	1-7	Meroploid	Pecoraro et al. (2011); Bremer & Dennis (1996)
Lactococcus lactis spp.	1-2	Mono-diploid	Michelsen et al. (2010)
Microcystis spp.	Up to 10	Oligoploid	Kurmayer & Kutzenberger (2003)
Neisseria gonorrhoeae	3	Oligoploid	Tobiason & Seifert (2006)

Neisseria lactamica	1	Monoploid	Tobiason & Seifert (2010)
Neisseria meningitidis	4	Diploid	Tobiason & Seifert (2010)
Periplaneta americana	10-18	Polyploid	López-Sánchez et al. (2008)
Prochlorococcus spp.	1	Monoploid	Vaulot et al. (1995)
Pseudomonas putida	14-20	Meroploid	Pecoraro et al. (2011)
Synechococcus elongatus sp. PCC7942	3-6	Oligoploid	Griese et al. (2011); Mori et al. (1996); Jain et al. (2012)
Synechococcus sp. PCC6301	1-6	Mono-oligoploid	Binder & Chisholm (1990)
Synechococcus sp. WH7803	2-4	Oligoploid	Binder & Chisholm (1995); Griese et al. (2011)
Synechococcus sp. WH7805	1	Monoploid	Binder & Chisholm (1995)
Synechococcus sp. WH8101	1	Monoploid	Armbrust et al. (1989)
Synechococcus sp. WH8103	1-2	Monoploid	Binder & Chisholm (1995)
Synechocystis sp. PCC6803	12-218	Polyploid	Griese et al. (2011); Labarre et al. (1989)
Thermus thermophiles HB8	4-5	Oligoploid	Ohtani et al. (2010)
Wolinella succinogenes	1	Monoploid	Pecoraro et al. (2011)

Table 5.1. Chromosome copy number and ploidy status of bacteria.

Work on *E. coli* has shown that there is a high DNA content per cell in fast growing cultures (doubling time of ~20 min) due to the replication time being longer than the doubling time (Bremer and Dennis, 1996; Cooper and Helmstetter, 1968). Hence, cells divide before the chromosomes have time to finish the replication process. In this case the organism is called meroploid because the cells appear to have several copies, but the organism has one copy of the chromosome since these copies are still attached by the terminus. Table 5.1 summarises all bacteria that the chromosome copy number and ploidy status have been tested directly or indirectly, including the few cases of meroploidy.

Similarly, multiple copies of the chromosome have also been detected in Archaea. Interestingly, there seems to be a difference between Euryarchaeota and Crenarchaeota, in which members of Euryarchaeota possess multiple copies of the chromosome with no monoploid representatives, whilst the Chrenarchaeota contain only monoploid species. Likewise, all methanogenic Archaea investigated so far possess a range of chromosome copy numbers, but none of them are monoploid, in agreement with the previous statement since methanogenic Archaea belong to the Euryarchaeota phylum (Hildenbrand et al., 2011).

Recent work on halophilic Archaea members of the Euryarchaeota, found that they possess multiple copies of the chromosome and that the number of copies decreases in *Halobacterium salinarum* and *Haloferax volcanii* when they enter into stationary phase (Breuert et al., 2006).

In contrast, in the monoploid hyperthermophilic acidophile *Sulfolobus*, a member of the Chrenarchaeota, changes in chromosome copy number are related to both temperature and growth state (Hildenbrand et al., 2011; Hjort and Bernander, 1999).

Zerulla et al. (2014) found that the number of *H. volcanii* chromosomes decreases during growth in the absence of phosphate. The number of chromosome copies can be restored by adding phosphate back to the medium, suggesting that DNA is being used as a source of phosphate as well as other unknown genetic benefits of polyploidy (for review see Zerulla & Soppa 2014).

In cyanobacteria, chromosome copy number has been determined using three different methods, all based on establishing the amount of DNA per cell: i) flow cytometry, though studies have rather focused on understanding the cyanobacterial cell cycle rather than ploidy status *per se*, ii) spectrophotometric and/or colorimetric quantification of extracted DNA and iii) qPCR. The first two methods are imprecise since the amount of DNA per cell depends on the use of controls with a similar genome/cell size or relies on a similar efficiency of DNA extraction. Griese et al., (2011) determined chromosome copy number in cyanobacteria using a qPCR approach previously designed by Pecoraro et al. (2011). Unfortunately, specific details of these methods were not mentioned in either publication, hence it is not possible to replicate their methodologies.

Methodological differences can be considerable, since in the freshwater cyanobacterium *Synechocystis* sp. PCC 6803, the number of chromosomes was calculated to be 12, on average, using a DNA extraction method (Labarre et al., 1989) whilst using qPCR the chromosome copy number increased to 218 (Table 5.1) during the exponential phase of growth (Griese et al., 2011). It is worth pointing out though, that the ploidy status of *Synechocystis* sp. PCC 6803 was assessed not only by different methods but also with 22 years between the studies and in two different laboratories. Thus, whether the qPCR approach is more precise than DNA extraction or merely that the strain has 'changed' in ploidy status over time and between laboratories is arguable.

Overall, ploidy in cyanobacteria has been found to vary between genera and species (see Table 5.1). From all the strains studied so far a clear difference between freshwater and marine cyanobacteria can be observed, with freshwater representatives being poly- and oligoploid and marine strains monoploid. However, the exception is marine *Synechococcus* sp. WH7803 that seems to have between 2 to 4 copies of the chromosome, determined both by flow cytometry and qPCR (Binder and Chisholm, 1995; Griese et al., 2011).

In this work, an in-depth study of the oligoploid strain *Synechococcus* sp. WH7803 was performed with the intention of understanding the mechanisms underlying regulation of ploidy in marine *Synechococcus* and its relationship with cyanophage resistance.

5.2 Objective

The aim of this study was to optimise a qPCR-based approach to assess chromosome copy number in marine *Synechococcus* and to determine if ploidy level varies during growth and under P deplete conditions in *Synechococcus* sp. WH7803.

5.3 Results

5.3.1 Optimisation of methods for assessment of chromosome copy number in *Synechococcus* sp. WH7803

The chromosome copy number of *Synechococcus* sp. WH7803 was determined using two independent methods: qPCR (section 2.15) and a fluorometric quantification of total genomic DNA (section 2.16).

A first approach to elucidate chromosome copy number in *Synechococcus* sp. WH7803 was a fluorometric quantification of genomic DNA, which gave an average of 6.4±1.8 copies of the chromosome per cell (Table 5.2).

Experiment	Chromosome copy number
1	7.9±0.49
2	7.9±1.23
3	5.2±0.83
4	4.6±0.60

Table 5.2. Chromosome copy number of *Synechococcus* sp. WH7803 quantified by a fluorometric method.

Synechococcus sp. WH7803 genomic DNA per cell was quantified by a fluorometric method (section 2.16). Chromosome copy number is presented as an average of three technical replicates ± standard deviation.

The extraction of genomic DNA for this experiment was performed using a lysis buffer (section 2.13), which maintains a good integrity of genomic DNA (Fig. 5.1.A) but does not assure a replicable complete extraction of nucleic acids due to partial lysis of *Synechococcus* cells (Fig. 5.1.B). Therefore, a more accurate approach was performed by modifying the qPCR method designed by Pecoraro et al. (2011). In this case, genomic DNA was obtained by mechanical lysis of *Synechococcus* cells (Fig. 5.1) since high DNA integrity is not an absolute prerequisite when using qPCR.

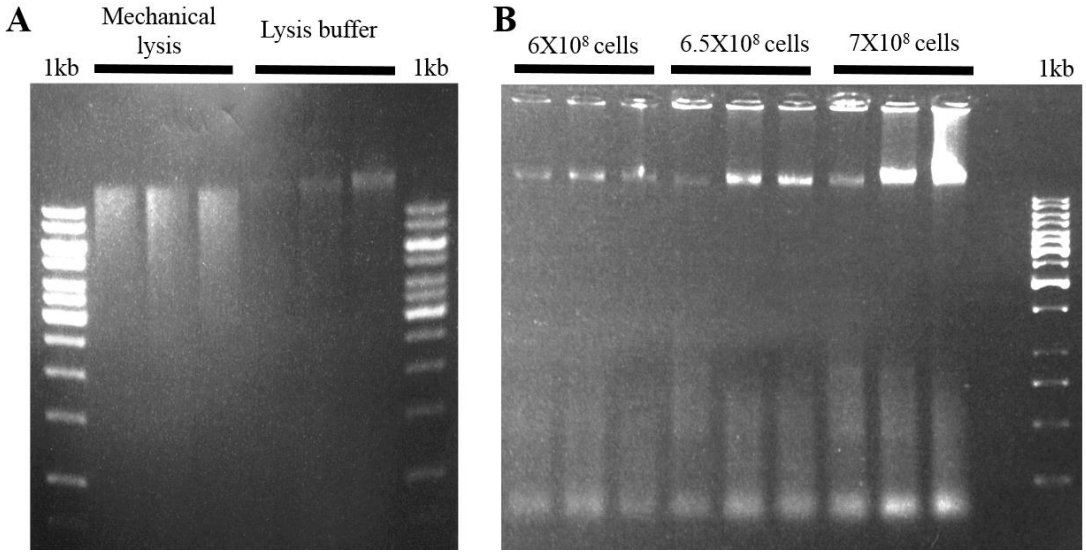


Figure 5.1. *Synechococcus* sp. WH7803 genomic DNA extraction.

Extracted DNA was separated using 1% (w/v) agarose gel electrophoresis stained with ethidium bromide for visualisation of (A) comparison of genomic DNA extraction using lysis buffer and mechanical lysis. (B) Genomic DNA extraction using lysis buffer of sorted cells. All experiments were run in triplicate. 1kb DNA ladder (Invitrogen).

Optimisation of the qPCR assay required determination of the detection limit, DNA extraction conditions and primer selection. The detection limit for this qPCR assay was 1,000 cells (Fig. 5.2.B), which was determined by flow cytometry of sorted *Synechococcus* sp. WH7803 cells (see section 2.15). Also, another step of optimisation was required because the salt concentration of ASW medium inhibited the qPCR

reaction, detecting an order of magnitude less material (Fig. 5.2.B). In order to avoid salt inhibition of the qPCR, different methods were tested for desalting the DNA samples before assay. The most efficient method, i.e. requiring the least amount of time for desalting and without losing DNA, was precipitation by centrifugation, re-suspension with nuclease-free water and mechanical lysis of cells (see section 2.15 and Fig. 5.2.B). Additionally, primer selection for the qPCR assay for *Synechococcus* sp. WH7803 and WH7805 was based both on their specificity to each strain and also that the same amplification efficiency for each primer set was obtained (Fig. 5.2.A).

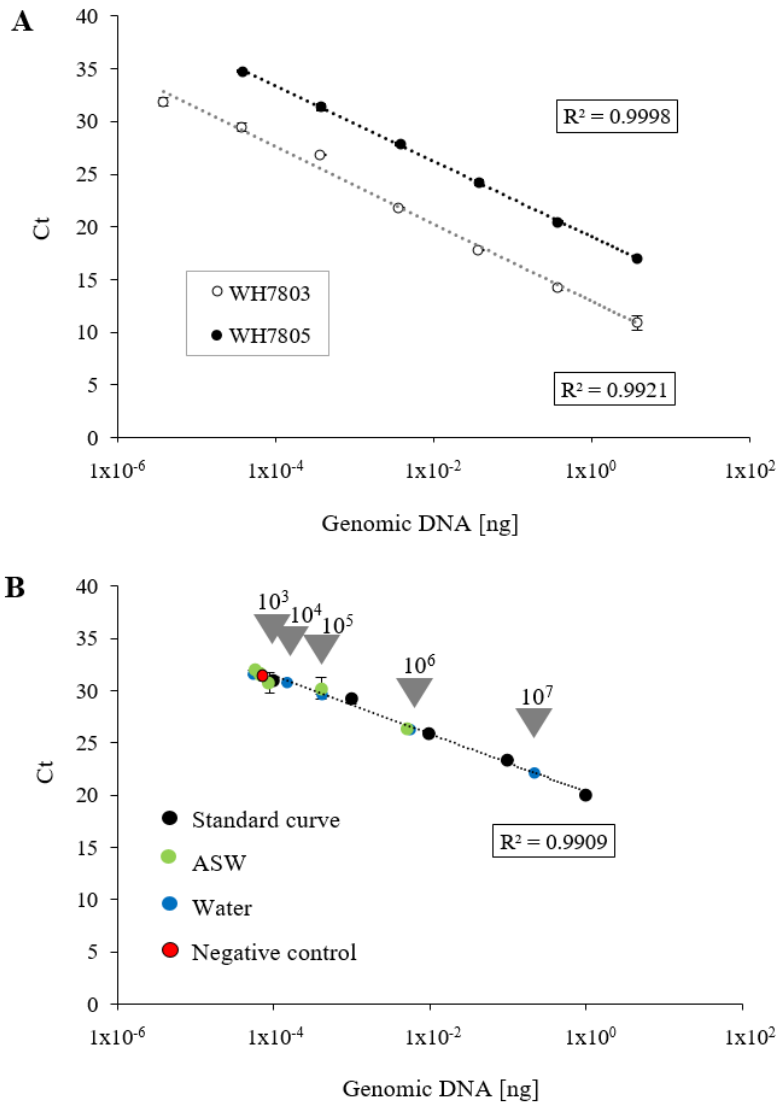


Figure 5.2. qPCR optimisation for quantification of chromosome copy number in *Synechococcus* sp. WH7803 and WH7805.

(A) Standard curves constructed with *Synechococcus* sp. WH7803 and WH7805 genomic DNA and primer sets listed in Table 2.10. (B) 10^0-10^7 sorted *Synechococcus* sp. WH7803 cells (grey arrows) were used for qPCR (see section 2.16) with ASW medium (green) or washed with nuclease free water (blue) to remove excess salt. The detection limit was determined to be 10^3 cells. *Synechococcus* sp. WH7803 genomic DNA standard curve (black); qPCR negative control (red), showing non-specific amplification.

5.3.2 Determination of chromosome copy number in *Synechococcus* sp. WH7805

Due to its close phylogenetic proximity with *Synechococcus* sp. WH7803 (see Fig. 6.1) chromosome copy number was also determined for *Synechococcus* sp. WH7805. This strain was previously suggested to be monoploid based on flow cytometry, and in comparison to fluorescence characteristics and DNA content of *Synechococcus* sp. WH7803 (Binder and Chisholm, 1995).

Chromosome copy number was determined using both *Synechococcus* sp. WH7805 cells alone or with an equimolar amount of *Synechococcus* sp. WH7803 cells (see section 2.15) to assess the reproducibility of the nucleic acid extraction process. Both experiments gave the same results, with an average of 1.0±0.5 copies of the *Synechococcus* sp. WH7805 chromosome per cell but 4.1±2.9 copies for *Synechococcus* sp. WH7803 (Table 5.3), confirming that *Synechococcus* sp. WH7805 is a monoploid strain and that the assay is accurate in determining ploidy level.

Strain	Chromosome copy number
Synechococcus sp. WH7805	0.9±0.6
	1.2±0.4
Synechococcus sp. WH7803	4.1±3.3
	4.0±2.5

Table 5.3. *Synechococcus* spp. chromosome copy number quantified by qPCR.

The qPCR method used for quantifying chromosome copy number showed a clear difference in the average number of chromosomes per cell between *Synechococcus* sp. WH7803 and *Synechococcus* sp. WH7805. Again, since this is an average of biological replicates, the standard deviation between replicates is high, but this variation is higher

for *Synechococcus* sp. WH7803 (CV = 0.88) compared to *Synechococcus* sp. WH7805 (CV = 0.55). This can be explained as a function of the number of chromosome copies found per cell, since for a monoploid strain the only variation will be related to dividing cells, i.e. incomplete separation of the replicating chromosomes, but for an oligoploid strain like *Synechococcus* sp. WH7803 the variation indicates either a differential regulation of chromosomal replication, incomplete separation of replicating chromosomes, or both.

Given the higher chromosome copy number in *Synechococcus* sp. WH7803 compared to *Synechococcus* sp. WH7805 this could suggest an increase in cell size in the former strain. Cell size was determined by flow cytometry with a mix of cells from the exponential growth phase of both organisms stained with SYBR Green I (see section 2.3). Forward scatter as a proxy for cell size showed cells of both cultures were the same size (Fig. 5.3.A) whereas a shift in fluorescence was observed (Fig. 5.3.B), indicating that the DNA content was higher in *Synechococcus* sp. WH7803 but cell size was the same for both strains.

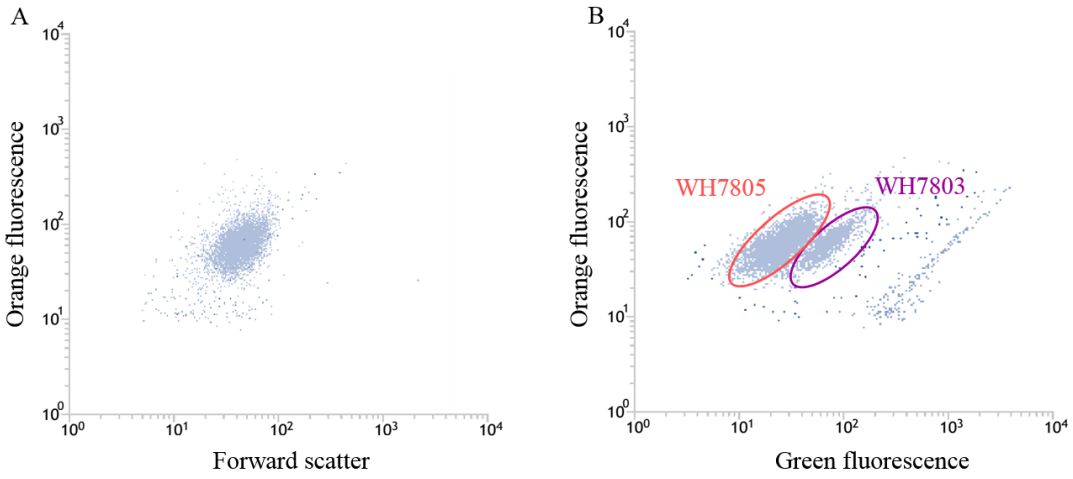


Figure 5.3. Cytogram of a mixed culture of *Synechococcus* sp. WH7805 and *Synechococcus* sp. WH7803 stained with SYBR green I.

(A) Orange fluorescence (580 nm), for detection of phycoerythrin, versus forward scatter. (B) Orange fluorescence versus green fluorescence (530 nm), for detection of SYBR green I DNA staining. *Synechococcus* sp. WH7803 (purple) and *Synechococcus* sp. WH7805 (red) are circled.

5.3.3 Determination of the chromosome copy number of *Synechococcus* sp. WH7803 throughout the growth phase and under P limitation

In order to investigate whether chromosome copy number changes during the growth of *Synechococcus* sp. WH7803, a growth curve was performed with quantification of chromosome copy number during three different growth phases: lag, mid exponential and late exponential phase (see Fig. 5.4). Analysis of individual replicates (Fig. 5.4.B) revealed that although chromosome copy number remained stable during growth, there was a large variation between replicates, explaining the large error bars associated with Fig. 5.4.C.

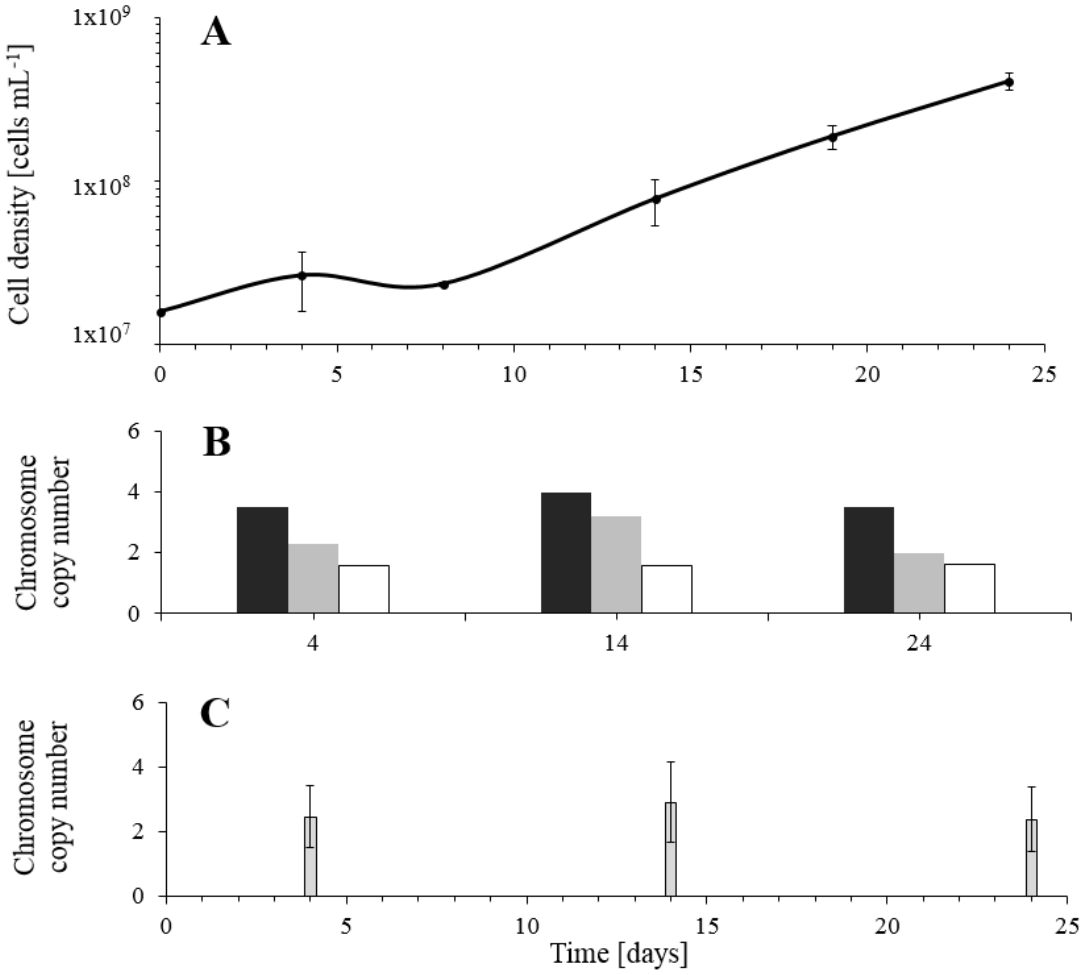


Figure 5.4. Determination of chromosome copy number during growth of *Synechococcus* sp. WH7803.

(A) Growth curve of *Synechococcus* sp. WH7803. (B) Chromosome copy number of three biological replicates measured at three different time points. (C) The average chromosome copy number of the three biological replicates (and three technical replicates) during growth. Error bars represent standard deviation.

During this experiment, growth of three biological replicates was monitored every four days and genome number assessed during lag, mid exponential and late exponential phase (Fig. 5.4). Ploidy level is maintained during growth of the culture, with a mean of 2.6±1.0 copies per cell, although the level is lower than measured in previous experiments, with only one of the replicates reaching 4 copies per cell. This experiment also showed a high variation between biological replicates (CV = 0.4), observed in

Fig. 5.4.B, but very low within technical replicates (CV = 0.08, 0.26 and 0.03 respectively). In conclusion, the chromosome copy number per cell is maintained during growth but appears to show a fairly wide variation between replicate cultures. In order to provide further clues into the underlying regulation of this process we also assessed the role of phosphorus (P) availability in this process, testing the idea that P limitation would reduce chromosome copy number.

In this respect, *Synechococcus* sp. WH7803 was grown under P deplete conditions (see section 2.1.1. and Table 2.1), causing a decrease in growth rate ($\mu = 0.07 \pm 0.003\ d^{-1}$) compared to standard growth conditions ($\mu = 0.18 \pm 0.01\ d^{-1}$). Chromosome copy number under P-deplete conditions was highly variable, with between 1.5 - 18 copies detected. This experiment showed inconsistency between biological replicates, but also during growth (Fig. 5.5).

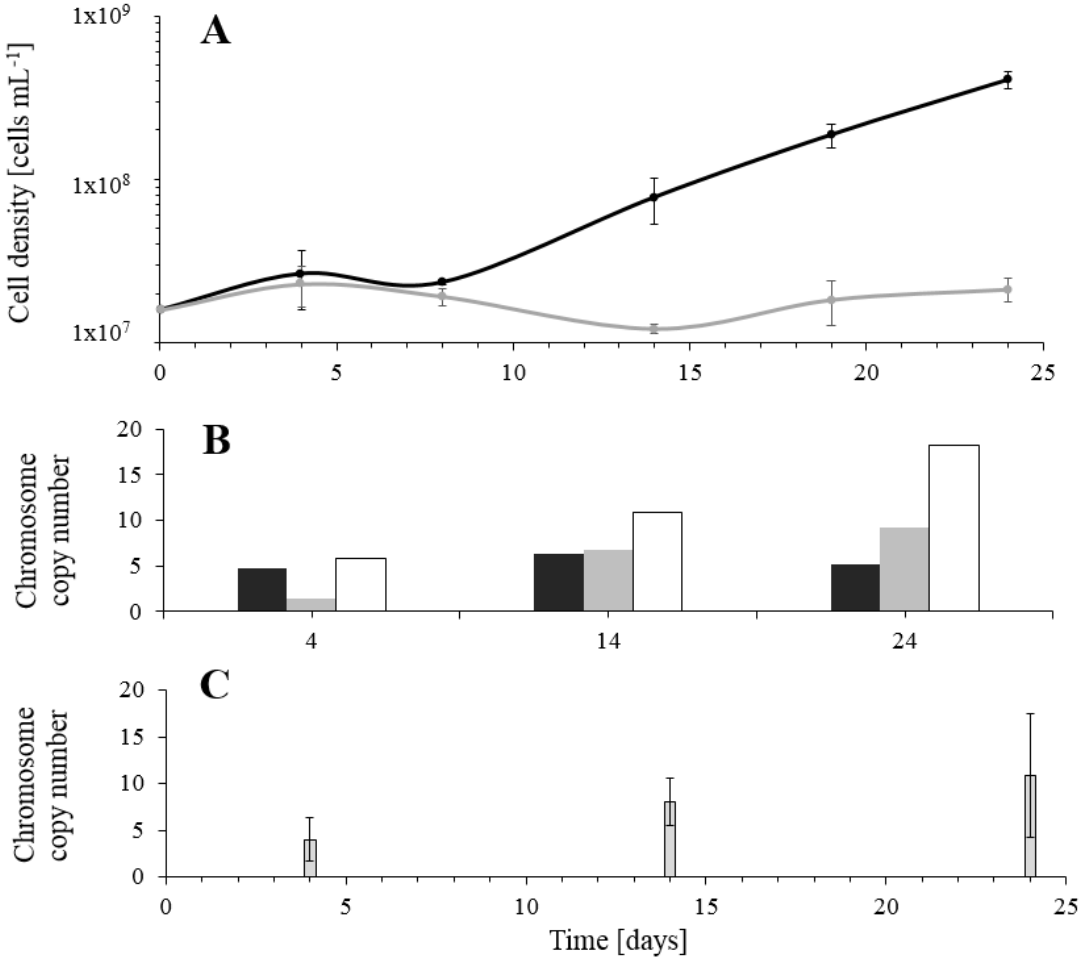


Figure 5.5. Determination of chromosome copy number during growth of *Synechococcus* sp. WH7803 under P replete and deplete conditions.

(A) Growth curve of *Synechococcus* sp. WH7803 grown in P replete (black) and deplete (grey) medium. (B) Chromosome copy number of three biological replicates grown under P deplete conditions, measured at three different time points. (C) The average chromosome copy number of the three biological replicates (and three technical replicates) during growth in P deplete medium. Error bars represent standard deviation.

In Fig. 5.5.B the chromosome copy number of three biological replicates is shown. Even though there appears to be an increase in chromosome copy number during prolonged P limitation of the culture, the results are inconclusive due to high variation of the replicates and no apparent growth of the culture (Fig. 5.5.A). Thus, as mentioned above when measuring chromosome copy number during the growth phase, where

variation between biological replicates was high (CV = 0.5), during P limitation variation within replicates was also high (CV = 0.15, 0.69, 0.53 respectively), potentially suggesting high cell variability in ploidy level.

The data was assessed using a 3 x 2 repeated measures mixed analysis of variance (ANOVA). The within replicates factor had three levels (times 4, 14 and 24) and the between replicates factor had two levels (P-deplete and replete). A non-significant result was found regarding chromosome copy variation related to growth of *Synechococcus* sp. WH7803 ($F(2,8) = 3.56$, $p = 0.078$) and this was also the case for the P-deplete and P-replete conditions ($F(1,4) = 5.83$, $p = 0.073$). Additionally, a non-significant interaction was found ($F(2,8) = 3.65$, $p = 0.075$).

5.4 Discussion

Synechococcus sp. WH7803 was confirmed to possess approximately four copies of the chromosome per cell in this study. This is in agreement with previous studies (Binder and Chisholm 1995, Griese et al., 2011), suggesting that ploidy status is preserved over time (as it is ~20 years since the first study) and between culture conditions (the Binder and Chisholm (1995) study used SN medium for culturing *Synechococcus*).

It is important to remember that copy number estimates represent an average, such that in theory, in individual cells the number of chromosome copies could be even more variable. This could explain the variation seen with fluorometric quantification of DNA (Table 5.2) but also to a lower extent with the qPCR assay (Table 5.3).

Although regulation of chromosome replication in polyploid bacteria has not yet been studied, the segregation of chromosomes has been investigated in freshwater cyanobacteria. In *Synechocystis* sp. PCC 6803, that possesses up to 218 chromosome copies (see Table 5.1), segregation occurs late during the cell cycle and it is in a random and passive way, most likely due to division of the cell by the septum (Schneider et al., 2007). Likewise, work performed on the freshwater cyanobacterium *Synechococcus elongatus* PCC 7942 indicates that the chromosomes arrange along the cell in a uniform way whereas their replication is asynchronous and not related to each

location (Jain et al., 2012). This evidence could also explain the variation mentioned above, since *Synechococcus* sp. WH7803 chromosomes could be replicating at different times producing cells with different numbers of chromosomes. If that is the case, then *Synechococcus* sp. WH7803 could be demonstrating differential regulation of the replication of each chromosome or it could be a process of pseudo-mero-oligoploidy of some chromosomes, i.e. an apparent higher number of chromosomes due to asynchronous replication or incomplete division of some of them, similar to what has been observed in monoploid bacteria such as *E. coli* and *P. putida* (Table 5.1).

Moreover, variation in the number of chromosome copies observed in the Archaea *H. salinarum* and *H. volcanii* and the freshwater cyanobacterium *Synechocystis* sp. PCC 6803 shows that there is a reduction in ploidy when cells enter stationary phase, indicating regulation related to the growth state of the culture (Breuert et al., 2006; Griese et al., 2011). In contrast, no significant change in chromosome copy number was observed throughout the *Synechococcus* sp. WH7803 growth curve (Fig. 5.4).

Likewise, *Synechococcus* sp. WH7803 grown under P-deplete conditions did not show a significant variation in the number of chromosome copies. This results contrast with what was observed by Zerulla et al. (2014) in the archaea *H. volcanii*, where the number of copies decreased in the absence of phosphate. The authors suggest that in this organism, as well as bacteria from the *Shewanella* genus (Pinchuk et al., 2008), DNA acts as a nutrient source and that they can grow with this molecule as the sole source of phosphorus, carbon and nitrogen. However, *Synechococcus* sp. WH7803 possesses an intracellular store of phosphate as polyphosphate (Mazard et al. 2012) suggesting no need for an extra source of phosphate from DNA and hence no reduction under limiting conditions.

Nonetheless, the results from this experiment to detect variation in chromosome copies under P-deplete conditions are not conclusive and only careful inferences should be made. For instance, the growth rate of *Synechococcus* sp. WH7803 under P-deplete conditions has been determined to be similar to P-replete conditions (Mazard et al. 2012). However, here (Fig. 5.5.A) a large difference in growth between these two conditions was observed, with virtually no growth under P-deplete conditions

($\mu = 0.07 \pm 0.003\ d^{-1}$), indicating that perhaps the cells are beyond being P stressed but rather are actually dying off. This could explain the variation in chromosome copy number since this could be an artefact of cells lysing and releasing their DNA into the medium and hence more DNA per cell would be calculated.

So far, there is no information regarding the ploidy status of *Synechococcus* sp. WH7803 grown under other environmental conditions, i.e. apart from P limitation. Moreover, to date, *Synechococcus* sp. WH7803 is the only known oligoploid marine cyanobacterium (Table 5.1), though from a limited number of strains analysed (Binder and Chisholm, 1995; Griese et al., 2011). Thus, the extent of oligoploid *Synechococcus*, and indeed other cyanobacteria in natural marine waters, is unknown. Clearly though, the phylogenetically closely related monoploid strain *Synechococcus* sp. WH7805 represents an excellent 'sister' organism to compare with WH7803, work which should bring new insights into how multiple copies of the chromosome are maintained and regulated.

The presence of multiple chromosomes in *Synechococcus* sp. WH7803 may also explain the variation in SNP frequency identified in chapter 4, in the genomes of the cyanophage resistant mutants obtained here and by Marston et al. (2012), though the conclusions of the latter study are insufficient given their assignment of mutations in specific *Synechococcus* sp. WH7803 genes despite their being SNPs in many different genes (see Chapter 4). How oligoploidy might play a role in cyanophage resistance is currently unclear. It might be that this organism is able to differentially regulate the expression of genes and/or replication of chromosomes containing genes with mutations conferring resistance to cyanophage infection. Alternatively, potentially dominant or recessive alleles exist in this organism, a scenario not known for prokaryotes to the best of my knowledge. Certainly, further experiments are required to resolve this fascinating aspect of *Synechococcus* biology.

Relating ploidy status back to my own work on cyanophage resistance mechanisms, clearly there is likely a difference in ploidy level between *Synechococcus* sp. WH7803 and *Prochlorococcus* spp. (the latter is suggested to be monoploid, see Vaulot et al. 1995). This is consistent with the WGS data obtained from cyanophage resistant mutants of *Synechococcus* sp. WH7803 (this work, chapter 4) and the conclusions of

Avrani et al., (2011) where the number of mutations or SNPs found in *Synechococcus* sp. WH7803 greatly exceeds those found in *Prochlorococcus* (the latter is in accordance with the single mutation/phenotype scenario). Therefore, experiments with the monoploid *Synechococcus* sp. WH7805 are imperative to understand this phenomenon.

Chapter 6

GENERATION AND CHARACTERISATION OF *SYNECHOCOCCUS* SP. WH7805 CYANOPHAGE RESISTANT MUTANTS

6.1. Introduction

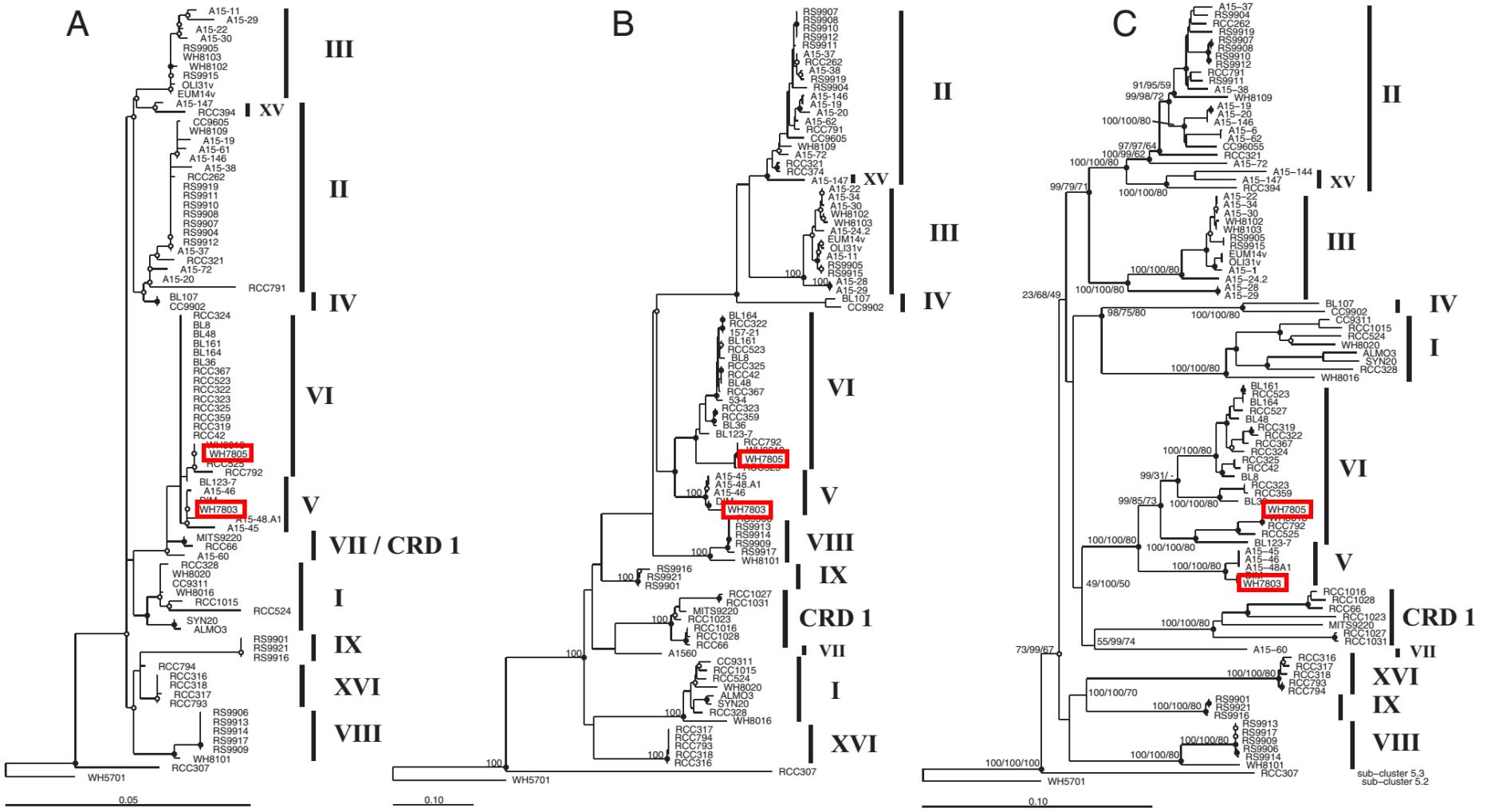
I have previously discussed in this thesis (Chapters 4 and 5) the possible role that ploidy might play in cyanophage resistance in *Synechococcus* sp. WH7803. Whilst the extent of this phenomenon is not clear, i.e. whether ploidy as a mechanism for cyanophage resistance is inherent to other members of the marine *Synechococcus* genus, to more directly assess mechanisms of cyanophage resistance in marine *Synechococcus*, and to disentangle the effects of ploidy from such resistance mechanisms, we focused on the monoploid strain *Synechococcus* sp. WH7805. Such a monoploid status of this strain was demonstrated using a quantitative PCR assay developed in this work (see Section 5.3.2), confirming earlier cell cycle analyses in this organism (Binder and Chisholm, 1995).

Unlike *Synechococcus* sp. WH7803, which has been used as a model strain for various aspects of marine *Synechococcus* biology (Liu et al., 1999; Marston et al., 2012; Scanlan, 2003; Traving et al., 2014), *Synechococcus* sp. WH7805 is much less studied. However, both strains were isolated in 1978 from the same location (33°45'N, 67°30'W) in the Sargasso Sea at 25m depth (Waterbury et al., 1986) and are closely related phylogenetically (Fuller et al., 2003; Urbach et al., 1998). Such close taxonomic relatedness is evidenced using several different genetic markers, e.g. the 16S rRNA gene (Fuller et. al., 2003; see Fig. 6.1) the high resolution marker *petB* (Mazard et al. 2012), and whole genome sequence analyses (Dufresne et al., 2008). These two *Synechococcus* strains belong to the same sub-cluster, 5.1B, but to different clades: *Synechococcus* sp. WH7803 (clade V) and *Synechococcus* sp. WH7805 (clade VI). Moreover, the two strains encompass different pigment types, with *Synechococcus* sp. WH7805 lacking PUB (pigment type 2) and *Synechococcus* sp. WH7803 possessing both PUB and PEB but with a low PUB:PEB ratio (< 0.4, pigment type 3a) (Six et al. 2007; Scanlan et al. 2009; see also Table 2.2).

As mentioned above the genomes of *Synechococcus* sp. WH7803 and WH7805 have been sequenced (Dufresne et al., 2008), although only for *Synechococcus* sp. WH7803 has it been closed, with 2 contigs for *Synechococcus* sp. WH7805. The genomes have similar size: 2.4 Mbp in *Synechococcus* sp. WH7803 and 2.6 Mbp in *Synechococcus* sp. WH7805 (Dufresne et al., 2008) but importantly they differ in their ploidy status,

with *Synechococcus* sp. WH7803 being oligoploid compared to the monoploid *Synechococcus* sp. WH7805 (see Chapter 5).

Due to the high similarity of these two *Synechococcus* strains and their different chromosome copy numbers makes *Synechococcus* sp. WH7805 an ideal model to examine mechanisms of cyanophage resistance without inherent problems associated with ploidy.


A
B
C
III
II
XV
IV
VI
V
VII / CRD 1
I
IX
XVI
VIII
CRD 1
VII
WH7805
WH5701
RCC307
0.05
0.10
sub−cluster 5.3
sub−cluster 5.2

(Previous page)

Figure 6.1. Phylogeny of marine *Synechococcus* based on (A) 16S rRNA, (B) 16S-23S ITS and (C) seven concatenated core genes (Multi locus sequence alignment) nucleotide sequences (from Mazard et al. 2012)

Phylogenetic trees constructed using neighbour joining, maximum likelihood and maximum parsimony with bootstrap values > 90 % (closed circles) or > 70 % (60 % for the 16S rRNA gene; open circles). The scale bar represents the number of substitutions per nucleotide position. *Synechococcus* sp. WH7803 and WH7805 are highlighted in red.

6.2. Objective

To obtain and characterise cyanophage resistant mutants of the monoploid strain *Synechococcus* sp. WH7805. Characterisation would involve phenotypic and genotypic (WGS) analysis of the derived mutant strains and comparison of these *Synechococcus* sp. WH7805 cyanophage resistant mutants with those obtained in *Synechococcus* sp. WH7803 (see Chapter 4).

6.3. Results

6.3.1. Selection of the cyanophage to be used for obtaining cyanophage resistant *Synechococcus* sp. WH7805 mutants

Cyanophage infection of *Synechococcus* sp. WH7805 has not previously been studied. In order to determine the infection of *Synechococcus* sp. WH7805, a range of different cyanophages was selected (Table 2.5), including the two cyanophages used to obtain the *Synechococcus* sp. WH7803 cyanophage resistant mutants (cyanophages S-PM2 and S-RSM42). The infection of eight cyanophages was assessed in solid media and four of them in liquid media (see Sections 2.4 and 2.6; Fig. 6.2).

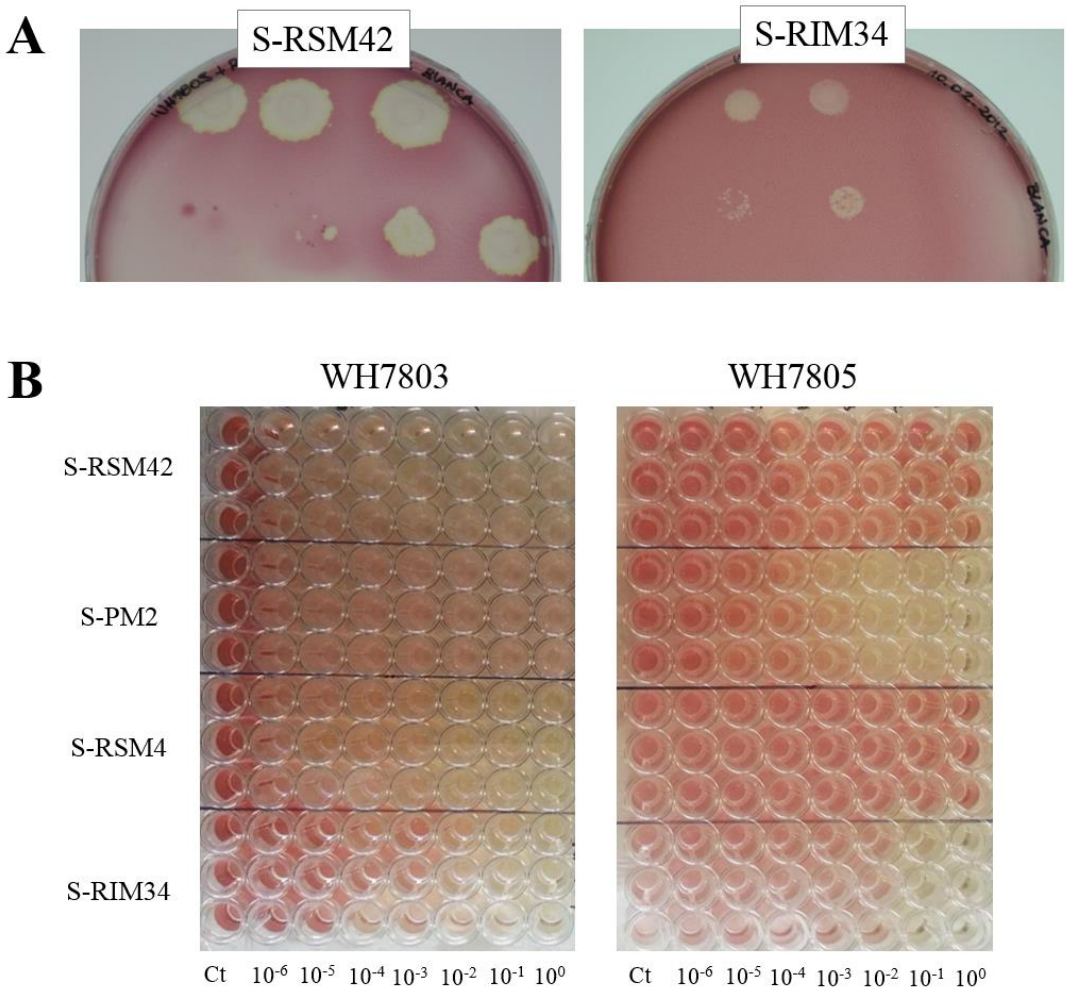


Figure 6.2. Cyanophage infection of *Synechococcus* sp. WH7805.

(A) Infection of *Synechococcus* sp. WH7805 by cyanophages S-RSM42 and S-RIM34 observed by plaques obtained with a spot assay (section 2.6). The 'wrinkly' morphology of S-RSM42 plaques indicates bacterial contamination of the phage lysate used, corroborated with contamination plates (see section 2.1.2). (B) Infection of *Synechococcus* sp. WH7805 (WH7805) by cyanophages S-RSM42, S-PM2, S-RSM4 and S-RIM34 observed by lysis of the culture. Rows indicate the 4 cyanophages tested in triplicate and columns indicate the dilution, from 10^0 to 10^{-6}. Ct: negative control of *Synechococcus* spp. without the cyanophage; WH7803: *Synechococcus* sp. WH7803, used as a positive control for infection.

Infection of *Synechococcus* sp. WH7805 in solid medium was determined by the detection of plaques with cyanophages S-RSM42 and S-RIM34 (Fig. 6.2.A) and complete lysis of the plate with cyanophages S-PM2 and S-RIM11 (no plaques were observed). Infection in liquid media was determined for cyanophages S-RIM34 and

S-PM2 (Fig. 6.2.B). A summary of infection range is given in Table 6.1, where cyanophages S-PM2 and S-RIM34 were the only ones tested for infection in both solid and liquid medium.

Cyanophage	Infection in solid medium	Infection in liquid medium
S-PM2	Yes	Yes
S-RSM4	No	No
S-RSM42	Yes	No
S-RSM57	No	N/T
S-RSM76	No	N/T
S-RIM1	No	N/T
S-RIM11	Yes	N/T
S-RIM34	Yes	Yes

Table 6.1. Cyanophages infecting *Synechococcus* sp. WH7805.

Infection of *Synechococcus* sp. WH7805 in solid medium determined by a spot assay (section 2.6) and in liquid medium by lysis of the culture. N/T = not tested.

The cyanophage selected for obtaining cyanophage resistant mutants in *Synechococcus* sp. WH7805 was S-PM2 because it can lyse in solid and liquid medium (Table 6.1 and Fig. 6.2) and also because it allows for direct comparison with the cyanophage resistant *Synechococcus* sp. WH7803 mutants (mutants R1, R2 and PHR are resistant to S-PM2, see Table 2.3). Although most of cyanophage resistant *Synechococcus* sp. WH7803 mutants are resistant to S-RSM42 (A-J, R1 and R2, see Table 2.3), unfortunately infection of *Synechococcus* sp. WH7805 in liquid medium was not observed with this cyanophage, hence the reason it was not selected for obtaining resistant mutants in this strain.

6.3.2. Phenotypic characterisation of spontaneous cyanophage resistant *Synechococcus* sp. WH7805 mutants

Ten cyanophage resistant *Synechococcus* sp. WH7805 mutants, obtained as single colonies on solid medium, were isolated following infection with cyanophage S-PM2 (see section 2.7). However, despite picking single colonies all were contaminated with heterotrophic bacteria (data not shown). The mutants JC2 and JC6 have an obvious clumping phenotype whereas mutants JC1, JC3, JC4 and JC10 clumped to a much lesser extent (Fig. 6.3). The clumping phenotype is similar to that observed in the cyanophage resistant *Synechococcus* sp. WH7803 mutants (Fig. 3.1) although no changes in pigmentation were observed (data not shown).

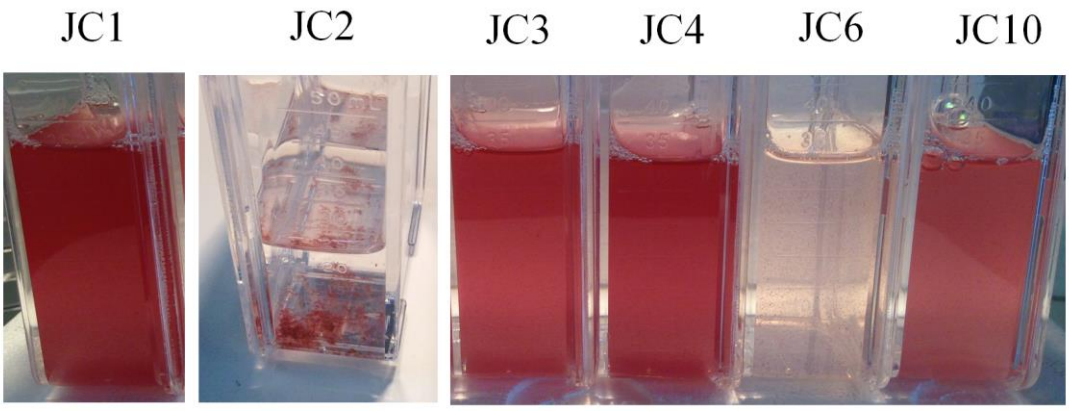
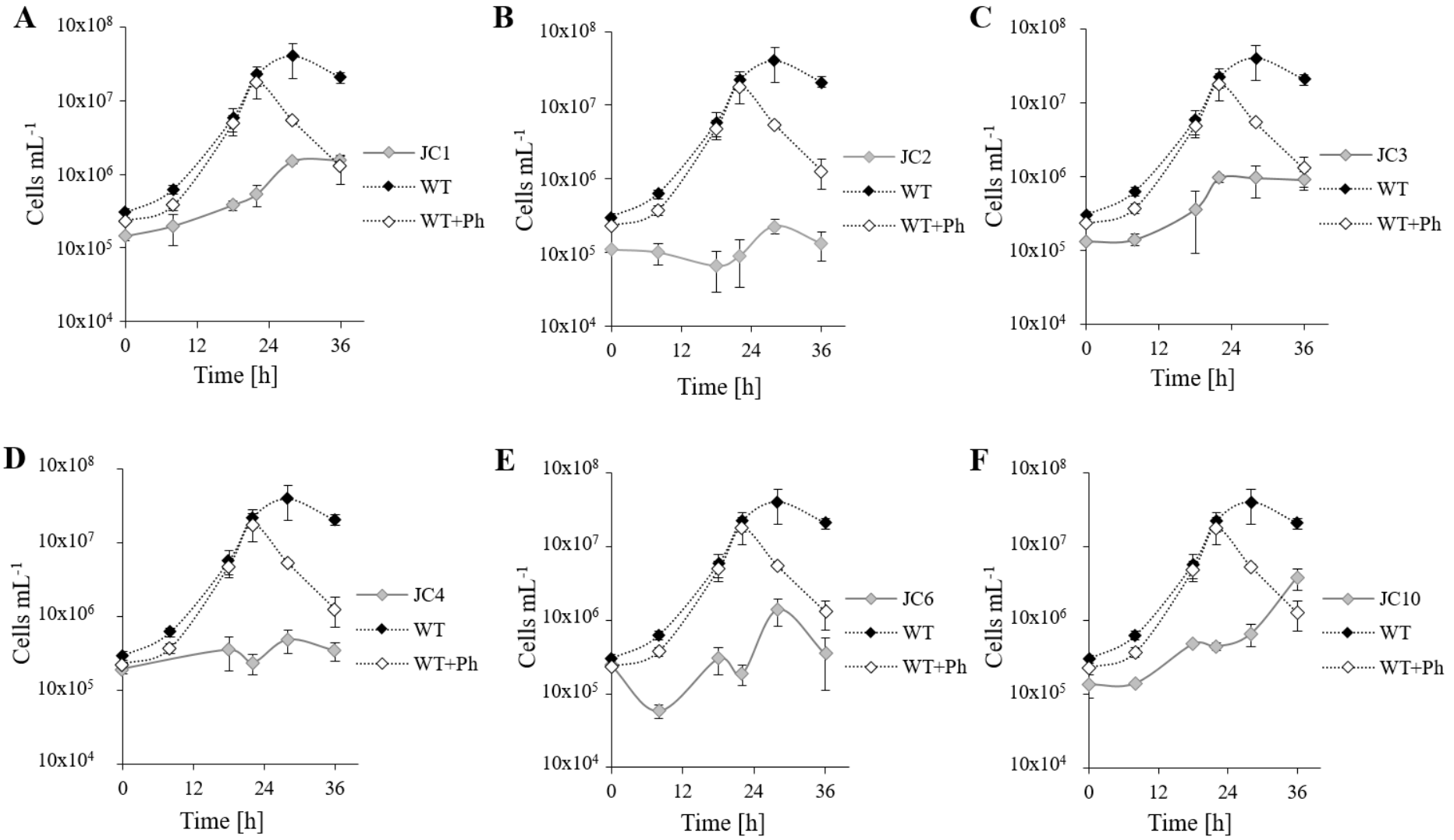


Figure 6.3. 'Clumping' phenotype of some of the 10 cyanophage resistant Synechococcus sp. WH7805 mutants isolated in this work.

Photographs were taken after 52 days of growth without agitation. Mutants JC1, JC3, JC4 and JC10 presented a less intense clumping phenotype (not observed in the photograph).

Growth rate analysis, of six of the ten cyanophage resistant *Synechococcus* sp. WH7805 mutants, was performed comparing growth with the wild type in the presence and absence of cyanophage S-PM2 (Fig. 6.4). Noteworthy, variation between biological replicates was higher in the mutants compared to the wild type, especially for mutants JC2 and JC6 (Fig. 6.4.B and E, respectively).


A
B
C
D
E
F
Cells mL-1
10x10^8
10x10^7
10x10^6
10x10^5
10x10^4
0
12
24
36
Time [h]
JC1
JC2
JC3
JC4
JC6
JC10
WT
WT+Ph

(Previous page)

Figure 6.4. Growth of cyanophage resistant *Synechococcus* sp. WH7805 mutants.

Cell density over time of cyanophage resistant *Synechococcus* sp. WH7805 mutants: JC1 (A), JC2 (B), JC3 (C), JC4 (D), JC6 (E) and JC10 (F). For each mutant growth was compared to un-infected (WT) and cyanophage S-PM2 infected *Synechococcus* sp. WH7805 (WT+Ph).

A fitness cost, compared to the WT, was observed with all the mutants studied (Table 6.2) and this growth deficiency led to the loss of mutants JC5, JC7, JC8 and JC9. For this reason, all further physiological characterisations were performed with mutants JC1, JC2, JC3, JC4, JC6 and JC10. Noteworthy, is that variation in growth between replicates is higher in the mutants compared to the wild type, especially for mutants JC2 and JC6 (CV = 0.30 and 0.35 respectively) (Table 6.2). This is likely related to the clumping phenotype of these resistant strains, and because cell number was measured using flow cytometry (see section 2.3).

Strain	μ [d^{-1}]	CV
Synechococcus sp. WH7805 JC1	0.10±0.03	0.26
Synechococcus sp. WH7805 JC2	0.13±0.04	0.30
Synechococcus sp. WH7805 JC3	0.10±0.03	0.27
Synechococcus sp. WH7805 JC6	0.12±0.04	0.35
Synechococcus sp. WH7805 JC10	0.10±0.02	0.15
Synechococcus sp. WH7805	0.20±0.01	0.03

Table.6.2. Growth rate of cyanophage resistant *Synechococcus* sp. WH7805 mutants and wild type.

Growth rate (μ) was calculated during exponential growth as an average of the 3 replicates ± standard deviation. Mutant JC4 did not reach exponential growth in this experiment, therefore the growth rate was not calculated. CV = coefficient of variation.

To study the attachment of cyanophage S-PM2 to the *Synechococcus* sp. WH7805 S-PM2 resistant mutants, an adsorption assay was performed (see section 2.8). These assays showed that all the cyanophage resistant *Synechococcus* sp. WH7805 mutants studied were resistant to adsorption by S-PM2 (Fig. 6.5), potentially suggesting that they have a similar mechanism of cyanophage resistance.

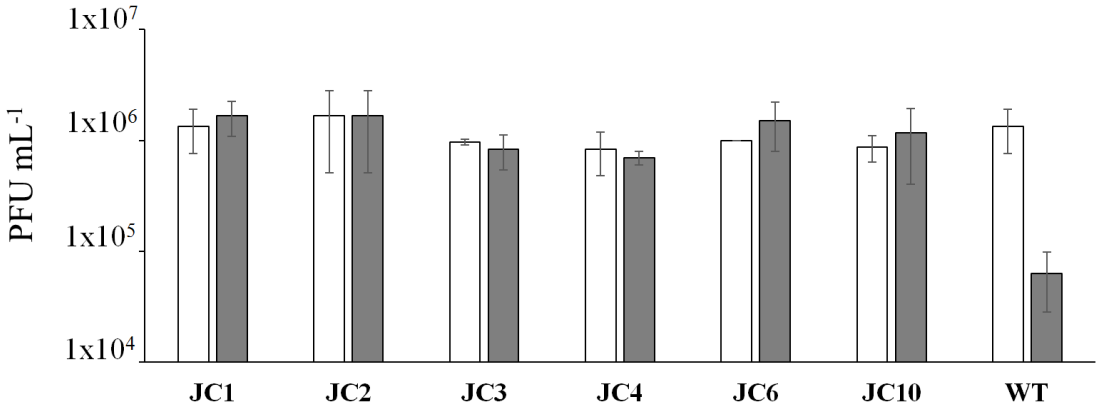


Figure 6.5. Adsorption assays of cyanophage-resistant *Synechococcus* sp. WH7805 mutants with cyanophage S-PM2.

The concentration of cyanophage S-PM2 is given at time 0 (white bars) and time 60 min (grey bars). JC1-JC10: S-PM2 resistant *Synechococcus* sp. WH7805 mutants; WT: wild type *Synechococcus* sp. WH7805. Bars are the average of three replicates with error bars as standard deviation.

LPS profiles (see section 2.9) of the cyanophage resistant *Synechococcus* sp. WH7805 mutants were also characterised. In contrast to cyanophage resistant *Synechococcus* sp. WH7803 mutants, all cyanophage resistant *Synechococcus* sp. WH7805 mutants assessed showed a different LPS profile compared to the wild type (Fig. 6.6).

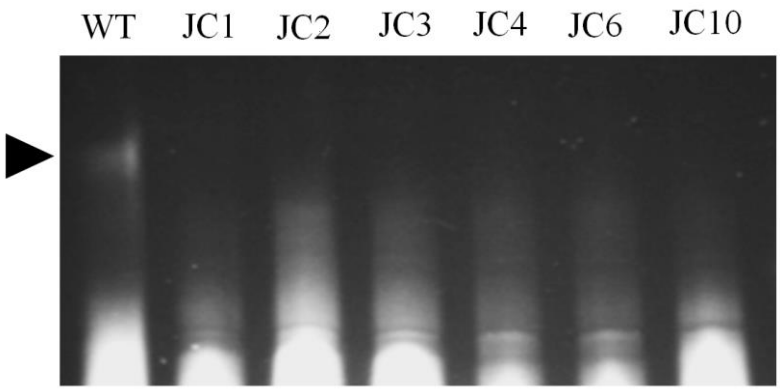


Figure 6.6. LPS profiles of the S-PM2 resistant *Synechococcus* sp. WH7805 mutants obtained using a LPS extraction kit (Intron Biotechnology).

JC1-JC10: S-PM2 resistant *Synechococcus* sp. WH7805 mutants; WT: wild type *Synechococcus* sp. WH7805. The arrow points to the most obvious difference in the LPS profiles of wild type and mutant strains. Note, the bottom of the gel, showing bands related to lipid A (which showed no difference between mutants and WT), was removed from the figure to highlight differences between the O-polysaccharide chains.

6.3.3. Whole genome sequencing of cyanophage resistant *Synechococcus* sp. WH7805 mutants

Whole genome sequencing was performed on 10 cyanophage resistant *Synechococcus* sp. WH7805 mutants and the wild type in order to identify the genetic differences between them and hence assign the genes possibly implicated in cyanophage resistance. Importantly, this approach would also identify whether the variation in frequency of mutations observed in the cyanophage-resistant *Synechococcus* sp. WH7803 mutants (Chapter 4) was also observed in this monoploid strain.

The design of the experiment included the re-sequencing of two wild type *Synechococcus* sp. WH7805 cultures: before and after obtaining the cyanophage resistant mutants (Fig. 6.7, Table 2.11). This was done in order to control for mutations accumulated over the course of the experiment, i.e. to follow parallel evolution of the cultures sequenced from the same origin.

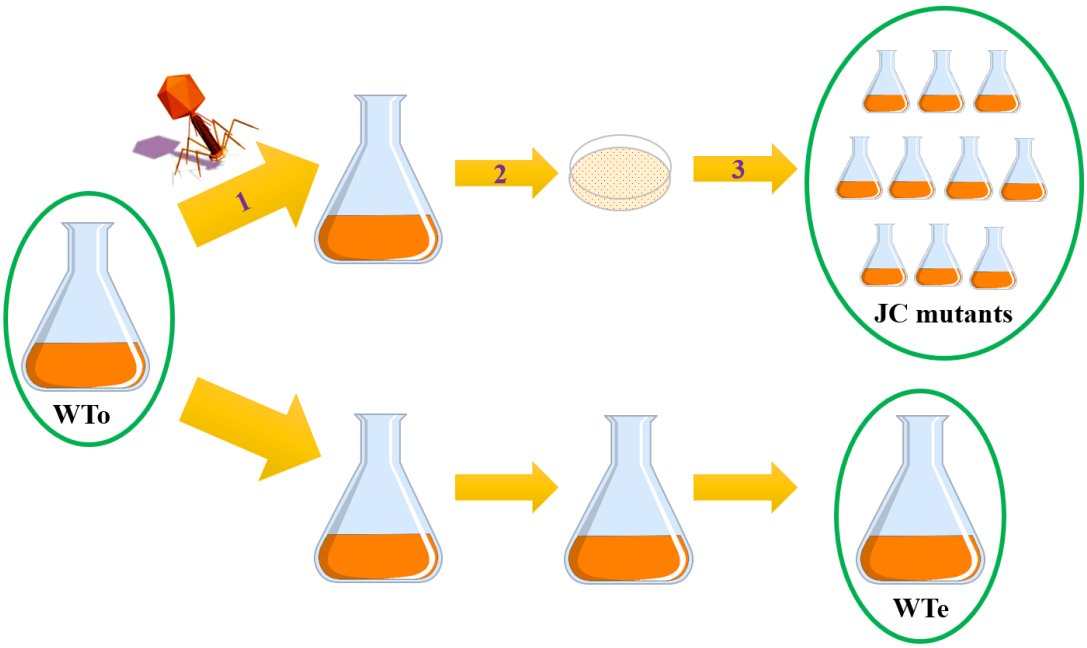


Figure 6.7. Schematic diagram of *Synechococcus* sp. WH7805 cultures selected for whole genome sequencing.

For isolation of cyanophage resistant mutants (see section 2.7) an axenic *Synechococcus* sp. WH7805 culture (WTo: original wild type) was divided in 2 aliquots. One aliquot was infected with cyanophage S-PM2 (1), the survivors were grown in solid media (2) and isolated colonies were grown in liquid ASW medium (3) and maintained with the presence S-PM2 (JC mutants). The other aliquot was continued growing in standard conditions (section 2.1.2) in parallel with the cyanophage resistant mutants (WTe: evolved wild type). In green are the cultures selected for whole genome sequencing.

Genomic DNA from all the *Synechococcus* sp. WH7805 cultures (i.e. WTo, WTe and all mutants) was extracted (see Section 2.13) and the integrity was visualised by agarose gel electrophoresis stained with ethidium bromide (Fig. 6.8) before sending samples for Illumina sequencing at the Centre for Genomic Research, University of Liverpool.

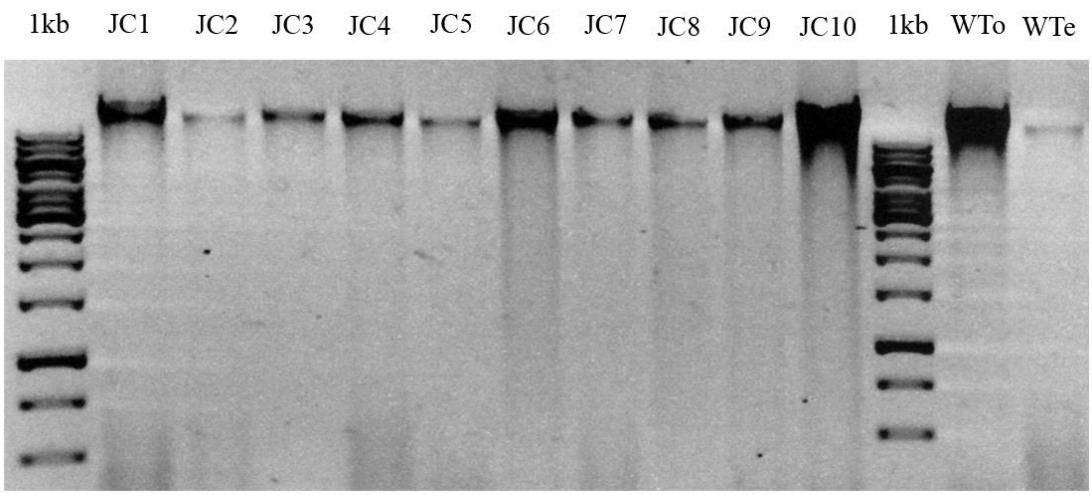


Figure 6.8. *Synechococcus* sp. WH7805 genomic DNA integrity visualised by agarose gel electrophoresis and ethidium bromide staining.

JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTo: original *Synechococcus* sp. WH7805 wild type from May 2014; WTe: evolved *Synechococcus* sp. WH7805 wild type from December 2014 (see Table 2.11 and Fig. 6.6); 1 kb ladder (Invitrogen).

Sequencing reads were obtained in fastq format and quality controlled using FastQC v.0.10.1 (see section 2.14). After trimming, the average number of reads obtained for all samples was $2.7 \times 10^6 \pm 4.6 \times 10^5$ (Fig. 6.9).

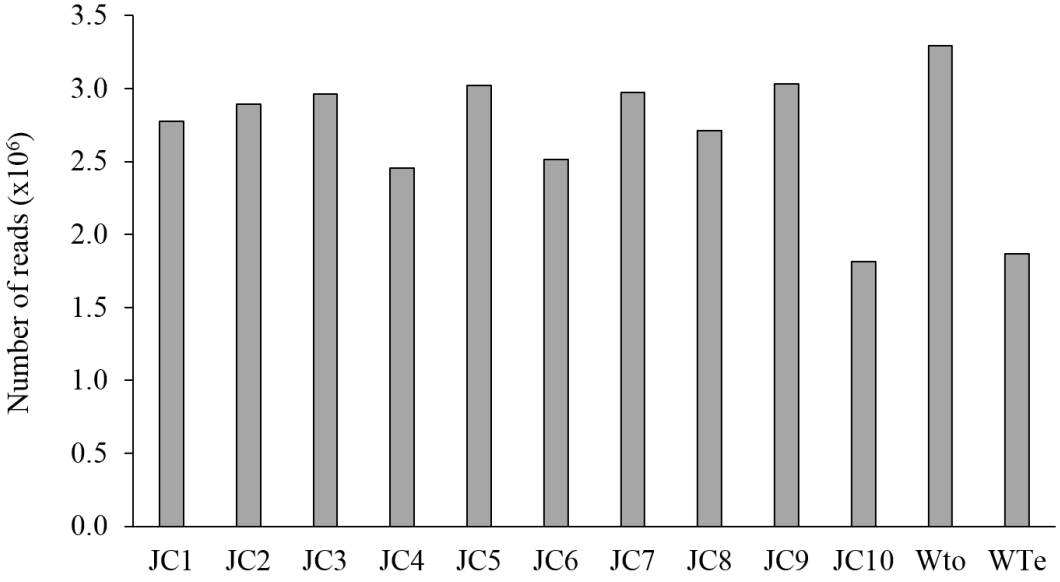


Figure 6.9. Total number of reads obtained from Illumina sequencing of each *Synechococcus* sp. WH7805 sample.

JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTo: original *Synechococcus* sp. WH7805 wild type; WTe: evolved *Synechococcus* sp. WH7805 wild type.

6.3.4. Alignment of sequencing reads against the published *Synechococcus* sp. WH7805 genome

Sequencing reads in fastq format were mapped using BWA mem (see Section 2.14.1) against the published genome sequence of *Synechococcus* sp. WH7805 (NZ_AAOK00000000.1), which has two contigs of 2,621,166 bp and 5,980 bp.

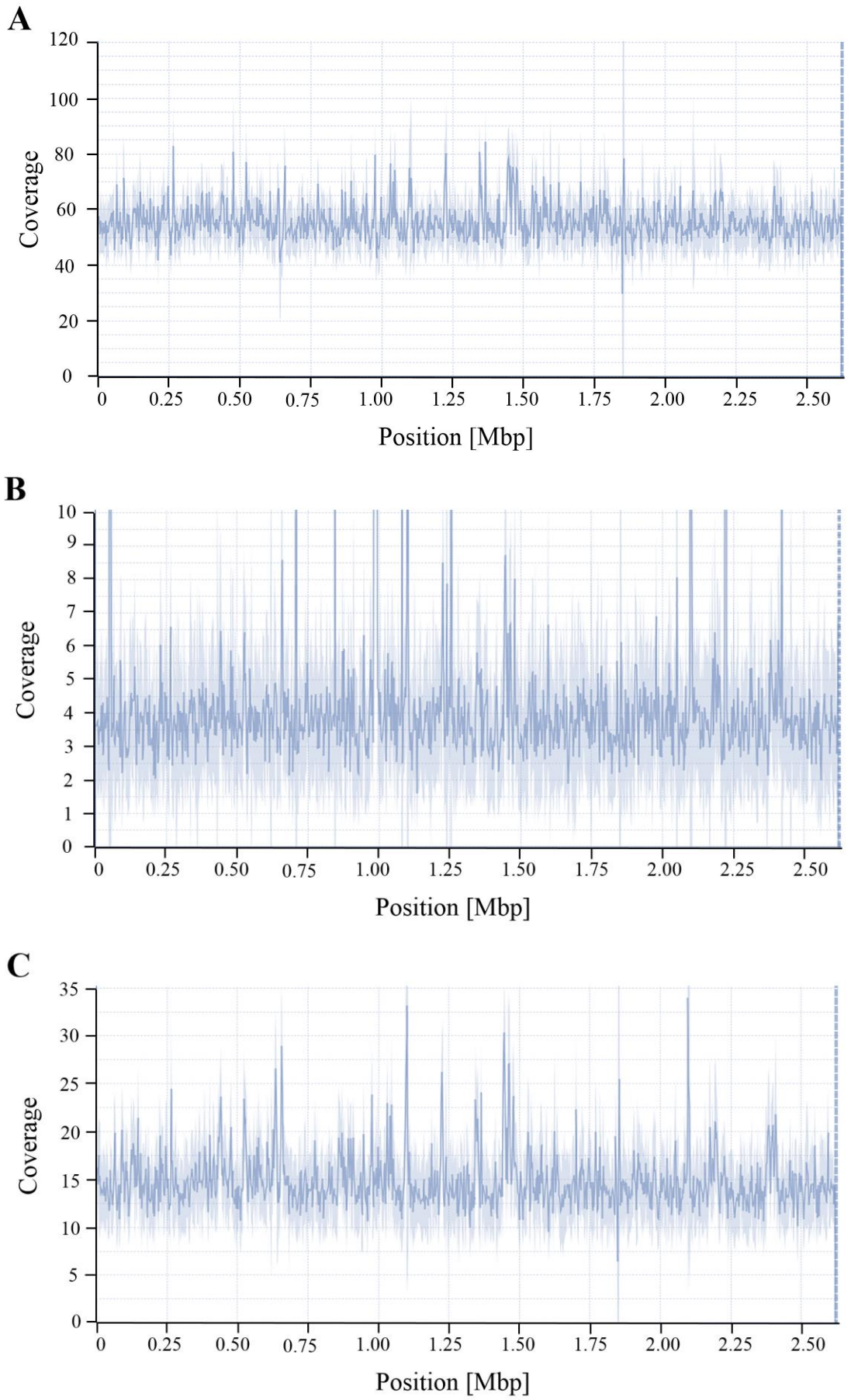
Strain	Mapped reads	Percentage of mapped reads as a proportion of the total reads for each sample [%]	Coverage mean
JC1	996,695	35.9	55.33±13.45
JC2	80,821	2.8	4.15±6.57
JC3	268,662	9.1	14.76±5.76
JC4	492,896	20.1	26.77±8.52
JC5	1,205,182	39.9	67.5±13.37
JC6	603,276	24.0	33.26±8.90
JC7	276,765	9.3	15.01±5.66
JC8	674,921	24.9	37.29±9.86
JC9	427,728	14.1	23.56±6.76
JC10	662,197	36.6	35.89±9.79
WTo	3,248,699	98.7	182.53±45.58
WTe	1,826,381	97.9	95.24±57.80

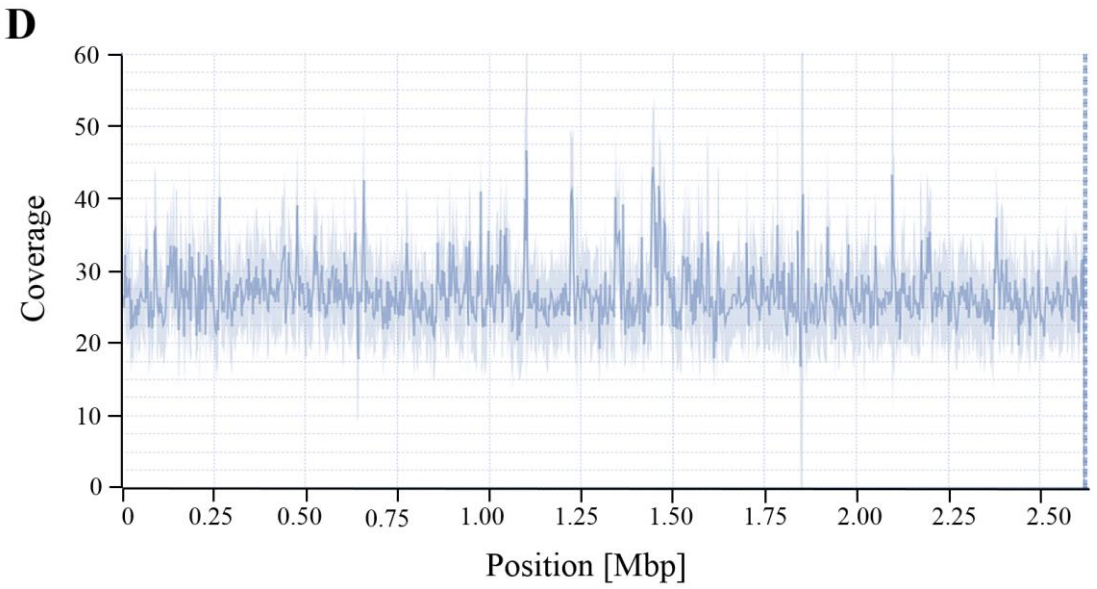
Table 6.3. Alignment of *Synechococcus* sp. WH7805 to the published reference sequence (NZ_AAOK00000000.1) using BWA mem.

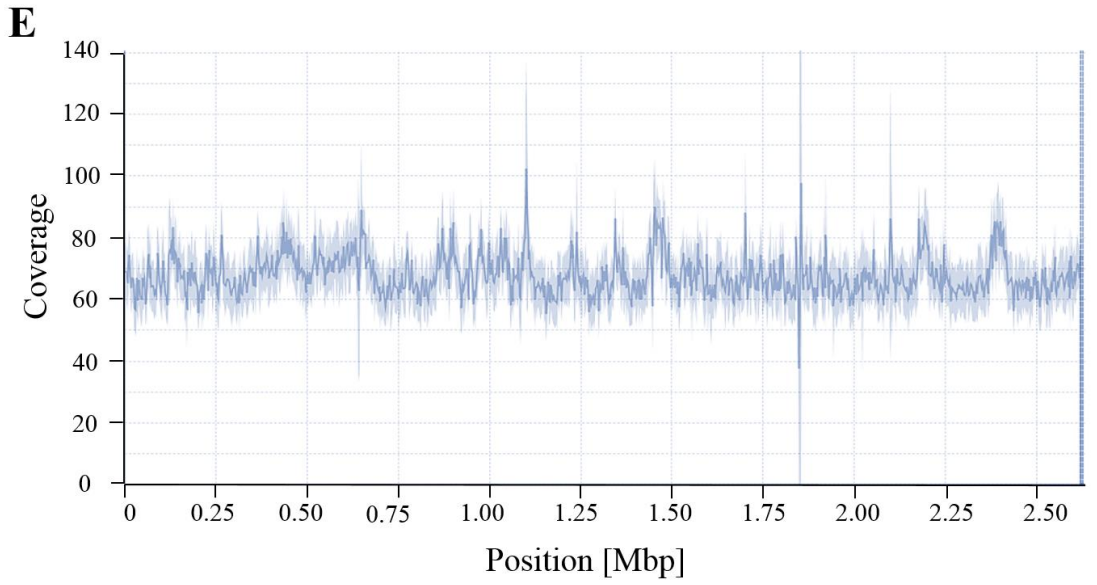
JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTo: original *Synechococcus* sp. WH7805 wild type; WTe: evolved *Synechococcus* sp. WH7805 wild type. The coverage was obtained using Qualimap (see Section 2.14|).

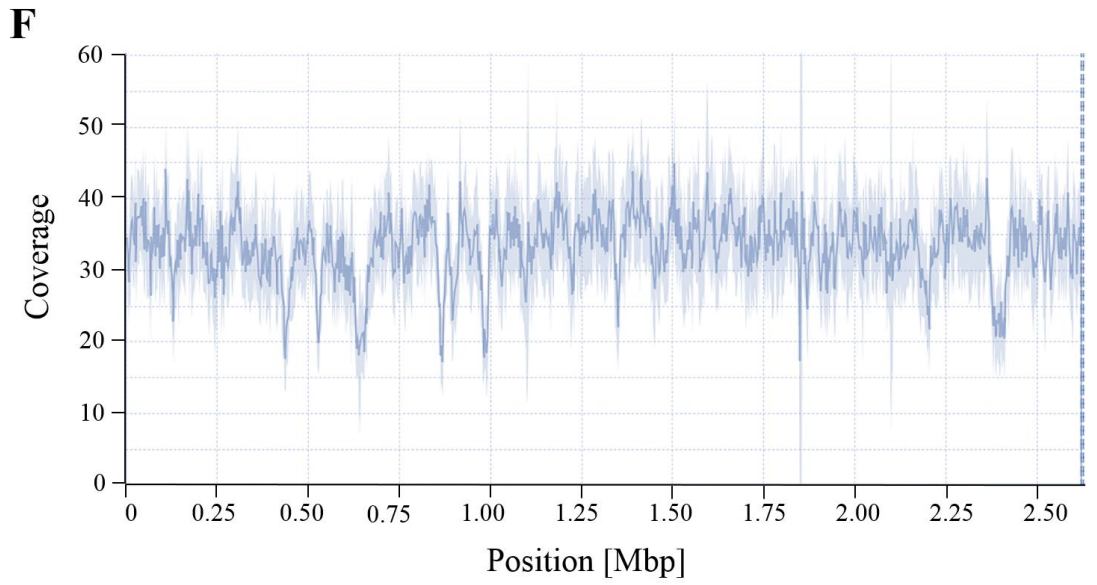
The lower number of aligned reads from the mutants compared to the wild type strains corresponds to heterotrophic bacteria contamination present in all the mutant cultures (Table 6.3).

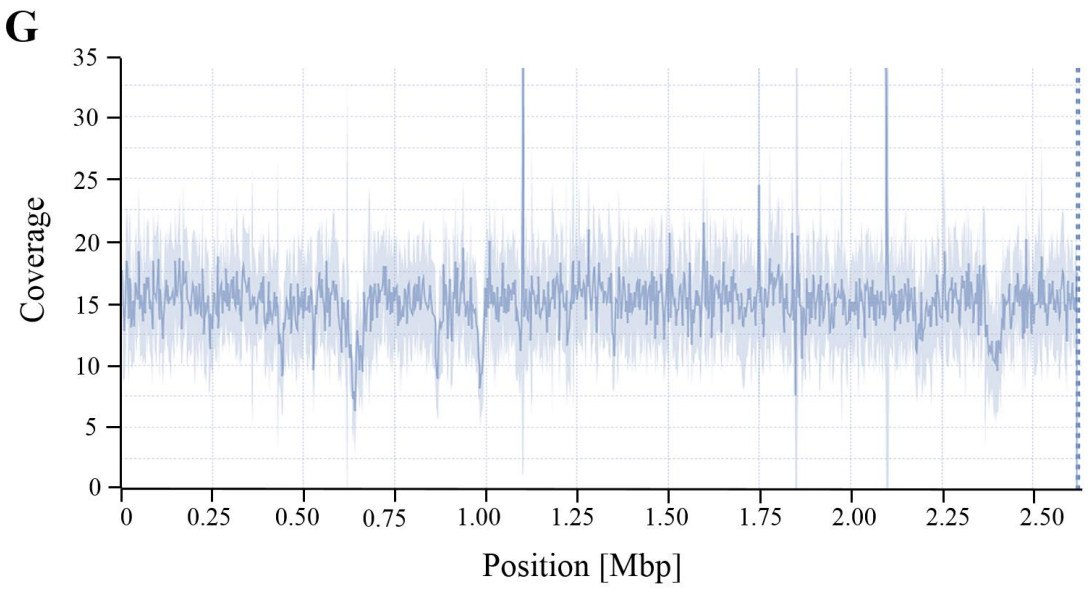
The sequencing coverage for each nucleotide in the *Synechococcus* sp. WH805 genome is shown in Fig. 6.10. The coverage ranges between ~15 to ~183 fold for all the samples, with the exception of mutant JC2 where the coverage was 4.15±6.57 fold (Fig. 6.10.B).

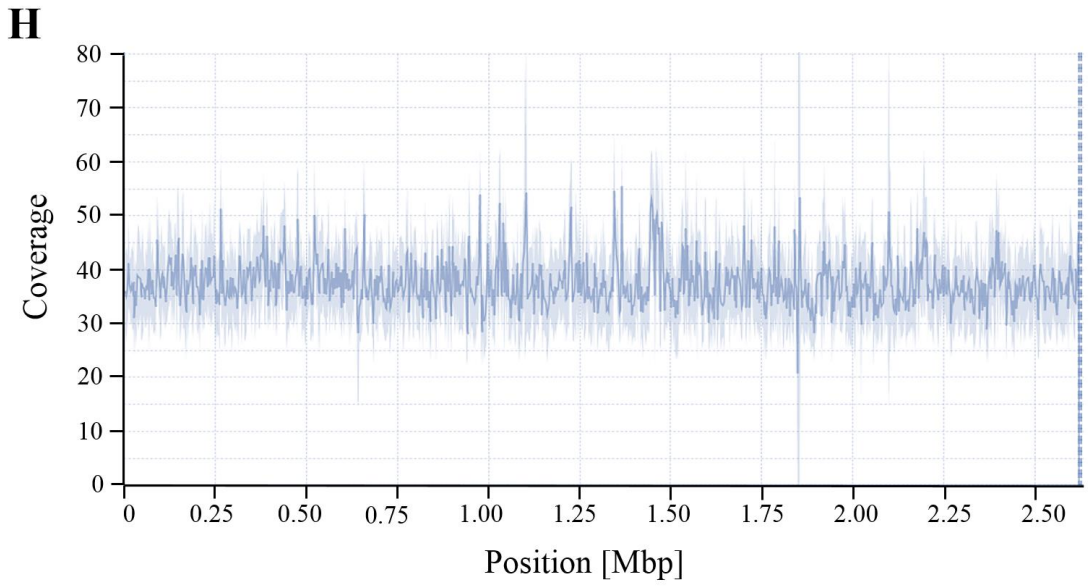

A
Coverage
120
100
80
60
40
20
0
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]
B
Coverage
10
9
8
7
6
5
4
3
2
1
0
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]
C
Coverage
35
30
25
20
15
10
5
0
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

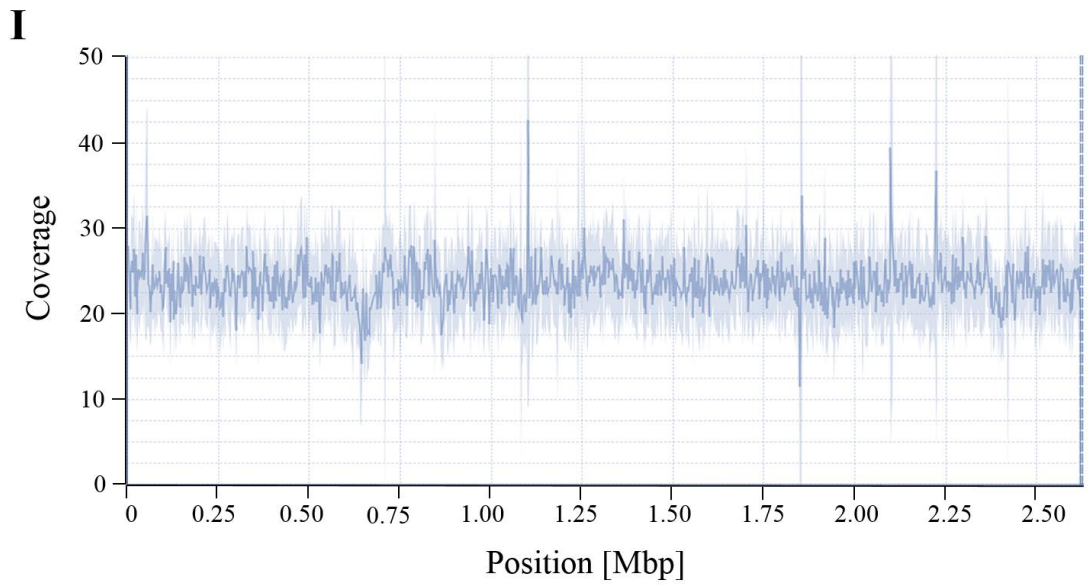

D
Coverage
0
10
20
30
40
50
60
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

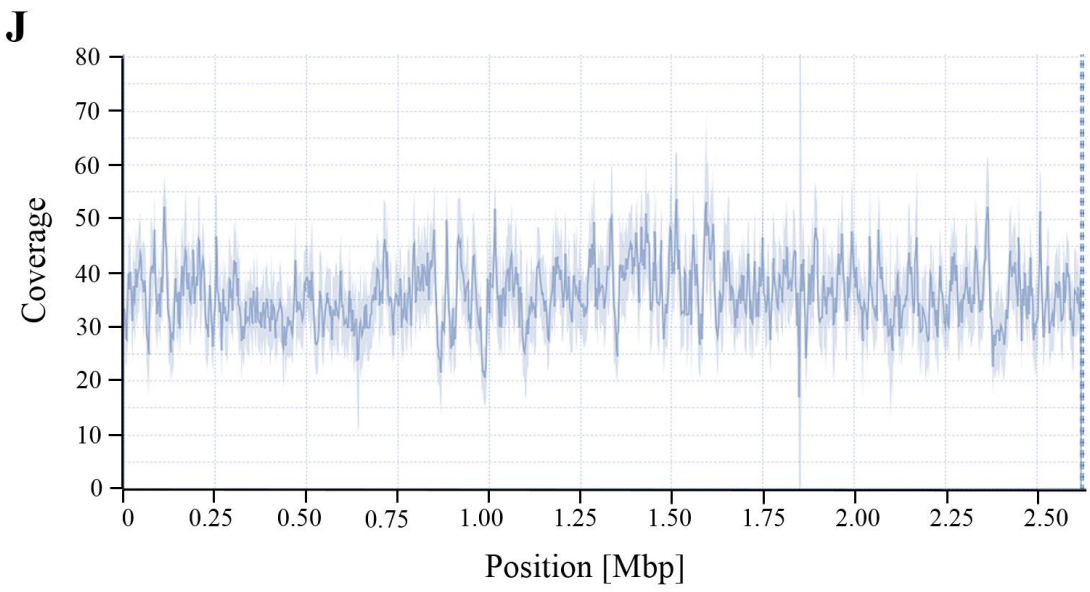

E
Coverage
0
20
40
60
80
100
120
140
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

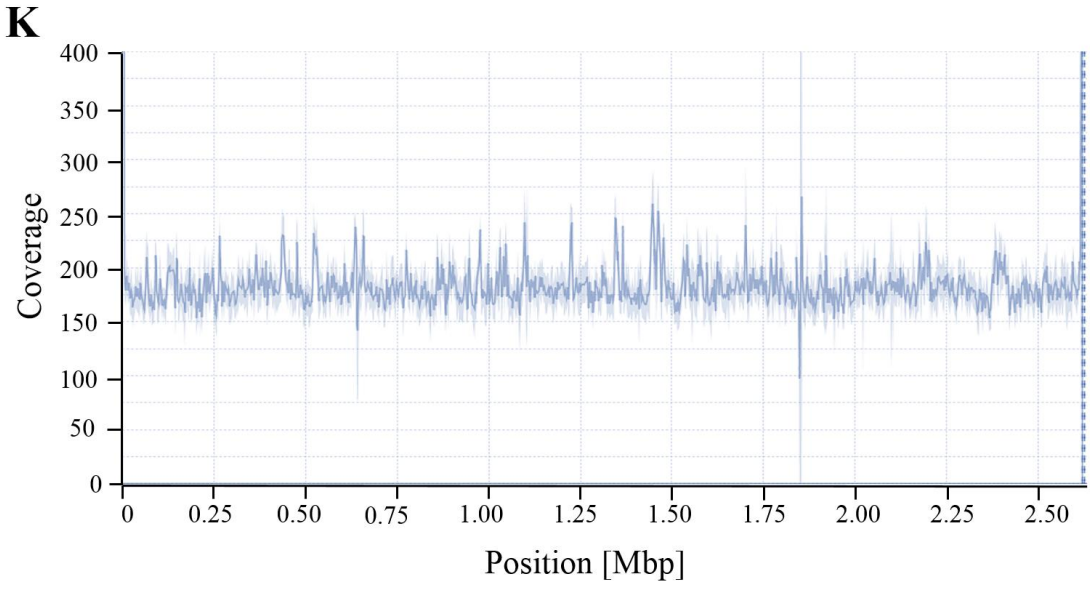

F
Coverage
0
10
20
30
40
50
60
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

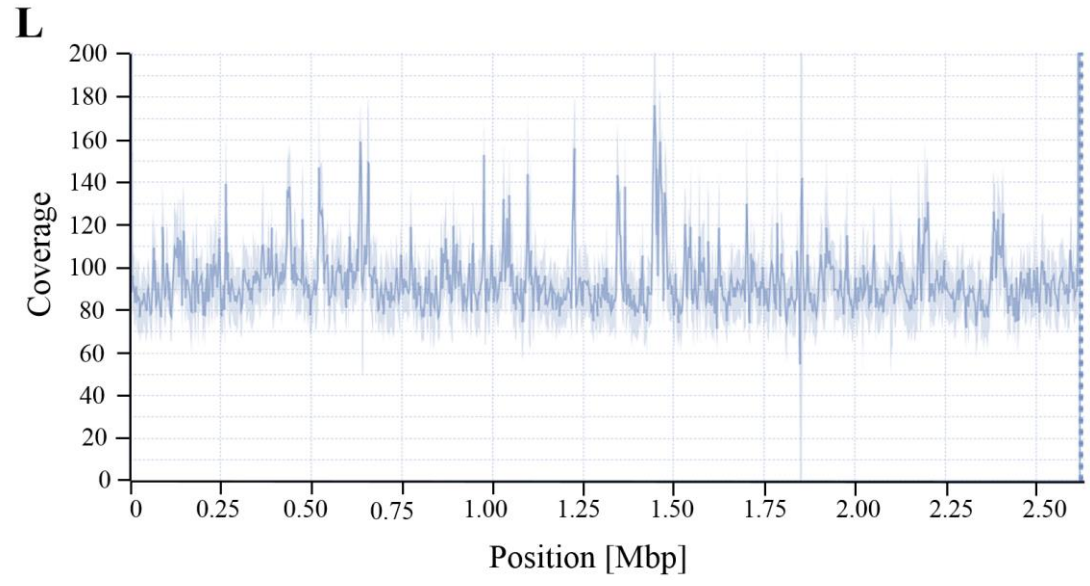

G
Coverage
0
5
10
15
20
25
30
35
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]


H
Coverage
0
10
20
30
40
50
60
70
80
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]


I
Coverage
0
10
20
30
40
50
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]


J
Coverage
0
10
20
30
40
50
60
70
80
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]


K
Coverage
0
50
100
150
200
250
300
350
400
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]


L
Coverage
0
20
40
60
80
100
120
140
160
180
200
0
0.25
0.50
0.75
1.00
1.25
1.50
1.75
2.00
2.25
2.50
Position [Mbp]

(Previous page)

Figure 6.10. Fold coverage of the different alignments across the genome of the published *Synechococcus* sp. WH7805 sequence.

(A) Mutant JC1; (B) Mutant JC2; (C) Mutant JC3; (D) Mutant JC4; (E) Mutant JC5; (F) Mutant JC6; (G) Mutant JC7; (H) Mutant JC8; (I) Mutant JC9; (J) Mutant JC10; (K) Original wild type (WTo); (L) Evolved wild type (WTe). The figures were modified from Qualimap output (see Section 2.14).

The BAM files (binary alignment map files; see Section 2.14.1) were processed and 'variants' were called using VarScan (Section 2.14.2). The number of mutations, including indels and SNPs, found in each of the *Synechococcus* sp. WH7805 cyanophage resistant mutants as well as the WTo and WTe strains, compared to the published *Synechococcus* sp. WH7805 genome sequence are summarised in Table 6.4. Only the JC9 mutant and the evolved wild type (WTe) possessed more than 100 mutations compared to the published WH7805 sequence. This suggests either an artefact or the simultaneous sequencing of multiple populations. In the case of the JC9 mutant it is highly unlikely that multiple populations are present since this mutant was obtained from a single colony. For this reason further methods were used to analyse the data in order to 'filter out' possible artifacts and detect 'true' mutations.

Strain	Number of Indels	Number of SNPs	Total number of mutations
JC1	15	14	29
JC2	10	43	53
JC3	17	69	86
JC4	14	24	38
JC5	15	82	97
JC6	11	67	78
JC7	17	77	94
JC8	13	83	96
JC9	22	181	203
JC10	12	40	52
WTe	23	557	580
WTo	23	25	48

Table 6.4. Number of mutations found in each of the *Synechococcus* sp. WH7805 cyanophage resistant mutants and the original and evolved WT culture.

This filtering focused on problems associated with genes that are present in multiple copies in the chromosome since this can cause read misalignment. As described previously (see section 4.3.3), *Synechococcus* sp. WH7805 has five copies of *psbA* (Dufresne et al., 2008) encoding the D1 protein of photosystem II. Hence, for each mutant and WT strain the mutations obtained were filtered by excluding all reads aligned to *psbA* from the analysis. Consequently, the high number of mutations in the JC9 mutant and the evolved wild type (WTe) plummeted when filtering out *psbA*, indicating a high number of false positives that correspond to misalignment of reads belonging to genes with multiple copies in the genome (Fig. 6.11).

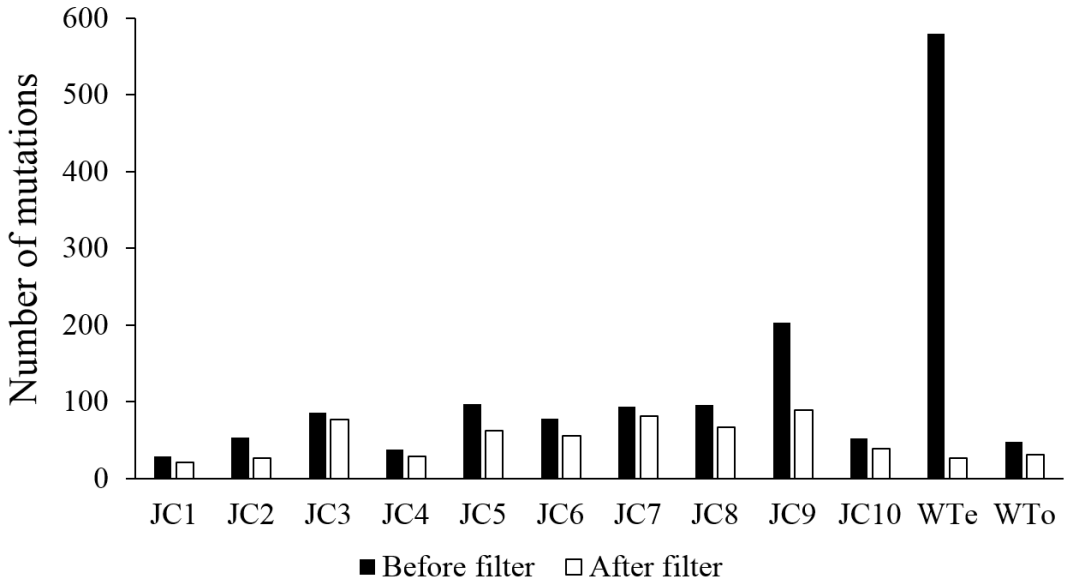


Figure 6.11. Number of mutations in whole genomes from each of the cyanophage-resistant and WT *Synechococcus* sp. WH7805 strains following 'filtering' for *psbA*.

JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTo: original *Synechococcus* sp. WH7805 wild type; WTe: evolved *Synechococcus* sp. WH7805 wild type.

After filtration of the sequences aligned to *psbA*, the variation in frequency of each mutation along the *Synechococcus* sp. WH7805 genome, in each mutant and WT strain, was assessed (compared to the published WT strain genome sequence). Fig. 6.12 shows the observed variation in frequency in all samples, which shows a very different profile compared to that seen in *Synechococcus* sp. WH7803 (Fig. 4.7) and similar to *Prochlorococcus marinus* MED4 (Fig. 4.8). In *Synechococcus* sp. WH7803 an accumulation of data points can be observed in the middle part of the graph meaning that a high number of mutations have variable frequency (Fig. 4.7) whereas in *Synechococcus* sp. WH7805 (Fig. 6.12) and *Prochlorococcus* (Fig. 4.8) the data is mostly distributed in the higher and lower extremes of the frequency.

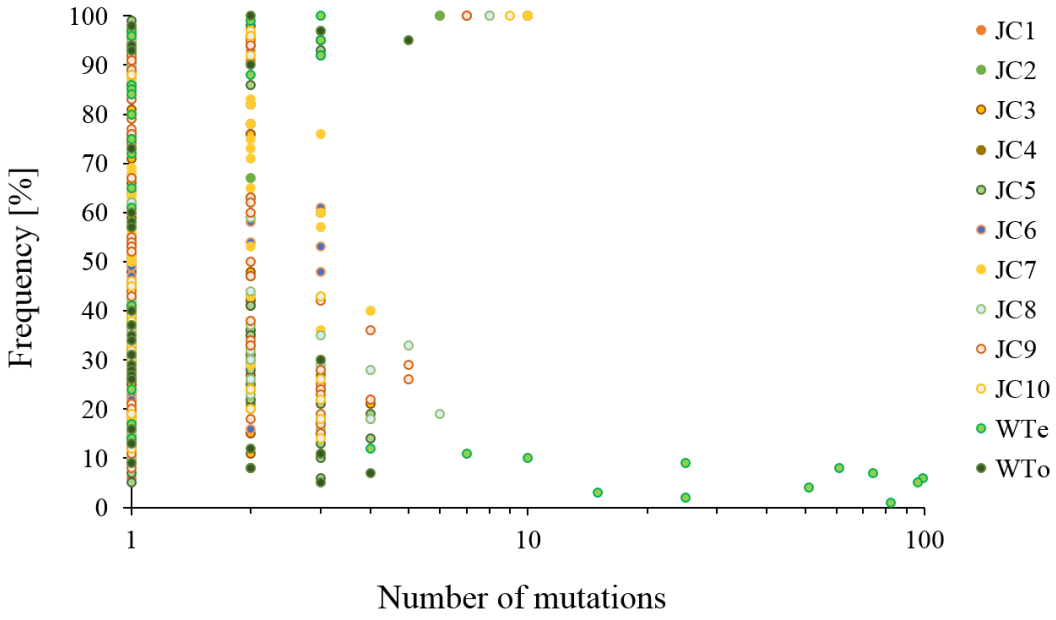


Figure 6.12. Frequency of mutations found in *Synechococcus* sp. WH7805 cyanophage-resistant and WT strains compared to the published genome sequence.

JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTo: original *Synechococcus* sp. WH7805 wild type; WTe: evolved wild type. The frequency is the percentage of reads possessing the mutation compared to the total number of aligned reads. Frequencies of reads aligned to *psbA* were filtered out.

6.3.5. Identification of possible resistance-giving mutations in *Synechococcus* sp. WH7805

In order to identify the genes responsible for cyanophage resistance in *Synechococcus* sp. WH7805, firstly genes with mutations only present in both of the wild type strains (WTe and WTo) were discarded. In total, the wild type *Synechococcus* sp. WH7805 contained mutations in 31 genes (compared to the original published sequence), of which 13 were also present in all or most of the cyanophage-resistant mutants (Table 6.5). Based on the data analysis of *Synechococcus* sp. WH7803 and *Prochlorococcus* WGS (see Chapter 4) and the frequency of mutations profile obtained for *Synechococcus* sp. WH7805 (Fig. 6.12), only mutations with a frequency > 10 % were considered further (Table 6.6).

Gene ID	Gene product	JC1	JC2	JC3	JC4	JC5	JC6	JC7	JC8	JC9	JC10	WTe	WTo
125	Hypothetical protein	97		92	100	92	100	100	97	96	93	92	97
245	Homoserine kinase	100	100	100	100	100	100	100	100	100	100	100	99
1147	Hypothetical protein	89		100	94	93	92	88	90	92	92	95	95
2587	Branched-chain amino acid aminotransferase	100	100	100	100	100	100	100	100	96	100	100	93
3217	Phosphoribulokinase	97	97	100	100	100	100	100	100	100	100	100	100
3657	Histidine kinase	84		82	100	86	91	83	75	89	85	85	90
3991	Hypothetical protein	100	100	100	100	94	100	92	93	90	100	98	90
4216	Transcription-repair coupling factor							46				13	
4946	Putative nicotinamide nucleotide transhydrogenase, subunit beta	95	100	100	93	100	100	100	95		100		60
5781	Hypothetical protein	95		100	100	100	100	100	100	100	100	96	98
6401	Putative glycerol kinase	100	100	100	100	100	100	100	100	100	100	100	100
6916	Possible fructosamine 3-kinase	100	100	100	100	100	100	83	97	100	94	92	45
7576	Global nitrogen regulatory protein *ntcA*, CRP family of transcriptional regulators	100	100	100	100		100		94	100	100		29
7766	Probable oxidoreductase	100	100	100	100	100	100	100	100	100	100	98	99

Gene ID	Gene product	JC1	JC2	JC3	JC4	JC5	JC6	JC7	JC8	JC9	JC10	WTe	WTo
7851	Hypothetical protein	100	100	100	100	100	100	88	100	100	100		58
7956	Hypothetical protein	100	100	100	100	100	100	100	100	100	97	100	95
8852	Putative RND family multidrug efflux transporter	91		100	94	92	93	82	92	94	96	84	95
9489	Putative phospho-N-acetylmuramoyl-pentapeptide-transferase						25	41					11
9564	Hypothetical protein	100	100	95	100	100	100	100	100	100	100		57
9919	Putative ABC multidrug efflux transporter					98						72	30
10024	Short-chain dehydrogenase/reductase (SDR) superfamily	96		100	100	96		100	96	95	100	95	95
10638	Large exoprotein involved in heme utilization or adhesion				26	100	32	47	32	32	17		80
12283	Periplasmic trypsin-like serine protease	88		82	100	91		80	97	91	92	93	95
13863	ABC-type Fe3+ transport system, membrane component			100	100	100	97	100	100	95	100		100

Table 6.5. Mutation frequency in genes where there were mutations in both *Synechococcus* sp. WH7805 cyanophage resistant mutants and re-sequenced wild type.

JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTo: original *Synechococcus* sp. WH7805 wild type; WTe: evolved wild type. Mutation frequency is the percentage of reads possessing the mutation compared to the total number of aligned reads. Frequencies of reads aligned to *psbA* and < 10 % were filtered out.

Four genes were identified (Table 6.6) that contained mutations across several of the mutants but which were absent in the wild type (WTe and WTo, Fig. 6.5): i) *tufA*, encoding elongation factor EF-Tu (gene ID WH7805_07276), ii) *ntcA*, encoding a global nitrogen regulatory protein (gene ID WH7805_07576), iii) a putative alpha-glycosyl transferase (gene ID WH7805_08977) and iv) *ureC*, encoding the urease alpha subunit (gene ID WH7805_09809). Of these genes, three contained silent (or synonymous) mutations, but the gene encoding a possible alpha-glycosyl transferase possessed potentially non-synonymous mutations (gene ID WH7805_08977).

Gene ID	Gene product	JC1	JC2	JC3	JC4	JC5	JC6	JC7	JC8	JC9	JC10
345	Catalase/peroxidase				32	22			40	32	
1492	Possible porin					100					
1932	Acetate-CoA ligase							47			
2072	Putative arsenite transporter, ACR3 family protein				32		18				
2417	Tyrosine binding protein				15			55			
2987	Possible exonuclease of the metallo-β-lactamase family	100	100								
3312	Dihydroxy-acid dehydratase							36	14	35	
3812	BgtA-like ABC-type uptake transporter for acidic and neutral polar amino acids (N-II), ATPase subunit	13	13					42			
3962	ABC transporter ATP-binding protein	11	11		19						
4121	Glyceraldehyde-3-phosphate dehydrogenase						26	50			
4556	Probable aminopeptidase N *pepN*		14					42			
5446	Stationary-phase survival protein SurE					89		83			
6056	GTP-binding protein LepA						28	75			
6296	Nitrogen regulatory factor *pipX*					98					

7246	Ferredoxin-dependent glutamate synthase	16	16					51	20		
7276	Elongation factor EF-Tu	20	20	38	45	28	22	40	39	55	24
7311	Ribonuclease E							29			
7396	Cell division protein FtsH2				30		21	40	18		
7716	Phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase			53							
7906	Deoxyribodipyrimidine photolyase (DNA photolyase) (photoreactivating enzyme)	22	22		30		15	69	19		
8016	Putative nucleotide sugar epimerase				26						
8932	Preprotein translocase SecA subunit				27				17		
8977	Possible alpha-glycosyltransferase, family 4	98	98	81	100	100			100	100	100
8992	Putative ABC-type polysaccharide/polyolphosphate transport system ATPase component							100			
9354	HI0933-like protein										14
9529	Molecular chaperone DnaK				24						18
9579	Aconitate hydratase						38	57	19	25	
9784	Asparagine synthase (glutamine-hydrolyzing)										17
9809	Urease alpha subunit, *ureC*			50		14	52	82	35	50	12

9819	Urease gamma subunit, *ureA*							32			
9954	Aspartyl-tRNA synthetase:GAD domain:trna synthetase, class I	15	15				20	48	27		
10438	Enolase									31	
10538	2-hydroxy-3-oxopropionate reductase				100						
10633	Metal dependent phosphohydrolase							42			
10863	Putative O-succinylbenzoate synthase										14
11063	NADH dehydrogenase subunit NdhF (chain 5 or L)							48			
12313	DNA-directed RNA polymerase, beta subunit						22	24			
12318	DNA directed RNA polymerase beta subunit							32		29	
12643	ATP synthase beta chain			44	13	41	56			38	
12893	Lipid-A-disaccharide synthase										19
12913	Leucyl aminopeptidase										18
13553	Cytochrome C oxidase			31							
13553	Possible cytochrome C oxidase subunit II			100							
13923	Porphobilinogen deaminase							40			

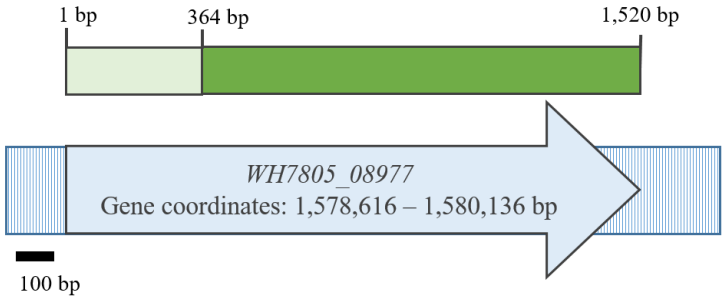
13938	Major RNA polymerase sigma factor, type I		52	
14053	Transpeptidase involved in septal peptidoglycan synthesis (peptidoglycan synthetase precursor)	99		
14068	Deoxyribodipyrimidine photolyase, possible cryptochrome *cry2*			32
14499	Uncharacterized membrane protein	100		

Table 6.6. Frequency of mutations in *Synechococcus* sp. WH7805 cyanophage resistant mutants.

JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants. The frequency is the percentage of reads possessing the mutation compared to the total number of aligned reads. Frequencies < 10 % were filtered out. The gene selected for Sanger sequencing is highlighted in red

To confirm these non-synonymous mutations found by Illumina sequencing in the possible alpha-glycosyltransferase gene of the *Synechococcus* sp. WH7805 cyanophage-resistant mutants, a region containing the mutations was amplified by PCR from each mutant and the WT (see Table 2.9 for details of the PCR primers used) using a high-fidelity Taq polymerase (see section 2.2) and Sanger sequenced (Fig. 6.13). Mutants JC2 and JC3 contained an adenine insertion at gene position 1,011 bp that causes a frameshift resulting in an early stop codon at gene position 1,045 bp. In the case of mutants JC4, JC8 and JC10, non-synonymous mutations caused by transversion occur as follows: JC4 in position 692 bp (T to G), JC8 in position 1,148 bp (C to A) and JC10 in position 1,174 bp (A to C). The potential mutations in mutants JC1 (gene position 409 bp), JC5 (gene position 100 bp) and JC9 (gene position 409 bp) could not be confirmed due to the low quality of the sequencing at the extreme ends of the Sanger sequences. Mutants JC6 and JC7 did not have any mutations in that gene predicted by WGS and confirmed by Sanger sequencing of the region selected (Table 6.5 and Fig. 6.13).

A



B

```
JC1          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC2          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC3          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC4          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRRQAAGPARRFFVSEDARQHYELSILGTLS
JC5          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC6          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC7          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC8          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC9          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
JC10         QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
WTe          QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS
Ref. WH7805  QRCERLSYLSEIDGLICANNCFLNSIELAKRKHPHSLHIDLRGFDGLITTSPLNIKPVNTRLFVQTVHDLIPLDYQRTRDHLPCFTRRLQAAGPARRFFVSEDARQHYELSILGTLS

JC1          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC2          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKKYSVLHHRQTSE*
JC3          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKKYSVLHHRQTSE*
JC4          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC5          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC6          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC7          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC8          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC9          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
JC10         NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
WTe          NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV
Ref. WH7805  NPNRLQRVVTQSPSLQFPGDCLDWEARTRQVRVLSSNQSQLHALKPCSYFLFNSSVVPHKNLLFALKAFMESDLEQKNIRFCITGKPQNDEYSEAVRSVAAKHKSVIFTGYV

JC1          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC2
JC3
JC4          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC5          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC6          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC7          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC8          DEATKRQLYLNDLALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC9          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
JC10         DEATKRQLYLNALALLSPSLLEGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
WTe          DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
Ref. WH7805  DEATKRQLYLNALALLSPSLI EGFGIPVLDAACLGLSAIASPLGSHREIQAMHDFKDHVMLCSTLVTSDWASAMRLIALKNELSLSELSPKRQRKRLNQMRAERIQRYRHYQTL
```

(Previous page)

Figure 6.13. Confirmation of mutations in gene WH7805_08977 by Sanger sequencing in cyanophage resistant mutants and wild type *Synechococcus* sp. WH7805

(A) Schematic representation of the gene WH7805_08977 (blue arrow) with the gene position of the PCR product sequenced using Sanger sequencing (green; primers are given in Table 2.9). (B) Amino acid alignment of a putative alpha-glycosyl transferase (gene ID WH7805_08977) in *Synechococcus* sp. WH7805 highlighting the position of frameshifts and non-synonymous changes in specific cyanophage-resistant mutants (red). The amino acid sequence is derived by translation of sequenced PCR products. JC1-10: cyanophage resistant *Synechococcus* sp. WH7805 mutants; WTe: *Synechococcus* sp. WH7805 wild type; Ref. WH7805: reference sequence of *Synechococcus* sp. WH7805 (NZ_AAOK00000000.1).

6.4. Discussion

Herein, several cyanophage resistant *Synechococcus* sp. WH7805 mutants were obtained and characterised. Initially, several different cyanophages were assessed in order to determine their infectivity on *Synechococcus* sp. WH7805 before attempting to obtain cyanophage resistant mutants. This was successfully accomplished for cyanophages S-PM2 and S-RIM34 where infection in both liquid and solid media was observed (Table 6.1). This was particularly important since it allowed us to use cyanophage S-PM2 to obtain resistant mutants allowing a direct comparison with the *Synechococcus* sp. WH7803 S-PM2 resistant mutants (mutants R1, R2 and PHR, see Table 2.3).

The fitness cost associated with phage resistance has been widely studied in bacterial systems, where it is mostly observed as a decreased growth rate compared to the sensitive (wild-type) strain, although this is sometimes only observable during direct competition experiments and is dependent on the nutrient conditions (Avrani et al., 2011; Bohannan et al., 1999; Martiny et al., 2014). The reduced growth rate in these cyanophage-resistant mutants has been explained by a lower nutrient uptake due to modification of membrane components such as protein transporters or LPS that could alter the hydrophobicity of the cell (Lennon et al., 2007; Lenski, 1988b). Accordingly, in this study it was found that of the ten cyanophage resistant *Synechococcus* sp.

WH7805 mutants obtained only six were able to survive (mutants JC1, JC2, JC3, JC4, JC6 and JC10), with all having a decreased growth rate compared to the wild type (Fig. 6.4 and Table 6.2) in experiments performed in nutrient replete ASW medium without direct competition.

Moreover, the clumping phenotype observed in the cyanophage resistant *Synechococcus* sp. WH7805 mutants JC2 and JC6 is also observable to a lower degree in mutants JC1, JC3, JC4 and JC10 (see Fig. 6.3), which correlates with the modification detected in LPS profiles (Fig. 6.6). Work in *E. coli* has shown that the auto-aggregation phenotype is linked to truncated LPS and increased cell hydrophobicity (Nakao et al., 2012).

The cost of resistance has been associated with modification of cell surface properties, especially membrane proteins and LPS. In *E. coli* B mutants resistant to bacteriophage T3, mutations in a glucosyltransferase I enzyme that catalyses the addition of glucose to the outer core LPS produced a non-functional enzyme, and thus a truncated LPS (Perry et al., 2015). This was associated with bacteriophage resistance since LPS has been identified as a receptor for bacteriophage T3 adsorption. Noteworthy, all mutations found in the *waaG* gene encoding the glucosyltransferase I in T3 resistant *E. coli* B mutants were different, including frameshifts from indels and early stop codons (Perry et al., 2015).

The cyanophage resistant *Synechococcus* sp. WH7805 mutants obtained were characterised by a similar phenotype to that observed in cyanophage resistant *Prochlorococcus* mutants (Avrani et al., 2011), i.e. all of the mutants obtained (Fig. 6.5) were resistant to the adsorption of the cyanophage. This is consistent with the observed modifications to the LPS profiles of these mutants compared to the wild type *Synechococcus* sp. WH7805 (Fig. 6.6). Moreover, it was interesting, and pleasing, to find that mutations in a putative glycosyltransferase (gene ID WH7805_08977) were the major difference between mutant and wild type following whole genome sequencing of resistant strains.

Mutations in this gene (WH7805_08977) are absent in all the wild type *Synechococcus* sp. WH7805 genomes sequenced but 'shared' between eight of the ten mutants i.e. mutations are present in the same gene but not necessarily in the same location within

this gene. This is suggestive of a 'common' receptor based modification mechanism for becoming cyanophage resistant in *Synechococcus* sp. WH7805, similar to what has recently been proposed in *Prochlorococcus* (Avrani et al., 2011; Avrani & Lindell 2015).

Using Cyanorak (http://abims.sb-roscoff.fr/cyanorak/; Scanlan et al., 2009), the gene identified in several of the cyanophage resistant *Synechococcus* sp. WH7805 mutants (gene ID WH7805_08977), and annotated as a putative glycosyltransferase, contains orthologs in other marine *Synechococcus* and *Prochlorococcus* strains but with a patchy distribution (present in 3 of 41 *Prochlorococcus* genomes and 6 of 51 *Synechococcus* genomes). Within *Synechococcus,* the gene is present mostly in strains that are closely related to WH7805 phylogenetically, i.e. strains WH7803, PROS-7-1, A15-60 and A18-25c (Table 6.7). Noteworthy, no mutations were found in the ortholog of this gene present in *Synechococcus* sp. WH7803 (gene ID *synWH7803_0139*) after WGS analysis of cyanophage resistant mutants (see Chapter 4).

Regarding the function of this putative glycosyltransferase we can only speculate that this enzyme may be involved in LPS biosynthesis given the altered LPS profile of mutants (Fig. 6.3), but further, more extensive biochemical characterisation of the LPS of the mutant and wild type WH7805 strains would be required to confirm this and especially since, thus far, biochemical characterisation of the LPS of marine *Synechococcus* is limited to only a few strains (Snyder et al., 2009). Indeed, given the general sugar-modifying ability of glycosyltransferases it is possible that this enzyme might be involved in the glycosylation of other cell surface exposed features, particularly specific outer membrane proteins, as has been observed in the SwmA protein of *Synechococcus* sp. WH8102 (Brahamsha, 1996b).

Given that mutants JC6 and JC7 do not have mutations in this WH7805_08977 gene, further in-depth analysis of all the genes identified by whole genome sequencing is required, particularly confirming mutations via Sanger sequencing to identify the genes involved in cyanophage resistance in these particular mutants.

Strain	Clade	Gene ID
Synechococcus		
ROS8604	Ib	00104
WH8016	Ib	01809
WH7803	V	0139
PROS-7-1	VIb	00159
A15-60	VIIa	00152
A18-25c	VIIa	00119
Prochlorococcus		
MIT0701	LLIV	01502
MIT0702	LLIV	01761
MIT0703	LLIV	02360

Table 6.7. Orthologs of the possible glycosyltransferase (gene ID WH7805_08977) from *Synechococcus* sp. WH7805 in other marine picocyanobacteria

Noteworthy is that several genes contained mutations in all of the cyanophage resistant *Synechococcus* sp. WH7805 mutants, as well as the two re-sequenced wild type strains (WTo and WTe, see Table 6.5). This suggests sequencing errors in the originally published *Synechococcus* sp. WH7805 sequence (Genbank accession NZ_AAOK00000000.1). In this respect, further analysis of the *Synechococcus* sp. WH7805 cyanophage-resistant mutants (Tables 6.5 and 6.6) revealed that mutants JC1 and JC2 show very similar profiles, but with mutant JC2 not possessing mutations in genes encoding a histidine kinase, a periplasmic trypsin-like serine protease, a putative nicotinamide nucleotide transhydrogenase, putative multidrug efflux transporter and a short-chain dehydrogenase/reductase (SDR) superfamily protein. The fact that these mutations are present in all other *Synechococcus* sp. WH7805 mutant and WT strains, i.e. all except mutant JC2, may be due to the low sequence coverage of this mutant

(only ~4 fold, compared to 14-182-fold for the other strains; see Table 6.3 and Fig. 6.10.B).

The prevalence of adsorption mutants in *Synechococcus* sp. WH7805 and the mutations found in a glycosyltransferase that appears to be implicated in a modified LPS phenotype indicate that the mechanisms of cyanophage resistance in this strain are likely to be related to cell surface modification, in agreement to what has been suggested for *Prochlorococcus*. However, to truly confirm the association of this enzyme in cyanophage resistance it would be necessary to develop a genetic system for *Synechococcus* sp. WH7805 and construct a mutant in the WH7805_08977 gene.

Chapter 7

PROTIST GRAZING AND VIRAL INFECTION OF MARINE *SYNECHOCOCCUS*

7.1 Introduction

During the course of this work, the importance of viral infection of marine *Synechococcus* has been discussed. In this respect, although viral lysis is a major factor potentially controlling bacterial population abundance, precise rates have not yet been determined, though estimates show up to 50% of daily viral lysis in planktonic systems, a similar rate to grazing, the other main biotic cause of mortality of these organisms (Danovaro et al., 2011; Fuhrman and Noble, 1995; Proctor and Fuhrman, 1990; Weitz et al., 2015). Thereby, predation of marine *Synechococcus* plays a central role in terms of their mortality in pelagic systems while also serving to regenerate nutrients and transport energy to higher trophic levels. This is caused by grazing (i.e. predation or consumption) by eukaryotic protists as well as bacteriophage (i.e. viral) lysis (Fuhrman, 1999; Pernthaler, 2005; Rodriguez-Valera et al., 2009; Sherr and Sherr, 2002; Suttle, 2007).

However, in spite of high predation of marine *Synechococcus*, population sizes are fairly stable showing no extinction (Flombaum et al., 2013). Such an observation can be explained by rates of *Synechococcus* production counteracting this high mortality, but also can be due to phage resistance and prey-selectivity mechanisms, or a mixture of these.

Due to the small size of marine *Synechococcus*, mesozooplankton and small copepods are not able to directly consume them, so the carbon flux is more likely to move to these trophic levels from nanoflagellate and small ciliate grazers < 5 µm in size that feed on these primary producers (Guillou et al., 2001). Although *Synechococcus* has been determined to be grazed at a low level by heterotrophic nanoflagellates, frequently called “poor quality food” (Christaki et al., 2002), Worden and Binder (2003) demonstrated a grazing rate of about 50 % prey population consumption in 9 h, showing the high impact grazing can have on picocyanobacterial removal.

Thus far, however, little is known about the prey selectivity of grazers. It is often thought to be undirected, with grazing rates dependent mainly on grazer–prey contact rates, which are a function of cell density, cell motility and prey cell size (Boenigk et al., 2001; Corno and Jürgens, 2006; Gonzalez, 1996; Matz and Jürgens, 2005; Matz and Kjelleberg, 2005; Matz et al., 2002).

It is generally known that protists can graze on a variety of bacterial species, but viruses are usually highly specific for certain hosts (Fuhrman and Noble, 1995; Fuhrman and Suttle, 1993). Prey evasion of protist grazing can be achieved by a variety of mechanisms, such as high motility, reduction in cell size, toxin release, cell aggregation by formation of an exopolymer and filamentation, among others (Apple et al., 2011; Gonzalez, 1996; Pernthaler, 2005).

Zwirglmaier et al. (2009) showed that grazing selection of marine *Synechococcus* is manifest at the strain level (Fig. 7.1) occurring at both the stage of ingestion and digestion. Ingestion and digestion of a certain prey can be clearly distinguished by the growth of the grazer, meaning that a digested prey will be assimilated hence supporting the growth of the grazer, compared to only ingestion when the prey cell concentration decreases but with no increase in grazer cell concentration.

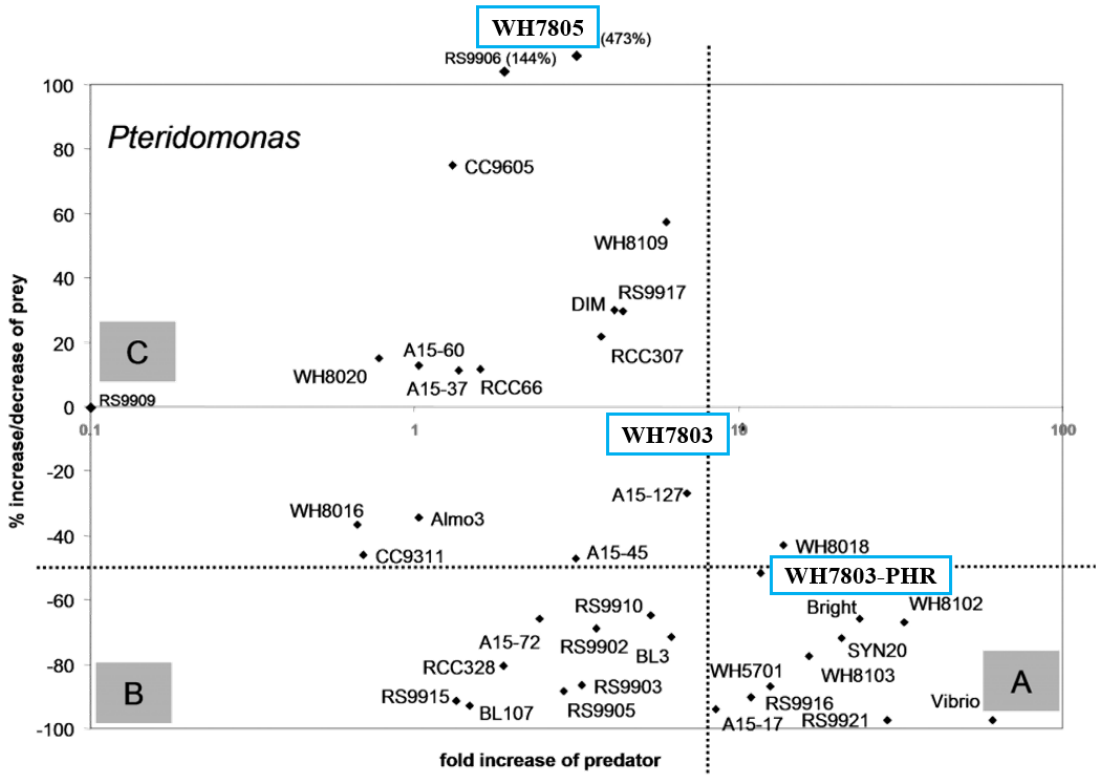


Figure 7.1. Grazing of *P. danica* on marine *Synechococcus* (modified from Zwirglmaier et al. 2009)

Increase in grazer (predator) cell numbers versus the percentage of decrease in prey (*Synechococcus* spp.) cell numbers. Positive grazing (ingestion and digestion of a specific *Synechococcus* strain) was defined as a more than 8-fold increase in grazer numbers (dashed line, X axis) while a decrease in more than 50 % of prey was defined as the threshold for prey mortality due to grazing (dashed line, Y axis). Thus, prey strains in quadrant A support growth of grazers (ingestion and digestion of prey), strains in quadrant B are ingested, but do not support growth and are therefore presumably not digested, while for strains in quadrant C there is no interaction between grazer and prey. *Synechococcus* strains used in this work are highlighted in blue: *Synechococcus* sp. WH7803 and WH7805 in quadrant C and the cyanophage-resistant *Synechococcus* sp. WH7803-PHR in quadrant A.

Furthermore, Zwirglmaier et al. (2009) also showed that the cyanophage resistant *Synechococcus* sp. WH7803 mutant PHR was more susceptible to grazing compared to the wild type (Fig. 7.2.A) and that this mutant possessed a modified LPS profile (Fig. 7.2.B), lacking a significant portion of its O-polysaccharide units, (see also Fig. 3.4). In this regard, cell surface properties seem to be a key aspect (Snyder et al., 2009; Zwirglmaier et al., 2009).

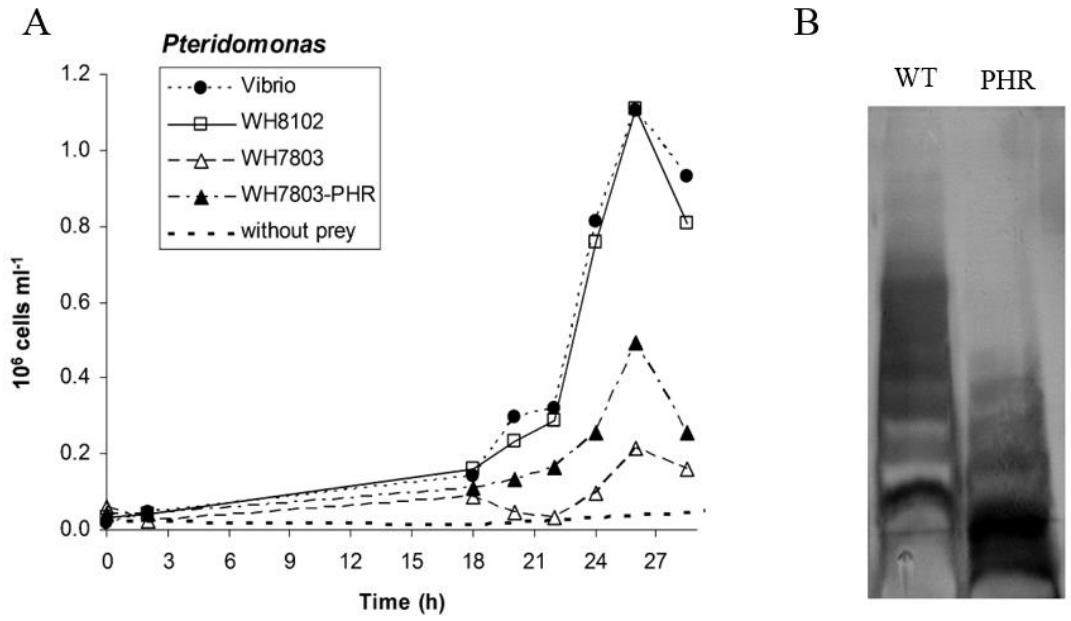


Figure 7.2. Grazing and LPS profile of *Synechococcus* sp. WH7803 and a cyanophage resistant mutant (from Zwirglmaier et al. 2009)

(A) Growth curves of *P. danica* feeding on different prey strains. Note that the cyanophage resistant *Synechococcus* sp. WH7803 mutant (PHR) supports growth of the grazer while the wild type (WH7803) does not. Growth on *Synechococcus* sp. WH8102 and *Vibrio* are used as a positive control for grazing. (B) SDS-PAGE showing differences in LPS profiles of cyanophage resistant *Synechococcus* sp. WH7803 mutant (PHR) and wild type (WT). The LPS profiles indicate greater heterogeneity in the wild type in comparison with the mutant, indicating that a significant portion of the mutant's O-polysaccharide side-chain units have been lost.

The role of LPS in phage resistance/infection has previously been discussed in this thesis. With respect to grazing, LPS modification seems to be related to physicochemical properties of the cell surface of the prey, which could affect contact rates with the grazer. In this regard, mutations in the LPS biosynthetic pathway of the freshwater *Synechococcus* sp. PCC 7942 conferred resistance to grazing by the Amoeba HGG1 (Simkovsky et al., 2012) meaning that cell surface properties can directly affect protist grazing. The mechanism for grazing resistance related to LPS is different from phage resistance, i.e. there are no known receptors involved in the selection of the prey, but rather resistance seems to be related to a change in prey cell

hydrophobicity. Accordingly, Monger et al. (1999) demonstrated that an augmented prey cell hydrophobicity in *Prochlorococcus* and *Synechococcus* increased protist predation, and studies carried out using a mutant of the soil bacterium *Bradyrhizobium japonicum* lacking the LPS O-polysaccharide showed an increase in hydrophobicity compared to the wild type (Park and So, 2000).

Moreover, cell surface hydrophobicity has been implicated as a possible factor for selective grazing (Matz and Jürgens, 2001; Monger et al., 1999) while the 'food quality' of the prey, in terms of C:N:P ratio, may be another discriminating factor (John and Davidson, 2001; Shannon et al., 2007).

In marine *Synechococcus*, changes in cell surface properties can be directly related to genome organization, since the genes involved in carbohydrate modification, such as glycosyltransferases, glycoside hydrolase gene families and other enzymes possibly related to O-polysaccharide biosynthesis, are located in genomic islands, variable parts of the genome showing a deviation in the tetranucleotide frequency compared with the core genome, and which could be the result of lateral gene transfer events (Dufresne et al., 2008).

Although the case for *Synechococcus* sp. WH7803 seems to be more complicated than only modifications in the LPS to avoid grazing and/or phage infection, an evident relation between protist grazing and phage infection has been observed, but little is known about it. In this regard, recent research in the Scanlan lab showed that in the absence of phage infection, the phagotrophic nanoflagellate *Cafeteria roenbergensis* ingested both *Synechococcus* sp. WH7803 and *Synechococcus* sp. WH8102, but only *Synechococcus* sp. WH8102 supported growth of the flagellate. Nevertheless, once *Synechococcus* sp. WH7803 was infected with cyanophage S-RSM4, *C. roenbergensis* cell numbers again increased significantly, implying both ingestion and digestion of this cyanophage-infected *Synechococcus* (Zwirglmaier et al., unpublished data). This change in palatability of the prey could have an important ecological role, related to the control of phage abundance by removing infected cells before they burst, which opens the question regarding the universality of this phenomena.

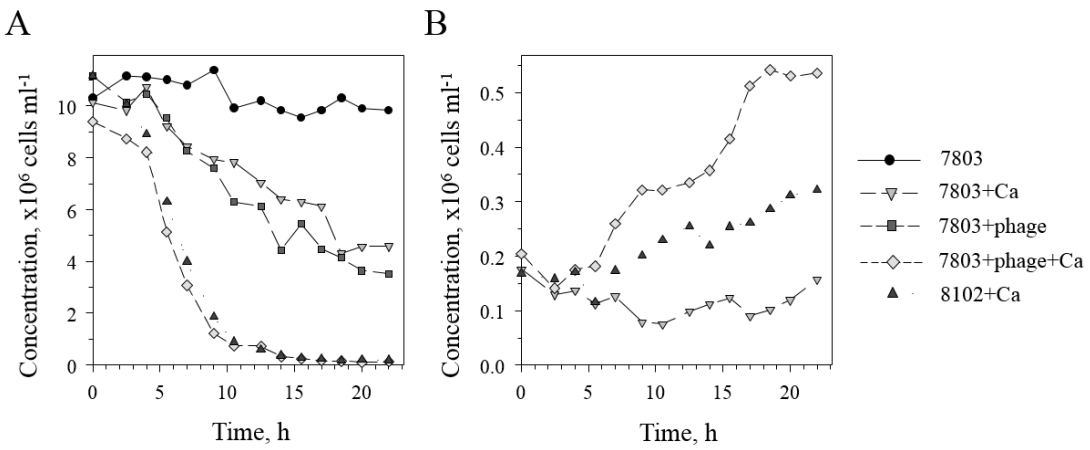


Figure 7.3 Grazing of *Synechococcus* spp. by *C. roenbergensis* predator (Zwirglmaier et al., unpublished data).

(A) Prey concentrations: *Synechococcus* sp. WH7803 (7803), *Synechococcus* sp. WH7803 in the presence of *C. roenbergensis* (7803+Ca); cyanophage-infected *Synechococcus* sp. WH7803 (7803+phage); cyanophage-infected *Synechococcus* sp. WH7803 in the presence of *C. roenbergensis* (7803+phage+Ca); *Synechococcus* sp. WH8102 in the presence of *C. roenbergensis* (8102+Ca). (B) Concentration changes of *C. roenbergensis* fed on *Synechococcus* sp. WH7803 (7803+Ca), infected *Synechococcus* sp. WH7803 (7803+phage+Ca) or *Synechococcus* sp. WH8102 (8102+Ca, positive control) as food source. Data from one representative experiment is shown (Zwirglmaier et al., unpublished data).

To examine the interactions between the two major biotic causes of bacterial mortality in the marine pelagic water column, viral lysis and protistan grazing, affecting trophic dynamics in the oceans and planetary biogeochemical cycles, it is necessary to study the mechanisms underpinning their interactions, particularly focusing on the role that the host cell surface plays in this process. Given that marine *Synechococcus* appear to have the genetic potential for a variety of grazer evasion capabilities (Jones et al., 2006), an approach based on biochemical and molecular biology techniques should give important insights into this problem.

In this chapter, grazing of *Synechococcus* sp. WH7803 (as prey) by the nanoflagellates *C. roenbergensis* and *P. danica* was studied. In order to confirm and study the generality of the phenomena observed by Zwirglmaier et al. (2009), grazing of *Synechococcus* sp. WH7803 was investigated in comparison with cyanophage infected and resistant prey as well as to *Synechococcus* sp. WH7805.

7.2 Objective

The aim of this work was to investigate the relationship between phage infection/resistance and grazing of *Synechococcus* sp. WH7803 by the nanoflagellates *C. roenbergensis* and *P. danica*. Also, grazing of cyanophage infected *Synechococcus* sp. WH7805 was also studied as a comparison grazer-prey system.

7.3 Results and discussion

7.3.1 Determination of growth conditions of *P. danica* and *C. roenbergensis* for grazing experiments

The protist grazers used in this work were *Pteridomonas danica* and *Cafeteria roenbergensis* (Fig. 7.4), as described in section 2.2.5.

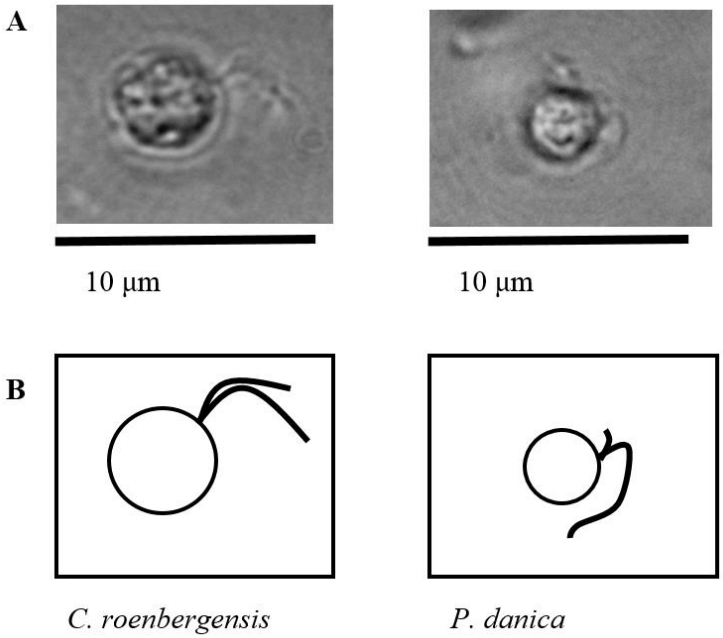


Figure 7.4. Structure of the protist grazers *C. roenbergensis* and *P. danica*.

(A) Light microscopy. (B) Schematic diagram of the microscopy image in (A) showing the two flagella.

Culturing of these organisms was not trivial since their feeding behaviour is poorly known and growth rates depend on the food source available and with grazing decreasing at high nutrient concentrations in the medium. In order to standardise the culturing and experimental conditions, growth curves using different cell concentration of prey and grazer were carried out (Fig. 7.5). Initial experiments used the heterotrophic bacterium *Halomonas* sp. as the prey strain. *Halomonas* sp. has been used before for grazing experiments performed in foraminifera (Mojtahid et al., 2011) and shown to be a 'good quality' food supporting growth of *P. danica* (Fig. 7.5) and *C. roenbergensis* (data not shown).

The high variability between biological and technical replicates was a constant problem during this work and hindered the reproducibility and interpretation of data. In Fig. 7.5 the growth of *P. danica* with different concentrations of prey and grazer cells is shown as a representative of both grazers, i.e. *P. danica* and *C. roenbergensis*.

The optimal concentration of grazers for grazing experiments was between 1×10^4 and 5×10^5 cells mL^{-1}, as shown by an increase in grazer concentration linked with a decrease in *Halomonas* prey cell concentration (Fig 7.5.B). The prey concentration selected for grazing experiments was $\sim 10^7$ cells mL^{-1} because although there was a larger increase in grazer cell concentration when using 5×10^6 cells mL^{-1} prey, the prey concentration remained stable during the experiment showing that either the growth rate of the prey was counteracting mortality by grazing or that the grazers were feeding on nutrients or other bacteria present in the medium.

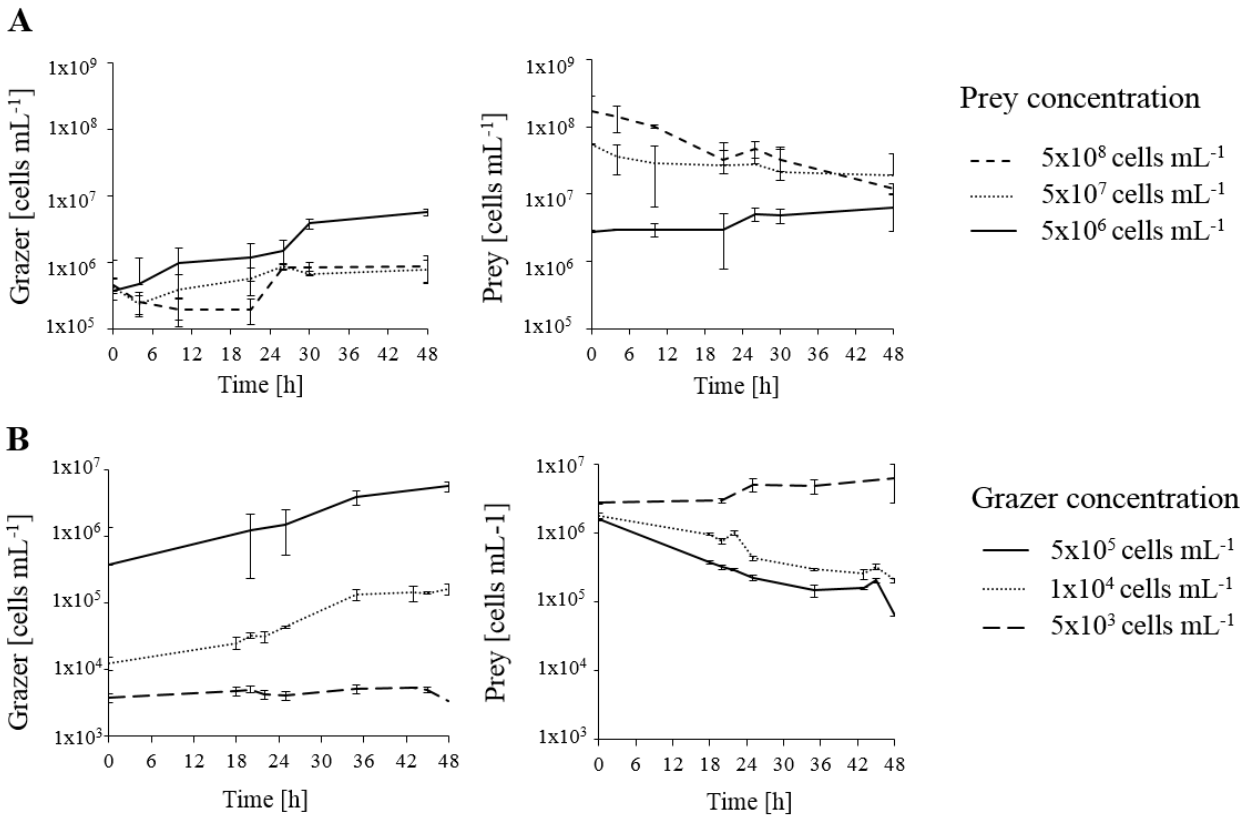


Figure 7.5. Growth curve of *P. danica* with *Halomonas* sp. as food source.

The starting culture had different prey (A) or grazer (B) cell concentrations. (A) Starting grazer concentration: 5 x 10^5 cells mL^{-1}; (B) Starting prey concentration: 5 x 10^6 cells mL^{-1}. These experiments were performed at 23°C, although experiments at 13°C showed the same results with a timescale of days instead of hours. Error bars indicate standard deviation of 3 biological replicates.

Moreover, prey and grazer cell size also plays an important role in terms of feeding behaviour. *Halomonas* sp., used in these initial experiments as prey to optimise grazer feeding conditions, varies in size and tends to elongate when grown in liquid conditions, so different pore size filters (0.8 and 1.2 μm) were used to determine the effect of prey size on grazing rates. The size of the filters was selected based on the size of marine *Synechococcus* (~0.8 μm) and *Halomonas* sp. in liquid medium (> 1 μm). The resulting growth curves, described in Fig. 7.6, showed no difference related to the filters used.

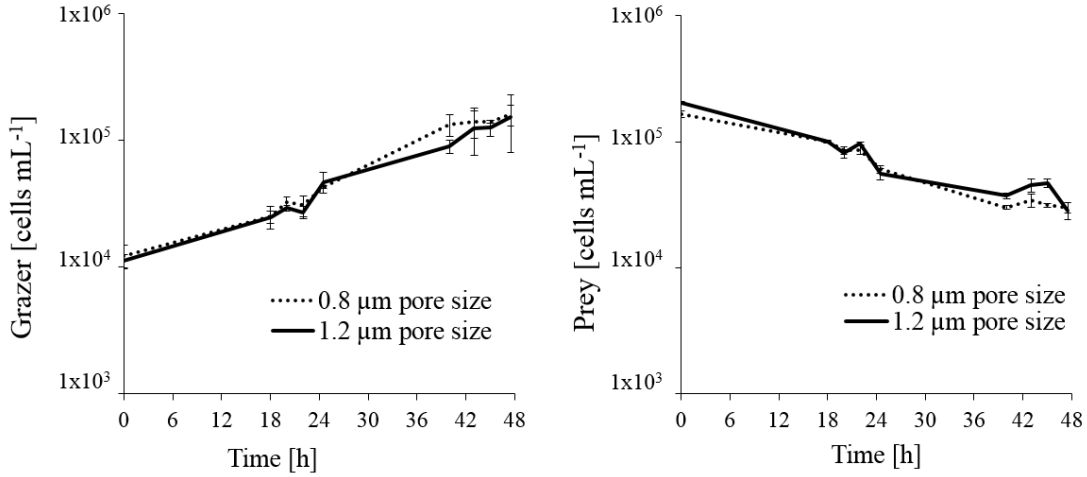


Figure 7.6. Growth curve of *P. danica* cultured with *Halomonas* sp. filtered through different pore size filters.

Error bars indicate standard deviation of 3 biological replicates.

Protist cell size, as determined using microscopy, showed high variability for *C. roenbergensis* cells, from 1 to 10 μm, whilst *P. danica* showed a more stable size around 3 μm. In order to achieve a homogeneous grazer cell size and to eliminate any remaining *Halomonas* cells on which the grazers are maintained, *P. danica* and *C. roenbergensis* cells were filtered through a 2 μm pore-size membrane (Fig. 7.7), which was selected based on microscopy and published data on the cell size of these organisms (Boenigk, 2002; Mohapatra and Fukami, 2004; Sekiguchi et al., 2002). The population of *P. danica*, as visualised using flow cytometry, is relatively uniform before filtration (Fig. 7.7.A) but showed an increase in both forward and side scatter

when analysing the filter (Fig. 7.7.B), suggesting clumping of cells, and almost no cells passed to the flow through (Fig. 7.7.C). All this indicates that this nanoflagellate has a homogeneous size. In contrast, *C. roenbergensis* showed a wide population size before filtration (Fig. 7.7.A) that gets tighter in the filter (Fig. 7.7.B) but with most of the cells passing through it (Fig. 7.7.B). This result confirms the observations made by microscopy, where *P. danica* has a stable cell size of ~3 µm whilst *C. roenbergensis* has a higher variation in size.

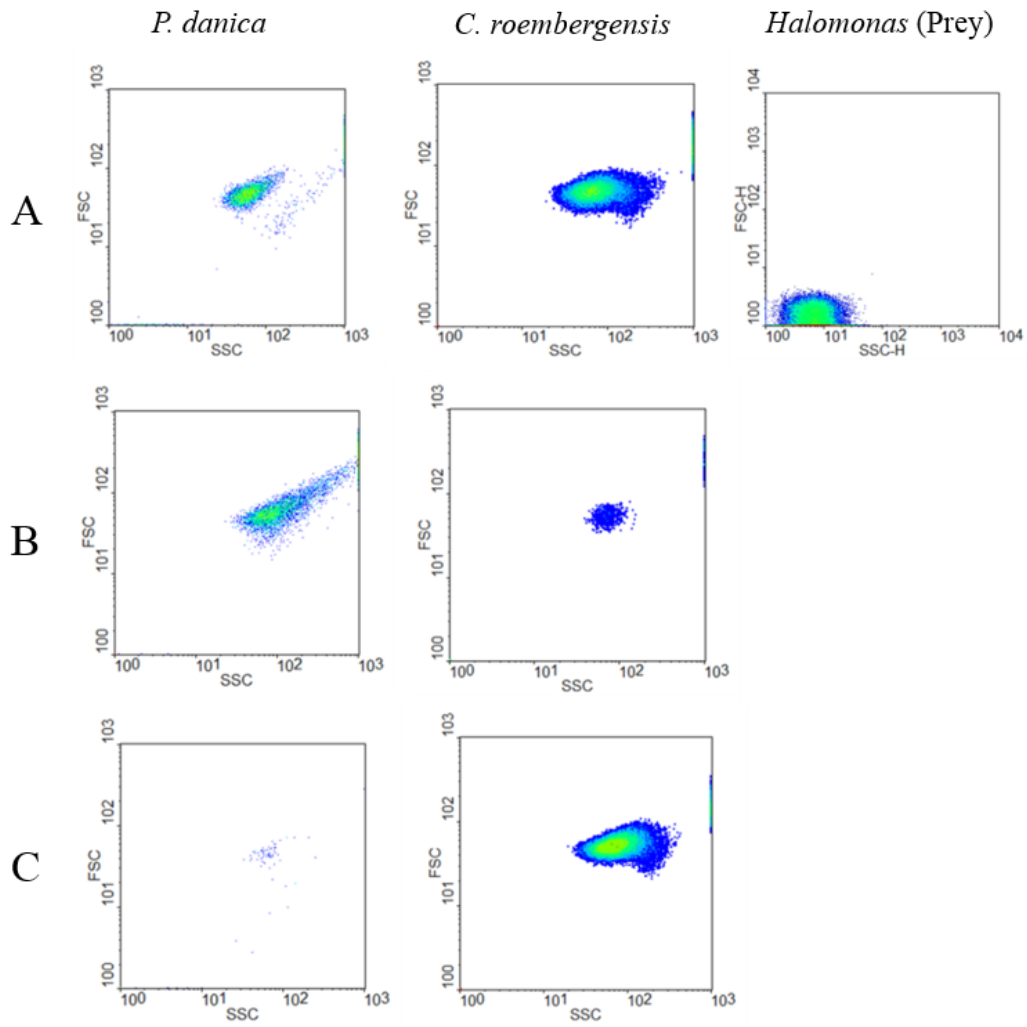


Figure 7.7. Size standardisation of *P. danica* and *C. roenbergensis* by filtration.

Flow cytometry analysis of cultures of *P. danica* (left column) and *C. roenbergensis* (middle column) filtered by a 2 µm pore size membrane. (A) before filtration, (B) cells in the filter after filtration and (C) cells in the flow through. A culture of *Halomonas* sp. (right column) is shown as size reference. FSC: forward scatter, SSC: side scatter.

Nonetheless, filtration of the grazers resulted in almost complete elimination of *Halomonas* sp. from the culture (data not shown), but since the feeding mechanism of *P. danica* and *C. roenbergensis* includes the use of a flagella to attract prey (Boenigk, 2002) filtration caused loss of the flagella resulting in a lag period before grazing resumed again. For this reason all subsequent grazing experiments carried out did not involve filtration of the grazers and instead cultures with the lowest numbers of *Halomonas* present were used.

7.3.2 Grazing of nanoflagellates on the cyanophage resistant *Synechococcus* sp. WH7803 mutant R2

Previously, Zwirglmaier et al. (2009) showed that grazing of the cyanophage resistant *Synechococcus* sp. WH7803 mutant PHR, resistant to cyanophage S-PM2, was higher than the wild type. Here, I extended this work by carrying out grazing experiments with the *Synechococcus* sp. WH7803 mutant R2, also resistant to S-PM2 and with a different LPS profile compared to the wild type (see section 3.3.1). This mutant was selected due to its phenotypic similarities with PHR, i.e. resistant to the same phage, an LPS profile different from the wild type, and high growth rate compared to the other cyanophage resistant mutants.

The results of grazing experiments on mutant R2 compared to the wild type are shown in Fig. 7.8. As explained above, the high variability of biological and technical replicates restricted the analysis, so individual replicates were studied separately.

Synechococcus sp. WH8102 was used as a positive control since this strain was already known to be grazed well by both nanoflagellates (Zwirglmaier et al., 2009). It is important to note that these experiments were performed over a 96 h period, compared to 27 h in the Zwirglmaier et al. (2009) study.

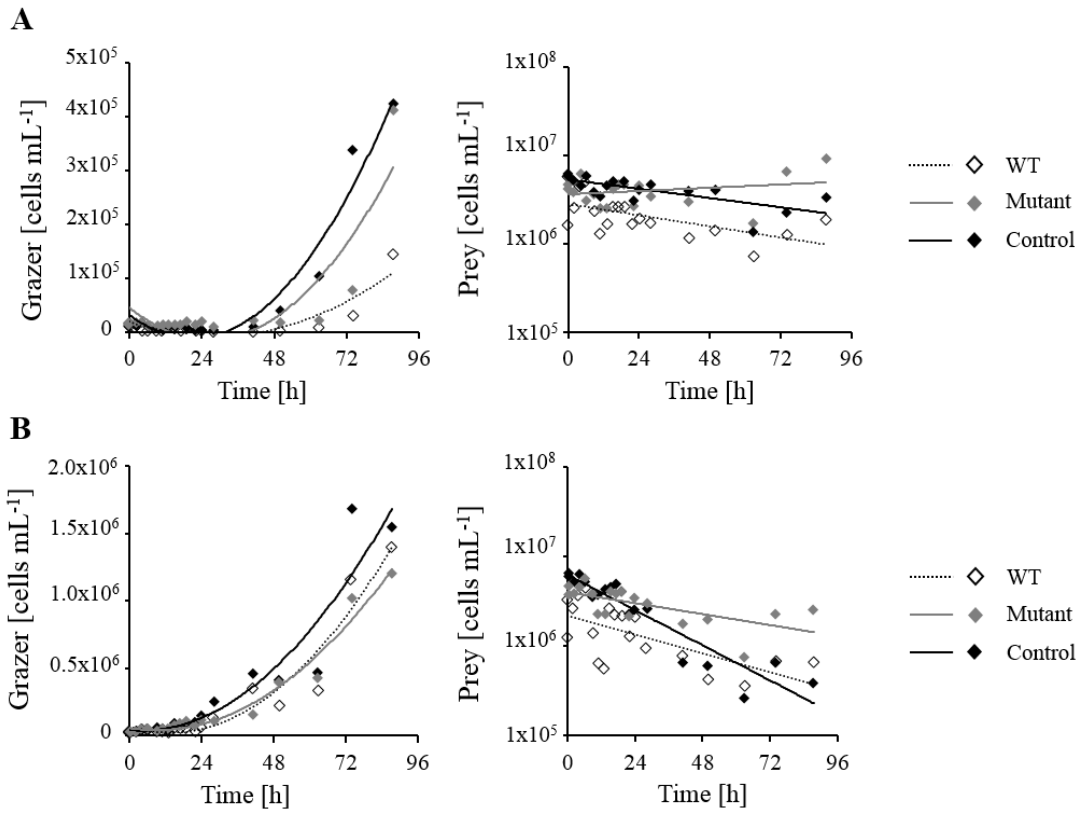


Figure 7.8. Grazing of cyanophage resistant *Synechococcus* sp. WH7803 mutant by *P. danica* and *C. roenbergensis*

Grazing of *P. danica* (A) and *C. roenbergensis* (B) fed with cyanophage resistant *Synechococcus* sp. WH7803 mutant R2 (Mutant), *Synechococcus* sp. WH7803 wild type (WT) or *Synechococcus* sp. WH8102 (Control) as a positive control for grazing. Data from one representative experiment and replicate is shown with lines of best fit.

Analysis of these grazing experiments showed that there was a clear difference in growth of *P. danica* and *C. roenbergensis* (Fig. 7.8 and Table 7.1). This difference was observed in a higher growth rate (up to 2.9 day^{-1}) and lower yield (up to 0.07) of *P danica* compared to *C. roenbergensis* (up to 1 day^{-1} and 6.3 respectively, see Table 7.1).

Grazing of *P. danica* resulted in the same yield but a higher growth rate when feeding on the cyanophage resistant mutant R2 compared to the wild type *Synechococcus* sp. WH7803 (Table 7.1) whilst *C. roenbergensis* showed a similar growth rate but a lower yield when feeding on the wild type *Synechococcus* sp. WH7803 compared to the mutant R2 (Table 7.1).

Moreover, the R2 mutant prey cell concentration increased over time, indicating that either *P. danica* was feeding on heterotrophic bacteria associated with the mutant or that the mutant growth rate was faster than the wild type in these conditions. Noteworthy, the growth rate of *P. danica* feeding on the R2 mutant was higher than the positive control, i.e. *Synechococcus* sp. WH8102, but the lag phase was 20 h longer. This could be due to the accumulation of heterotrophic bacteria, i.e. better 'quality food' supporting growth of the grazer.

Nonetheless, in spite of this data being a representative of the results obtained for several grazing experiments, this is the analysis of one experiment and replicate, due to the high variability between biological and technical replicates, and thus conclusions should be made carefully. Likewise, due to high variability and unstable growth of *C. roenbergensis* most grazing experiments were carried out with *P. danica*.

	P. danica		*C. roenbergensis*	
	μ [day^{-1}]	yield	μ [day^{-1}]	yield
WT	1.9	0.01	0.9	1.0
Mutant	2.9	0.01	1.0	0.6
Control	2.6	0.07	1.0	6.3

Table 7.1. Growth rate and yield of *P. danica* and *C. roenbergensis* feeding on different *Synechococcus* prey

Growth parameters are of one grazing experiment and replicate of *P. danica* and *C. roenbergensis* (see Fig. 7.8) feeding on *Synechococcus* sp. WH7803 (WT), *Synechococcus* sp. WH7803 mutant R2 (Mutant) and *Synechococcus* sp. WH8102 (Control). Growth rate (μ) was calculated during exponential growth. Growth yield was calculated as the number of grazer cells produces per number of prey cells before day three (72 h), where the prey cell numbers started to increase.

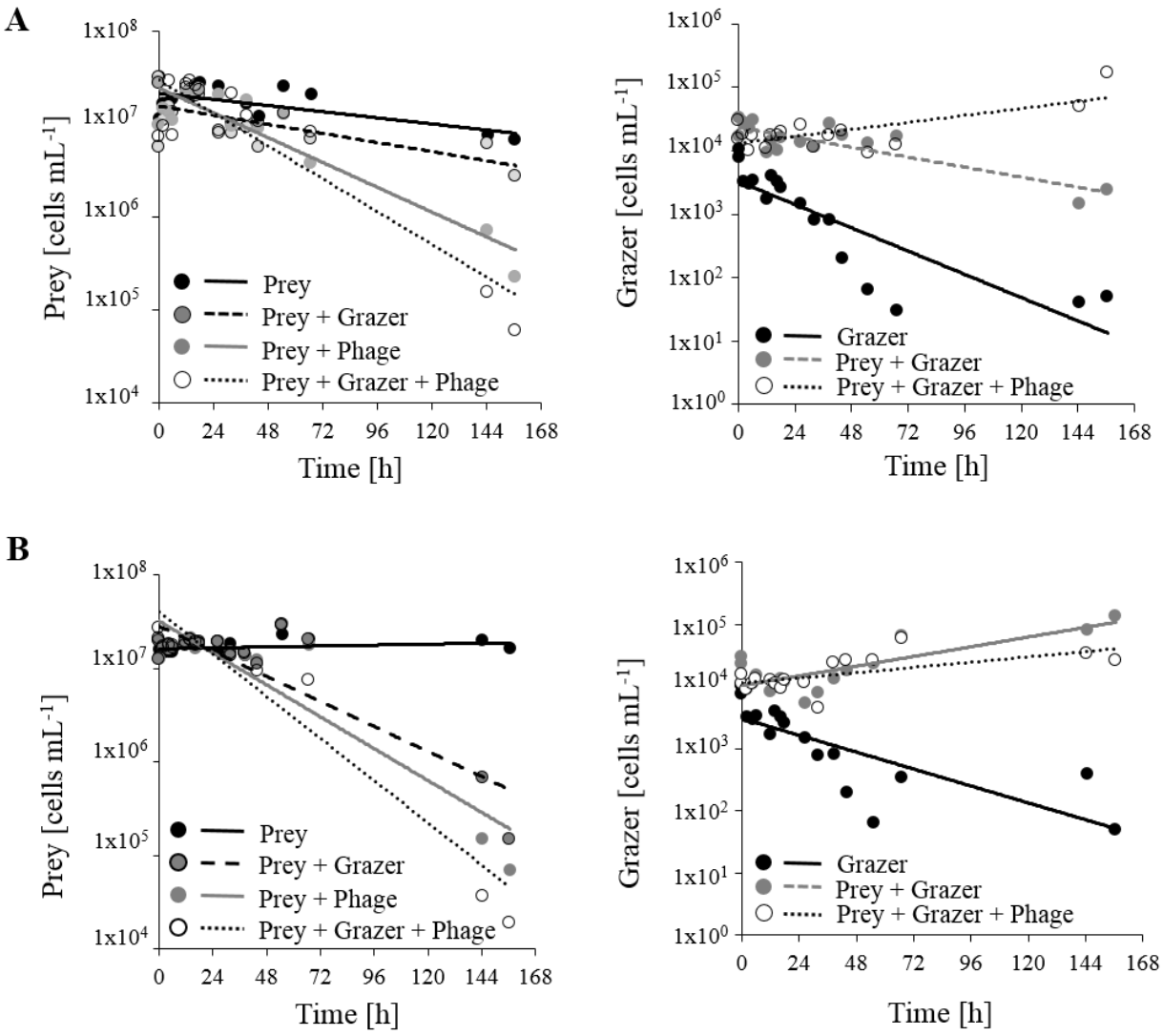
7.3.3 Grazing of cyanophage infected *Synechococcus* sp. WH7803 and *Synechococcus* sp. WH7805 by *P. danica*

To study the relationship between cyanophage infection and grazing of marine *Synechococcus,* grazing experiments of infected compared to uninfected prey were performed.

Grazing of *Synechococcus* sp. WH7803 can be observed in Fig. 7.9.A with a plummet in prey concentration corresponding to an increase in grazer concentration, compared to each individual predators separately, i.e. grazer and phage. The rise in *P. danica* concentration when feeding on infected prey shows that the decrease in prey concentration is not just a combined effect of both predators but the infected prey is supporting growth of the grazer. Fig. 7.9.B shows grazing on *Synechococcus* sp. WH8102 used as a positive control based on Zwirglmaier et al. (2009) (see Fig. 7.1).

Noteworthy again is the time scale of the experiment, which is ~7 days compared to 1 day used by Zwirglmaier et al. (unpublished data). This time scale is important since this may affect the grazing results given that the cyanophage infection cycle is probably shorter than that. Although the length of the infection cycle for S-RSM4 is not known, for cyanophage S-PM2 it has been determined to be ~9 h (Wilson et al., 1996). This would mean that infection of *Synechococcus* sp. WH7803 is faster than the grazing rate, releasing cyanophages and organic matter, i.e. lysed cells, into the medium. Unfortunately, there is no published information about whether *P. danica* is a restricted bacterivore or if it can acquire nutrients dissolved in the media or feed on phages, though it has been shown that nanoflagellates obtained from natural environments can feed on viruses and virus-like particles (Gonzalez and Suttle, 1993). Thus, the observed increase in grazer abundance cannot exclude growth due to acquisition of nutrients as a result of prey lysis, rather than actual ingestion and digestion of phage-infected *Synechococcus* prey.

.


A
Prey [cells mL-1]
1x10^8
1x10^7
1x10^6
1x10^5
1x10^4
0 24 48 72 96 120 144 168
Time [h]
Prey
Prey + Grazer
Prey + Phage
Prey + Grazer + Phage
Grazer [cells mL-1]
1x10^6
1x10^5
1x10^4
1x10^3
1x10^2
1x10^1
1x10^0
0 24 48 72 96 120 144 168
Time [h]
Grazer
Prey + Grazer
Prey + Grazer + Phage
B
Prey [cells mL-1]
1x10^8
1x10^7
1x10^6
1x10^5
1x10^4
0 24 48 72 96 120 144 168
Time [h]
Prey
Prey + Grazer
Prey + Phage
Prey + Grazer + Phage
Grazer [cells mL-1]
1x10^6
1x10^5
1x10^4
1x10^3
1x10^2
1x10^1
1x10^0
0 24 48 72 96 120 144 168
Time [h]
Grazer
Prey + Grazer
Prey + Grazer + Phage

(Previous page)

Figure 7.9. Grazing of infected and uninfected *Synechococcus* sp. WH7803 by *P. danica*

Grazing of *P. danica* (Grazer) fed with (A) *Synechococcus* sp. WH7803 (Prey) or (B) *Synechococcus* sp. WH8102 (Prey, positive control) uninfected or infected with the cyanophage S-RSM4 (Phage). The selection of S-RSM4 for this experiment was because it infects *Synechococcus* sp. WH7803 and WH8102. Data from one representative experiment and replicate is shown with lines of best fit.

Moreover, since *Synechococcus* sp. WH7805 was shown to be poorly grazed Zwirglmaier et al. (2009) and because it is phylogenetically close but genetically distinct from *Synechococcus* sp. WH7803, as discussed in previous chapters, grazing experiments were carried out using *Synechococcus* sp. WH7805, *P. danica* and cyanophage S-RIM34 to determine if the same phenomenon of increased palatability of a phage-infected prey was shown with another *Synechococcus* strain.

In this experiment there does not seem to be a difference of grazing on infected compared to uninfected cells (Fig. 7.10). This again can be related to the time scale of the experiment (in this case ~10 days) and the low preference of *P. danica* for *Synechococcus* sp. WH7805. Furthermore, a grazer control, using grazer fed only with phages, was included, but no difference was observed compared to the sample with the grazer alone.

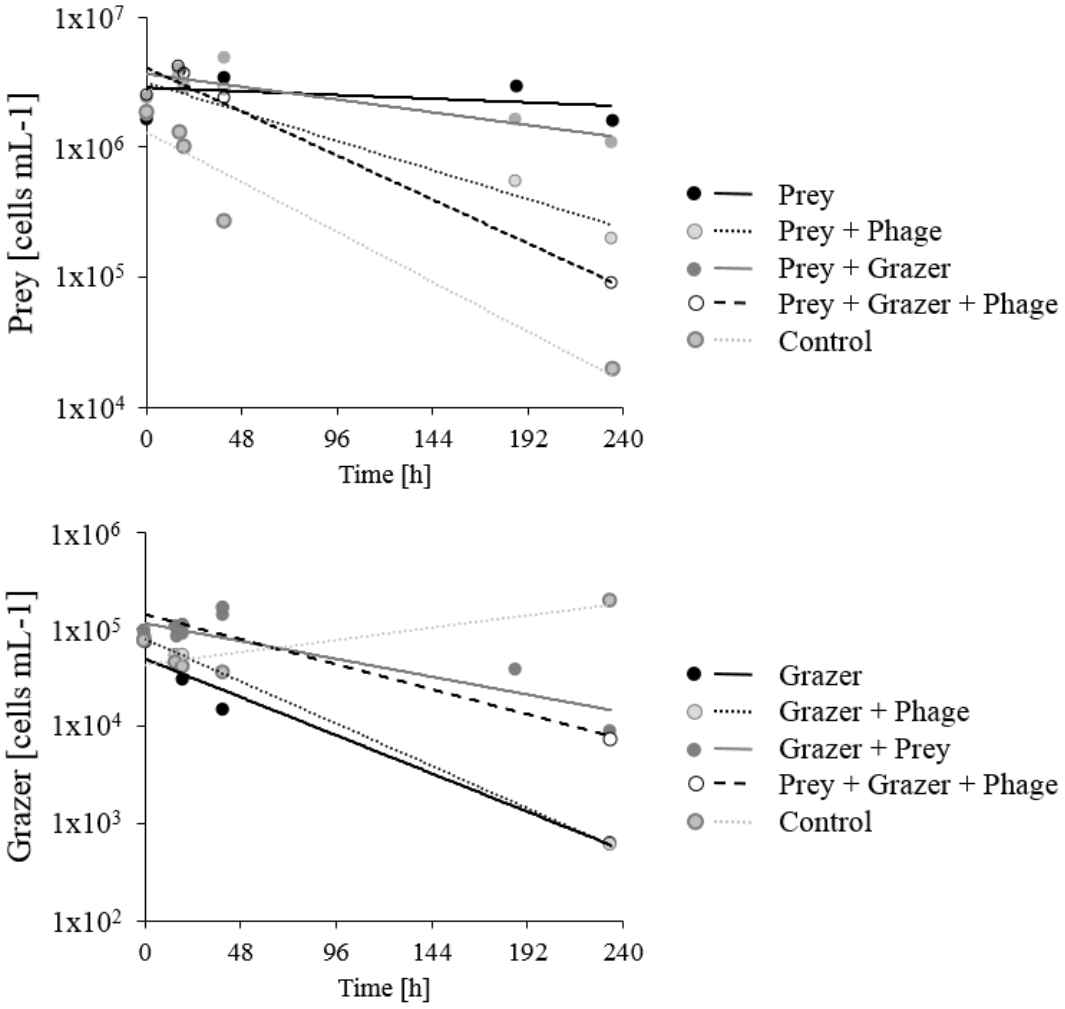


Figure 7.10 Grazing of infected and uninfected *Synechococcus* sp. WH7805 with cyanophage S-RIM34 by *P. danica*

Grazing of *P. danica* (Grazer) fed with *Synechococcus* sp. WH7805 (Prey) uninfected or infected with the cyanophage S-RIM34 (Phage). *Synechococcus* sp. WH8102 was used as a positive control for grazing (Control). Data from one representative experiment and replicate is shown with lines of best fit.

The results of grazing experiments with the nanoflagellates *P. danica* and *C. roenbergensis* feeding on *Synechococcus* resulted in inconclusive analyses due to the high variability of the system. For this reason, it is very important that future work to understand the relationship between grazing and viral infection of *Synechococcus* focus on the development of a new system, perhaps using grazers with higher growth rates and lower biological variability. Furthermore, the inclusion of the cyanophage

resistant *Synechococcus* sp. WH7805 mutants here characterised (see Chapter 6) could help resolve the generality of the phenomenon observed by Zwirglmaier et al. (2009).

Chapter 8

CONCLUDING REMARKS AND FUTURE DIRECTIONS

Here, the results of this thesis are discussed in the context of the original aims highlighted in section 1.7. Further, future approaches are considered that will allow for a better understanding of resistance mechanisms of marine *Synechococcus* to their co-occurring cyanophages.

Most notably, as a result of work performed in this thesis, a new angle on cyanophage resistance in marine *Synechococcus* was identified, implicating a role for ploidy status in potentially adding 'flexibility' to mechanisms of resistance that have not previously been described for bacteria.

In Chapter 3, cyanophage resistant *Synechococcus* sp. WH7803 mutants were phenotypically characterised, demonstrating that not all were adsorption mutants, which is the only known cyanophage-resistant mechanism so far reported for marine *Synechococcus* and *Prochlorococcus* (see Avrani et al. 2011; Marston et al. 2012) but rather these are likely to be replication defective mutants, i.e. affecting phage replication at some point after adsorption.

In those mutants that were adsorption defective, resistance was hypothesised to be linked to modification of cell surface properties, an idea supported by data from a recent *Prochlorococcus* study (Avrani et al., 2011) which found mutations located in genes related to cell wall and LPS biogenesis and membrane proteins. Unfortunately, none of these genes was confirmed to be the genetic basis of cyanophage resistance in these mutants due to the lack of a suitable genetic system in marine *Prochlorococcus* preventing a reverse genetic approach that could re-create such a phenotype.

In this thesis, characterisation of *Synechococcus* sp. WH7803 interposon mutants in the *synWH7803_0192* and *synWH7803_1767* genes, both involved in LPS biosynthesis, was performed, a strain where a genetic system is available (see Brahamsha 1996). However, mutants failed to segregate suggesting an essential nature of the genes concerned but not allowing any specific conclusion on the role of these genes in cyanophage resistance to be ascertained.

Analysis by WGS of cyanophage resistant *Synechococcus* sp. WH7803 mutants in Chapter 4 demonstrated a large number of mutations occurring in each mutant, as well as the re-sequenced wild type strain. This large number of mutations was mirrored by

mutations occurring at varying frequencies. However, even when just assessing mutations occurring at > 90 % frequency, no clear differences between mutants and the published genome of *Synechococcus* sp. WH7803 were observed. This phenomenon of a large number of mutations occurring at varying frequency was also observed using a previously published *Synechococcus* sp. WH7803 WGS dataset (Marston et al. 2012), casting doubt on the conclusions drawn in this manuscript. In contrast, use of the same bioinformatics pipeline with a previously published *Prochlorococcus* WGS dataset of a cyanophage resistant mutant (Avrani et al. 2011) gave results in accordance with those published, i.e. identified specific mutations occurring at 100 % frequency that could putatively be the molecular basis for resistance. Given the monoploid nature of *Prochlorococcus* but the oligoploid nature of *Synechococcus* sp. WH7803 led to the development of a hypothesis that ploidy level was somehow playing a role in cyanophage resistance in this organism.

To precisely determine the ploidy level in *Synechococcus* sp. WH7803 and a related strain *Synechococcus* sp. WH7805, a qPCR assay was optimised (Chapter 5) which showed that *Synechococcus* sp. WH7803 possessed on average four copies of its genome, whereas *Synechococcus* sp. WH7805 was monoploid. Subsequent growth phase experiments showed no significant variation in copy number as a function of early- mid- or late-exponential phase of growth leaving open the question whether ploidy status varies at all in *Synechococcus* sp. WH7803 or if it stays constant. Likewise, conclusions following assessing chromosome copy number during P-deplete growth suggest that chromosome copy number does not change, albeit in a manner perhaps contradicting logic since a decrease in P availability would suggest P would limit DNA synthesis and hence decrease ploidy level. Further work assessing ploidy level as a function of various environmental parameters is clearly needed here, but regardless of whether ploidy level is constant or varies as a function of the environment, this would imply some mechanism to maintain this number. How this is done is unknown, though there is some evidence in the freshwater cyanobacterium *Synechococcus* sp. PCC7942 that chromosomes are spatially ordered along the long axis of the cell, which along with accurate mid-cell septum placement allows for near-optimal segregation of chromosomes to daughter cells (Jain et al., 2012). In contrast,

in *Synechocystis* sp. PCC6803 segregation of chromosomes to daughters is random (Schneider et al., 2007).

Interestingly, in the qPCR experiments determining chromosome copy number performed here, a relatively wide variation between experimental replicates was observed, which was higher in *Synechococcus* sp. WH7803 compared to *Synechococcus* sp. WH7805 (see Table 5.3). This variation is likely related to the fact that the qPCR assay uses thousands of cells and provides an average DNA content per cell, rather than assessing the chromosome number in an individual cell. This variation suggests that individual cells may well have a different number of chromosomes, which on average is around four copies in *Synechococcus* sp. WH7803. Indeed, such variation would be in accordance with flow cytometry estimation of chromosome copy number following DNA staining which suggests that 1-10 copies of the chromosome occur in this organism (Binder and Chisholm, 1995). Though technically challenging, future work should be directed to determining chromosome copy number in individual *Synechococcus* sp. WH7803 cells. Furthermore, if WGS of single cells without DNA amplification could be performed this could help identify mutations at the single cell level. This would address whether individual cells in a population possess the same mutations with the same frequency or if there is variation between individuals. If the latter, is the process regulated and how?

Certainly the monoploid nature of *Synechococcus* sp. WH7805 suggests there is a marked difference in chromosome copy number regulation compared to *Synechococcus* sp. WH7803. This opens up several avenues of work comparing DNA replication and segregation mechanisms in these phylogenetically closely related strains. Potentially more interesting would be to assess the ploidy status across many more marine *Synechococcus* strains to assess how widespread such a feature is, and if possible to directly investigate the ploidy status of natural *Synechococcus* populations, which if observed would imply that ploidy was prevalent and important in nature.

Moreover, WGS analysis of cyanophage resistant mutants of the monoploid strain *Synechococcus* sp. WH7805 (Chapter 6) showed marked ‘simplicity’ in the mutations identified i.e. compared to that seen in *Synechococcus* sp. WH7803. Such ‘simplicity’ identified a gene (WH7805_08977) encoding a possible alpha-glycosyltransferase in

several of the mutants but not the wild type, which was confirmed by Sanger sequencing, that likely forms the molecular basis of cyanophage resistance in these mutants. Indeed, such a conclusion is in agreement with biochemical characterisation of these cyanophage resistant *Synechococcus* sp. WH7805 mutants, which also showed corresponding differences in LPS profiles and cyanophage adsorption efficiencies (section 6.3.2). That a potential glycosyltransferase is playing a role here is similar to that recently observed for *Prochlorococcus* cyanophage resistant mutants (Avrani et al., 2011) where mutations in glycosyltransferases and enzymes involved in LPS biosynthesis have been implicated. Taken together, the *Synechococcus* sp. WH7805 and *Prochlorococcus* WGS profiles of these monoploid strains supports the hypothesis that oligoploidy is responsible for the high variability and complex nature of the observed mutation frequencies found in *Synechococcus* sp. WH7803.

Further evidence for a role of ploidy in this process, potentially permitting genome flexibility, might be gleaned from the length of time required to initially obtain cyanophage resistant mutants for each of these *Synechococcus* strains. Thus, growth of surviving (cyanophage-resistant) *Synechococcus* sp. WH7805 mutants occurred after ca. 1 month whereas for *Synechococcus* sp. WH7803 mutants were only obtained ca. 1 year after infection (Spence, 2010). This might suggest that the multiple chromosomes present in *Synechococcus* sp. WH7803 prevent fixation of resistance-conferring mutations, which would not be possible in the monoploid strain *Synechococcus* sp. WH7805.

The advantages of this mechanism are arguable, although the cost of resistance seems to be slightly more pronounced in cyanophage resistant *Synechococcus* sp. WH7805 mutants, where growth rate decreased by up to 50 % (see Table 6.2) compared to 40 % in cyanophage resistant *Synechococcus* sp. WH7803 mutants (Spence, 2010). However, the need for repeated addition of cyanophage (ca. every three weeks) to the *Synechococcus* sp. WH7803 cyanophage resistant mutants but not to the *Synechococcus* sp. WH7805 mutants, to prevent them from lysing after transferring to fresh medium (data not shown) suggests that a lack of selective pressure soon reverts genotypes to wild type making them susceptible to infection.

Nevertheless, the work performed in this thesis on cyanophage resistant *Synechococcus* mutants is consistent with the 'Kill the Winner' hypothesis (see Section 1.6) in which the presence of the cyanophage (parasite) increases the diversity of the co-occurring host through resistance trade-offs. In this case, the trade-offs were observed as a reduced growth rate of the cyanophage resistant mutants and different LPS and adsorption profiles, as well as a clearly discernible clumping phenotype.

More broadly in future work it will be important to consider cyanophage-resistance and ploidy in the context of the complexity of natural marine microbial systems where *Synechococcus* 'interacts' with various other organisms, e.g. does ploidy play a role in grazing avoidance? Grazing of marine *Synechococcus* by nanoflagellates has been shown to be strain-selective Zwirglmaier et al. (2009), indicating that the grazer is somehow able to detect 'differences' between *Synechococcus* strains. Furthermore, cyanophage infection and resistance appears linked to grazing preference by these nanoflagellates (Zwirglmaier et al., 2009; Zwirglmaier et al., unpublished), although in work performed here (Chapter 7) the high biological variability in this predator-prey system prevented further insights to be made. To address this problem, future work should focus on establishing a 'stable' grazer-prey-cyanophage system that allows high reproducibility of experiments studying feeding behaviour in the context of cyanophage infection and resistance. Certainly, it is unknown how *Synechococcus* balances the cost of cyanophage resistance with increased grazing susceptibility Zwirglmaier et al. (2009) or the role that ploidy might play in genome flexibility/adaptability to these two mortality factors. As yet, no 'grazing resistant mutants' have been isolated, but cell surface properties are likely to play a role (see Zwirglmaier et al. 2009) and there may well be an additional role for 'ploidy' status here.

REFERENCES

Abedon, S.T. (2011). Lysis from without. Bacteriophage *1*, 46–49.

Abeyrathne, P.D., Daniels, C., Poon, K.K.H., Matewish, M.J., and Lam, J.S. (2005). Functional characterization of WaaL, a ligase associated with linking O-antigen polysaccharide to the core of *Pseudomonas aeruginosa* lipopolysaccharide. Journal of Bacteriology *187*, 3002–3012.

Ackermann, H.W. (2007). 5500 Phages examined in the electron microscope. Archives of Virology *152*, 227–243.

Ackermann, H.W. (2011). Bacteriophage taxonomy. Microbiology Australia *32*, 90–94.

Ackermann, H.W., and DuBow, M.S. (1987). Viruses of Prokaryotes: General properties of bacteriophages (CRC Press).

Agawin, N.S.R., Duarte, C.M., and Agustí, S. (1998). Growth and abundance of *Synechococcus* sp. in a Mediterranean Bay: Seasonality and relationship with temperature. Marine Ecology Progress Series *170*, 45–53.

Agawin, N.S.R., Duarte, C.M., and Agustí, S. (2000). Nutrient and temperature control of the contribution of picoplankton to phytoplankton biomass and production. Limnology and Oceanography *45*, 591–600.

Ahlgren, N.A., and Rocap, G. (2006). Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. Applied and Environmental Microbiology *72*, 7193–7204.

Angly, F.E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. a., Carlson, C., Chan, A.M., Haynes, M., Kelley, S., Liu, H., et al. (2006). The marine viromes of four oceanic regions. PLoS Biology *4*, 2121–2131.

Apple, J.K., Strom, S.L., Palenik, B., and Brahamsha, B. (2011). Variability in protist grazing and growth on different marine *Synechococcus* isolates. Applied and Environmental Microbiology *77*, 3074–3084.

Armbrust, E.V., Bowen, J.D., Olson, R.J., and Chisholm, S.W. (1989). Effect of light on the cell cycle of a marine *Synechococcus* strain. Applied and Environmental Microbiology *55*, 425–432.

Arrigo, K.R. (2005). Marine microorganisms and global nutrient cycles. Nature *437*, 349–355.

Avrani, S., and Lindell, D. (2015). Convergent evolution toward an improved growth rate and a reduced resistance range in *Prochlorococcus* strains resistant to phage. Proceedings of the National Academy of Sciences E2191–E2200.

Avrani, S., Wurtzel, O., Sharon, I., Sorek, R., and Lindell, D. (2011). Genomic island variability facilitates *Prochlorococcus*-virus coexistence. Nature *474*, 604–608.

Bailey, S., Clokie, M.R.J., Millard, A., and Mann, N.H. (2004). Cyanophage infection and photoinhibition in marine cyanobacteria. Research in Microbiology *155*, 720–725.

Bailly-Bechet, M., Vergassola, M., and Rocha, E. (2007). Causes for the intriguing presence of tRNAs in phages. Genome Research *17*, 1486–1495.

Barksdale, L., and Arden, S.B. (1974). Persisting bacteriophage infections, lysogeny, and phage conversions. Annual Review of Microbiology *28*, 265–299.

Behrenfeld, M.J., O'Malley, R.T., Siegel, D. a, McClain, C.R., Sarmiento, J.L., Feldman, G.C., Milligan, A.J., Falkowski, P.G., Letelier, R.M., and Boss, E.S. (2006). Climate-driven trends in contemporary ocean productivity. Nature *444*, 752–755.

Bergh, O., Børsheim, K.Y., Bratbak, G., and Heldal, M. (1989). High abundance of viruses found in aquatic environments. Nature *340*, 467–468.

Bertilsson, S., Berglund, O., Karl, D.M., and Chisholm, S.W. (2003). Elemental composition of marine *Prochlorococcus* and *Synechococcus*: Implications for the ecological stoichiometry of the sea. Limnology and Oceanography *48*, 1721–1731.

Bhaya, D., Davison, M., and Barrangou, R. (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annual Review of Genetics *45*, 273–297.

Biller, S.J., Schubotz, F., Roggensack, S.E., Thompson, A.W., Summons, R.E., and Chisholm, S.W. (2014). Bacterial vesicles in marine ecosystems. Science *343*, 183–186.

Binder, B.J., and Chisholm, S.W. (1990). Relationship between DNA cycle and growth rate in *Synecchococcus* sp. strain PCC 6301. Journal of Bacteriology *172(5)*, 2313–2319.

Binder, B.J., and Chisholm, S.W. (1995). Cell cycle regulation in marine *Synechococcus* sp. strains. Applied and Environmental Microbiology *61*, 708–717.

Blot, N., Wu, X.J., Thomas, J.C., Zhang, J., Garczarek, L., Böhm, S., Tu, J.M., Zhou, M., Plöscher, M., Eichacker, L., et al. (2009). Phycourobilin trichromatic phycocyanin from oceanic cyanobacteria is formed post-translationally by a phycoerythrobilin lyase-isomerase. Journal of Biological Chemistry *284*, 9290–9298.

Boehme, J., Frischer, M., Jiang, S., Kellogg, C., Pichard, S., Rose, J., Steinway, C., and Paul, J.H. (1993). Viruses, bacterioplankton, and phyloplankton in the southeastern Gulf of Mexico: distribution and contribution to oceanic DNA pools. Marine Ecology Progress Series *97*, 1–10.

Boenigk, J. (2002). Bacterivory by heterotrophic flagellates. Antonie van Leeuwenhoek *81*, 465–480.

Boenigk, J., Matz, C., Jurgens, K., and Arndt, H. (2001). Confusing selective feeding with differential digestion in bacterivorous nanoflagellates. The Journal of Eukaryotic Microbiology *48*, 425–432.

Bohannan, B.J.M., and Lenski, R.E. (1999). Effect of prey heterogeneity on the response of a model food chain to resource enrichment. The American Naturalist *153*, 73–82.

Bohannan, B.J.M., and Lenski, R.E. (2000). Linking genetic change to community evolution: Insights from studies of bacteria and bacteriophage. Ecology Letters *3*, 362–377.

Bohannan, B.J.M., Travisano, M., and Lenski, R.E. (1999). Epistatic interactions can lower the cost of resistance to multiple consumers. Evolution *53*, 292–295.

Bohannan, B.J.M., Kerr, B., Jessup, C.M., Hughes, J.B., and Sandvik, G. (2002). Trade-offs and coexistence in microbial microcosms. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology *81*, 107–115.

Bondy-Denomy, J., and Davidson, A.R. (2014). To acquire or resist: The complex biological effects of CRISPR-Cas systems. Trends in Microbiology *22*, 218–225.

Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E., and Wincker, P. (2015). Tara Oceans studies plankton at planetary scale. Science *348*, 873–873.

Boulanger, P. (2009). Purification of bacteriophages and SDS-PAGE analysis of phage structural proteins from ghost particles. In Methods in Molecular Biology (Clifton, N.J.), pp. 227–238.

Brahamsha, B. (1996a). A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. Applied and Environmental Microbiology *62*, 1747–1751.

Brahamsha, B. (1996b). An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus*. Proceedings of the National Academy of Sciences *93*, 6504–6509.

Bratbak, G., Heldal, M., Thingstad, T.F., and Tuomi, P. (1996). Dynamics of virus abundance in coastal seawater. FEMS Microbiology Ecology *19*, 263–269.

Braun, M., and Silhavy, T.J. (2002). Imp/OstA is required for cell envelope biogenesis in *Escherichia coli*. Molecular Microbiology *45*, 1289–1302.

Breitbart, M. (2012). Marine viruses: truth or dare. Annual Review of Marine Science *4*, 425–448.

Breitbart, M., and Rohwer, F. (2005). Here a virus, there a virus, everywhere the same virus? Trends in Microbiology *13*, 278–284.

Breitbart, M., Thompson, L., Suttle, C., and Sullivan, M.B. (2007). Exploring the vast diversity of marine viruses. Oceanography *20*, 135–139.

Bremer, H., and Dennis, P. (1996). Modulation of chemical composition and other parameters of the cell by growth rate. In *Escherichia coli* and *Salmonella typhimurium*. Neidhardt, F. (ed.). Washington, DC: American Society for Microbiology Press 1553.

Breuert, S., Allers, T., Spohn, G., and Soppa, J. (2006). Regulated polyploidy in halophilic archaea. PLoS ONE *1*, e92.

Brockhurst, M.A., Fenton, A., Roulston, B., and Rainey, P.B. (2006). The impact of phages on interspecific competition in experimental populations of bacteria. BMC Ecology *6*, 19.

Brockhurst, M.A., Morgan, A.D., Fenton, A., and Buckling, A. (2007). Experimental coevolution with bacteria and phage. The *Pseudomonas fluorescens*-Φ2 model system. Infection, Genetics and Evolution *7*, 547–552.

Brum, J.R., and Sullivan, M.B. (2015). Rising to the challenge: accelerated pace of discovery transforms marine virology. Nature Reviews Microbiology.

Brum, J.R., Ignacio-Espinoza, J.C., Roux, S., Doulcier, G., Acinas, S.G., Alberti, A., Chaffron, S., Cruaud, C., de Vargas, C., Gasol, J.M., et al. (2015). Patterns and ecological drivers of ocean viral communities. Science *348*, 1261498–1261498.

Brussaard, C.P.D. (2004). Optimization of procedures for counting viruses by flow cytometry. Applied and Environmental Microbiology *70*, 1506–1513.

Brussaard, C.P.D., Marie, D., and Bratbak, G. (2000). Flow cytometric detection of viruses. Journal of Virological Methods *85*, 175–182.

Brüssow, H., Canchaya, C., Hardt, W., and Bru, H. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiology and Molecular Biology Reviews: MMBR *68*, 560–602.

Bryan, M.J., Burroughs, N.J., Spence, E.M., Clokie, M.R.J., Mann, N.H., and Bryan, S.J. (2008). Evidence for the intense exchange of MazG in marine cyanophages by horizontal gene transfer. PLoS ONE *3*, 1–12.

Buckling, A., and Rainey, P.B. (2002). Antagonistic coevolution between a bacterium and a bacteriophage. Proceedings of the Royal Society B: Biological Sciences *269*, 931–936.

Campbell, A. (2003). The future of bacteriophage biology. Nature Reviews. Genetics *4*, 471–477.

Campbell, L., and Iturriaga, R. (1988). Identification of *Synechococcus* spp. in the Sargasso Sea by immunofluorescence and fluorescence excitation spectroscopy performed on individual cells. Limnology and Oceanography *33*, 1196–1201.

Carroll, L. (1871). Through the looking-glass and what Alice found there (Macmillan).

Casadaban, M.J., and Cohen, S.N. (1980). Analysis of gene control signals by DNA fusion and cloning in *Escherichia coli*. Journal of Molecular Biology *138*, 179–207.

Cheetham, B.F., and Katz, M.E. (1995). A role for bacteriophages in the evolution and transfer of bacterial virulence determinants. Molecular Microbiology *18*, 201–208.

Christaki, U., Courties, C., Karayanni, H., Giannakourou, A., Maravelias, C., Kormas, K.A., and Lebaron, P. (2002). Dynamic characteristics of *Prochlorococcus* and *Synechococcus* consumption by bacterivorous nanoflagellates. Microbial Ecology *43*, 341–352.

Clasen, J.L., Hanson, C. a., Ibrahim, Y., Weihe, C., Marston, M.F., and Martiny, J.B.H. (2013). Diversity and temporal dynamics of Southern California coastal marine cyanophage isolates. Aquatic Microbial Ecology *69*, 17–31.

Clement, J.-M., Lepouce, E., Marchal, C., and Hofnung, M. (1983). Genetic study of a membrane protein: DNA sequence alterations due to 17 *lamB* point mutations affecting adsorption of phage lambda. The EMBO Journal *2*, 77–80.

Clokie, M.R.J., Shan, J., Bailey, S., Jia, Y., Krisch, H.M., West, S., and Mann, N.H. (2006). Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. Environmental Microbiology *8*, 827–835.

Clokie, M.R.J., Millard, A., Letarov, A. V, and Heaphy, S. (2011). Phages in nature. Bacteriophage *1*, 31–45.

Comeau, A.M., Tremblay, D., Moineau, S., Rattei, T., Kushkina, A.I., Tovkach, F.I., Krisch, H.M., and Ackermann, H.W. (2012). Phage morphology recapitulates phylogeny: The comparative genomics of a new group of myoviruses. PLoS ONE *7*, 1–11.

Cooper, S., and Helmstetter, C.E. (1968). Chromosome replication and the division cycle of *Escherichia coli*. Journal of Molecular Biology *31*, 519–540.

Corno, G., and Jürgens, K. (2006). Direct and indirect effects of protist predation on population size structure of a bacterial strain with high phenotypic plasticity. Applied and Environmental Microbiology *72*, 78–86.

Crawford, J.T., and Goldberg, E.B. (1980). The function of tail fibers in triggering

baseplate expansion of bacteriophage T4. Journal of Molecular Biology *139*, 679–690.

Cunningham, B.R., Brum, J.R., Schwenck, S.M., Sullivan, M.B., and John, S.G. (2015). An inexpensive, accurate and precise wet-mount method for enumerating aquatic viruses. Applied and Environmental Microbiology *81*, 2995–3000.

Danovaro, R., Corinaldesi, C., Dell'Anno, A., Fuhrman, J. a., Middelburg, J.J., Noble, R.T., and Suttle, C.A. (2011). Marine viruses and global climate change. FEMS Microbiology Reviews *35*, 993–1034.

David, E.H., Jeremy, a Y., and Chris, W. (1998). Molecular basis for structural diversity in the core regions of the lipopolysaccharides of *Escherichia coli* and *Salmonella enterica*. Molecular Microbiology *30*, 221–232.

Davis, J.M., Mayor, J., and Plamann, L. (1995). A missense mutation in rpoD results in an A-signalling defect in Myxococcus xanthus. Molecular Microbiology *18*, 943–952.

Delbrück, M. (1940). The growth of bacteriophage and lysis of the host. The Journal of General Physiology 643–660.

Díez-Villaseñor, C., Almendros, C., García-Martínez, J., and Mojica, F.J.M. (2010). Diversity of CRISPR loci in *Escherichia coli*. Microbiology *156*, 1351–1361.

Doležel, J., Bartoš, J., Voglmayr, H., and Greilhuber, J. (2003). Nuclear DNA content and genome size of trout and human. Cytometry. Part A: The Journal of the International Society for Analytical Cytology *51*, 127–128; author reply 129.

Duffy, M. a., and Forde, S.E. (2009). Ecological feedbacks and the evolution of resistance. Journal of Animal Ecology *78*, 1106–1112.

Dufresne, A., Ostrowski, M., Scanlan, D.J., Garczarek, L., Mazard, S., Palenik, B., Paulsen, I.T., de Marsac, N.T., Wincker, P., Dossat, C., et al. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. Genome Biology *9*, R90.

Dy, R.L., Richter, C., Salmond, G.P.C., and Fineran, P.C. (2014). Remarkable mechanisms in microbes to resist phage infections. Annual Review of Virology *1*, 307–331.

Edwards, R., and Rohwer, F. (2005). Viral metagenomics. Nature Reviews Microbiology *3*, 504–510.

Emiola, A., George, J., and Andrews, S.S. (2015). A complete pathway model for lipid A biosynthesis in Escherichia coli. PLoS ONE *10*, e0121216.

Enav, H., Béjà, O., and Mandel-Gutfreund, Y. (2012). Cyanophage tRNAs may have a role in cross-infectivity of oceanic *Prochlorococcus* and *Synechococcus*

hosts. The ISME Journal *6*, 619–628.

Field, C.B. (1998). Primary production of the biosphere: integrating terrestrial and oceanic components. Science *281*, 237–240.

Filée, J., Tétart, F., Suttle, C.A., and Krisch, H.M. (2005). Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. Proceedings of the National Academy of Sciences *102*, 12471–12476.

Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S., and Salmond, G.P.C. (2009). The phage abortive infection system, ToxIN, functions as a protein-RNA toxin-antitoxin pair. Proceedings of the National Academy of Sciences *106*, 894–899.

Flombaum, P., Gallegos, J.L., Gordillo, R. a, Rincón, J., Zabala, L.L., Jiao, N., Karl, D.M., Li, W.K.W., Lomas, M.W., Veneziano, D., et al. (2013). Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. Proceedings of the National Academy of Sciences *110*, 9824–9829.

Forterre, P., Soler, N., Krupovic, M., Marguet, E., and Ackermann, H.W. (2013). Fake virus particles generated by fluorescence microscopy. Trends in Microbiology *21*, 1–5.

Fu, F.X., Warner, M.E., Zhang, Y., Feng, Y., and Hutchins, D. a. (2007). Effects of increased temperature and CO2 on photosynthesis, growth, and elemental ratios in marine *Synechococcus* and *Prochlorococcus* (Cyanobacteria). Journal of Phycology *43*, 485–496.

Fuhrman, J. a. (1999). Marine viruses and their biogeochemical and ecological effects. Nature *399*, 541–548.

Fuhrman, J. a., and Noble, R.T. (1995). Viruses and protists cause similar bacterial mortality in coastal seawater. Limnology and Oceanography *40*, 1236–1242.

Fuhrman, J. a., and Schwalbach, M. (2003). Viral influence on aquatic bacterial communities. Biological Bulletin *204*, 192–195.

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

Fuller, N.J., Marie, D., Vaulot, D., Post, A.F., and Scanlan, D.J. (2003). Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. Applied and Environmental Microbiology *69*, 2430–2443.

Fuller, N.J., West, N.J., Marie, D., Yallop, M., Rivlin, T., Post, A.F., Interuniversity, T., Sciences, M., Beach, C., and Scanlan, D.J. (2005). Dynamics of community structure and phosphate status of picocyanobacterial populations in the Gulf of Aqaba, Red Sea. Limnology and Oceanography *50*, 363–375.

Fuller, N.J., Tarran, G. a., Yallop, M., Orcutt, K.M., and Scanlan, D.J. (2006). Molecular analysis of picocyanobacterial community structure along an Arabian Sea transect reveals distinct spatial separation of lineages. Limnology and Oceanography *51*, 2515–2526.

García-Alcalde, F., Okonechnikov, K., Carbonell, J., Cruz, L.M., Götz, S., Tarazona, S., Dopazo, J., Meyer, T.F., and Conesa, A. (2012). Qualimap: Evaluating next-generation sequencing alignment data. Bioinformatics *28*, 2678–2679.

Garczarek, L., Dufresne, A., Blot, N., Cockshutt, A.M., Peyrat, A., Campbell, D. a, Joubin, L., and Six, C. (2008). Function and evolution of the *psbA* gene family in marine *Synechococcus*: *Synechococcus* sp. WH7803 as a case study. The ISME Journal *2*, 937–953.

Glazer, a N., and Clark, J.H. (1986). Phycobilisomes: macromolecular structure and energy flow dynamics. Biophysical Journal *49*, 115–116.

Gonzalez, J.M. (1996). Efficient size-selective bacterivory by phagotrophic nanoflagellates in aquatic ecosystems. Marine Biology *126*, 785–789.

Gonzalez, J.M., and Suttle, C.A. (1993). Grazing by marine nanoflagellates on viruses and virus-sized particles: ingestion and digestion. Marine Ecology Progress Series *94*, 1–10.

Griese, M., Lange, C., and Soppa, J. (2011). Ploidy in cyanobacteria. FEMS Microbiology Letters *323*, 124–131.

Grossman, A.R., Schaefer, M.R., Chiang, G.G., and Collier, J.L. (1993a). The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiological Reviews *57*, 725–749.

Grossman, A.R., Schaefer, M.R., Chiang, G.G., and Collier, J.L. (1993b). Environmental effects on the light-harvesting complex of cyanobacteria. Journal of Bacteriology *175*, 575–582.

Guillou, L., Jacquet, S., Chrétiennot-Dinet, M.J., and Vaulot, D. (2001). Grazing impact of two small heterotrophic flagellates on Prochlorococcus and Synechococcus. Aquatic Microbial Ecology *26*, 201–207.

Hambly, E., Tétart, F., Desplats, C., Wilson, W.H., Krisch, H.M., and Mann, N.H. (2001). A conserved genetic module that encodes the major virion components in both the coliphage T4 and the marine cyanophage S-PM2. Proceedings of the National Academy of Sciences *98*, 11411–11416.

Hancock, R.E.W., and Reeves, P. (1976). Lipopolysaccharide-deficient, bacteriophage-resistant mutants in *Escherichia coli* K-12. Journal of Bacteriology *127*, 98–108.

Hansen, M.T. (1978). Multiplicity of genome equivalents in the radiation-resistant bacterium *Micrococcus radiodurans*. Journal of Bacteriology *134*, 71–75.

Hara, S., Koike, I., Terauchi, K., Kamiya, H., and Tanoue, E. (1996). Abundance of viruses in deep oceanic waters. Marine Ecology Progress Series *145*, 269–277.

Heldal, M., Scanlan, D.J., Norland, S., Thingstad, T.F., and Mann, N.H. (2003). Elemental composition of single cells of various strains of marine *Prochlorococcus* and *Synechococcus* using X-ray microanalysis. Limnology and Oceanography *48*, 1732–1743.

Heller, K., and Braun, V. (1979). Accelerated adsorption of bacteriophage T5 to *Escherichia coli* F, resulting from reversible tail fiber-lipopolysaccharide binding. Journal of Bacteriology *139*, 32–38.

Henry, J.T., and Crosson, S. (2011). Ligand-binding PAS domains in a genomic, cellular, and structural context. Annual Review of Microbiology *65*, 261–286.

Hildenbrand, C., Stock, T., Lange, C., Rother, M., and Soppa, J. (2011). Genome copy numbers and gene conversion in methanogenic archaea. Journal of Bacteriology *193*, 734–743.

Hjort, K., and Bernander, R. (1999). Changes in cell size and DNA content in *Sulfolobus* cultures during dilution and temperature shift experiments. Journal of Bacteriology *181*, 5669–5675.

Hofer, J.S., and Sommaruga, R. (2001). Seasonal dynamics of viruses in an alpine lake: Importance of filamentous forms. Aquatic Microbial Ecology *26*, 1–11.

Holst, O. (2007). The structures of core regions from enterobacterial lipopolysaccharides - an update. FEMS Microbiology Letters *271*, 3–11.

Hu, J.C., and Gross, C. a (1988). Mutations in *rpoD* that increase expression of genes in the mal regulon of *Escherichia coli* K-12. Journal of Molecular Biology *203*, 15–27.

Hu, B., Yang, G., Zhao, W., Zhang, Y., and Zhao, J. (2007). MreB is important for cell shape but not for chromosome segregation of the filamentous cyanobacterium *Anabaena* sp. PCC 7120. Molecular Microbiology *63*, 1640–1652.

Hu, B., Margolin, W., Molineux, I.J., and Liu, J. (2015). Structural remodeling of bacteriophage T4 and host membranes during infection initiation. Proceedings of the National Academy of Sciences 201501064.

Huang, S., Wang, K., Jiao, N., and Chen, F. (2012a). Genome sequences of siphoviruses infecting marine *Synechococcus* unveil a diverse cyanophage group and extensive phage-host genetic exchanges. Environmental Microbiology *14*, 540–558.

Huang, W., Li, L., Myers, J.R., and Marth, G.T. (2012b). ART: A next-generation sequencing read simulator. Bioinformatics *28*, 593–594.

Hurwitz, B.L., and Sullivan, M.B. (2013). The Pacific Ocean Virome (POV): A marine viral metagenomic dataset and associated protein clusters for quantitative viral ecology. PLoS ONE *8*.

Hyman, P., and Abedon, S.T. (2009). Bacteriophage (overview). In Encyclopedia of Microbiology, (Elsevier), pp. 322–338.

Ignacio-Espinoza, J.C., Solonenko, S. a., and Sullivan, M.B. (2013). The global virome: Not as big as we thought? Current Opinion in Virology *3*, 566–571.

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. Journal of Bacteriology *169*, 5429–5433.

Jain, I.H., Vijayan, V., and O'Shea, E.K. (2012). Spatial ordering of chromosomes enhances the fidelity of chromosome partitioning in cyanobacteria. Proceedings of the National Academy of Sciences *109*, 13638–13643.

Jardillier, L., Zubkov, M. V., Pearman, J., and Scanlan, D.J. (2010). Significant CO2 fixation by small prymnesiophytes in the subtropical and tropical northeast Atlantic Ocean. The ISME Journal *4*, 1180–1192.

Jia, Y. (2009). An investigation into the adsorption of cyanophages to their cyanobacterial hosts. PhD Thesis, University of Warwick.

Jia, Y., Shan, J., Millard, A., Clokie, M.R.J., and Mann, N.H. (2010). Light-dependent adsorption of photosynthetic cyanophages to *Synechococcus* sp. WH7803. FEMS Microbiology Letters *310*, 120–126.

Jiang, S.C., and Paul, J.H. (1995). Viral contribution to dissolved DNA in the marine environment as determined by differential centrifugation and kingdom probing. Applied and Environmental Microbiology *61*, 317–325.

Jiang, W., Li, Z., Zhang, Z., Baker, M.L., Prevelige, P.E., and Chiu, W. (2003). Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. Nature Structural Biology *10*, 131–135.

John, E.H., and Davidson, K. (2001). Prey selectivity and the influence of prey carbon:nitrogen ratio on microflagellate grazing. Journal of Experimental Marine Biology and Ecology *260*, 93–111.

Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M.S., and Chisholm, S.W. (2006). Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. Science *311*, 1737–1740.

Jones, H., Ostrowski, M., and Scanlan, D.J. (2006). A suppression subtractive hybridization approach reveals niche-specific genes that may be involved in predator avoidance in marine *Synechococcus* isolates. Applied and Environmental Microbiology *72*, 2730–2737.

Jover, L.F., Effler, T.C., Buchan, A., Wilhelm, S.W., and Weitz, J.S. (2014). The elemental composition of virus particles: implications for marine biogeochemical cycles. Nature Reviews Microbiology *12*, 519–528.

Kitten, T., and Barbour, a. G. (1992). The relapsing fever agent *Borrelia hermsii* has multiple copies of its chromosome and linear plasmids. Genetics *132*, 311–324.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., Mclellan, M.D., Lin, L., Miller, C. a., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Research 568–576.

Komaki, K., and Ishikawa, H. (2000). Genomic copy number of intracellular bacterial symbionts of aphids varies in response to developmental stage and morph of their host. Insect Biochemistry and Molecular Biology *30*, 253–258.

Kovalyova, I. V., and Kropinski, A.M. (2003). The complete genomic sequence of lytic bacteriophage gh-1 infecting *Pseudomonas putida* - Evidence for close relationship to the T7 group. Virology *311*, 305–315.

Kretz, C.B., Bell, D.W., Lomas, D. a., Lomas, M.W., and Martiny, A.C. (2015). Influence of growth rate on the physiological response of marine *Synechococcus* to phosphate limitation. Frontiers in Microbiology *6*, 6–11.

Kurmayer, R., and Kutzenberger, T. (2003). Application of real-time PCR for quantification of microcystin genotypes in a population of the toxic cyanobacterium *Microcystis* sp. Applied and Environmental Microbiology *69*, 6723–6730.

Labarre, J., Chauvat, F., and Thuriaux, P. (1989). Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium *Synechocystis* strain PCC 6803. Journal of Bacteriology *171*, 3449–3457.

Labrie, S.J., Samson, J.E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. Nature Reviews Microbiology *8*, 317–327.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods *9*, 357–359.

Lennon, J.T., Khatana, S.A.M., Marston, M.F., and Martiny, J.B.H. (2007). Is there a cost of virus resistance in marine cyanobacteria? The ISME Journal *1*, 300–312.

Lenski, R.E. (1988a). Experimental studies of pleiotropy and epistasis in *Escherichia coli*. II. compensation for maladaptive effects associated with resistance to virus T4. Evolution *42*, 433–440.

Lenski, R.E. (1988b). Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. variation in competitive fitness among mutants resistant to virus T4.

Evolution *42*, 425–432.

Lenski, R.E., and Levin, B.R. (1985). Constraints on the coevolution of bacteria and virulent phage: A model, some experiments, and predictions for natural communities. The American Naturalist *125*, 585.

Lewis, T., Loman, N.J., Bingle, L., Jumaa, P., Weinstock, G.M., Mortiboy, D., and Pallen, M.J. (2010). High-throughput whole-genome sequencing to dissect the epidemiology of *Acinetobacter baumannii* isolates from a hospital outbreak. Journal of Hospital Infection *75*, 37–41.

van der Ley, P., de Graaff, P., and Tommassen, J. (1986). Shielding of Escherichia coli outer membrane proteins as receptors for bacteriophages and colicins by O-antigenic chains of lipopolysaccharide. Journal of Bacteriology *168*, 449–451.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2 *00*, 3.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G.T., Abecasis, G., and Durbin, R. (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S., Ignacio-Espinosa, J.C., Roux, S., Vincent, F., et al. (2015). Determinants of community structure in the global plankton interactome. Science *348*, 1262073–1262073.

Lindell, D. (2014). The genus *Prochlorococcus*, phylum cyanobacteria. In The Prokaryotes - Other Major Lineages of Bacteria and The Archaea, E. Rosenberg, E.F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson, eds. (Springer Berlin Heidelberg), pp. 829–845.

Lindell, D., Jaffe, J.D., Johnson, Z.I., Church, G.M., and Chisholm, S.W. (2005). Photosynthesis genes in marine viruses yield proteins during host infection. Nature *438*, 86–89.

Liu, H., Bidigare, R.R., Laws, E., Landry, M.R., and Campbell, L. (1999). Cell cycle and physiological characteristics of *Synechococcus* (WH7803) in chemostat culture. Marine Ecology Progress Series *189*, 17–25.

Long, A., McDaniel, L.D., Mobberley, J., and Paul, J.H. (2008). Comparison of lysogeny (prophage induction) in heterotrophic bacterial and *Synechococcus* populations in the Gulf of Mexico and Mississippi River plume. The ISME Journal *2*, 132–144.

López-Sánchez, M.J., Neef, A., Patiño-Navarrete, R., Navarro, L., Jiménez, R., Latorre, A., and Moya, A. (2008). Blattabacteria, the endosymbionts of cockroaches, have small genome sizes and high genome copy numbers. Environmental Microbiology *10*, 3417–3422.

Lu, M., and Henning, U. (1989). The immunity (*imm*) gene of *Escherichia coli* bacteriophage T4. Journal of Virology *63*, 3472–3478.

Lu, M., and Henning, U. (1994). Superinfection exclusion by T-even type coliphages. Trends in Microbiology *2*, 137–139.

Lu, J., Chen, F., and Hodson, R.E. (2001). Distribution, isolation, host specificity and diversity of cyanophages infecting marine Synechococcus spp. in river estuaries. Applied and Environmental Microbiology *67*, 3285–3290.

Lu, M., Stierhof, Y.D., and Henning, U. (1993). Location and unusual membrane topology of the immunity protein of the *Escherichia coli* phage T4. Journal of Virology *67*, 4905–4913.

Maldonado, R., Jimenez, J., and Casadesus, J. (1994). Changes of ploidy during the *Azotobacter vinelandii* growth cycle. Journal of Bacteriology *176*, 3911–3919.

Mann, N.H. (2003). Phages of the marine cyanobacterial picophytoplankton. FEMS Microbiology Reviews *27*, 17–34.

Mann, N.H., Cook, A., Millard, A., Bailey, S., and Clokie, M.R.J. (2003). Marine ecosystems: bacterial photosynthesis genes in a virus. Nature *424*, 741.

Mann, N.H., Clokie, M.R.J., Millard, A., Cook, A., Wilson, W.H., Wheatley, P.J., Letarov, A., and Krisch, H.M. (2005). The genome of S-PM2, a 'photosynthetic' T4-type bacteriophage that infects marine *Synechococcus* strains. Journal of Bacteriology *187*, 3188–3200.

Marston, M.F., and Sallee, J.L. (2003). Genetic diversity and temporal variation in the cyanophage community infecting marine *Synechococcus* species in Rhode Island's Coastal Waters. Applied and Environmental Microbiology *69*, 4639–4647.

Marston, M.F., Pierciey, F.J., Shepard, A., Gearin, G., Qi, J., Yandava, C., Schuster, S.C., Henn, M.R., and Martiny, J.B.H. (2012). Rapid diversification of coevolving marine *Synechococcus* and a virus. Proceedings of the National Academy of Sciences *109*, 1–6.

Martin, A.P., Zubkov, M. V, Burkill, P.H., and Holland, R.J. (2005). Extreme spatial variability in marine picoplankton and its consequences for interpreting Eulerian time-series. Biology Letters *1*, 366–369.

Martiny, J.B.H., Riemann, L., Marston, M.F., and Middelboe, M. (2014). Antagonistic coevolution of marine planktonic viruses and their hosts. Annual Review of Marine Science *6*, 393–414.

Masterson, J. (1994). Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. Science *264*, 421–424.

Matz, C., and Jürgens, K. (2001). Effects of hydrophobic and electrostatic cell

surface properties of bacteria on feeding rates of heterotrophic nanoflagellates. Applied and Environmental Microbiology *67*, 814–820.

Matz, C., and Jürgens, K. (2005). High motility reduces grazing mortality of planktonic bacteria. Applied and Environmental Microbiology *71*.

Matz, C., and Kjelleberg, S. (2005). Off the hook - How bacteria survive protozoan grazing. Trends in Microbiology *13*, 302–307.

Matz, C., Boenigk, J., Arndt, H., and Jurgens, K. (2002). Role of bacterial phenotypic traits in selective feeding of the heterotrophic nanoflagellate *Spumella* sp. Aquatic Microbial Ecology *27*, 137–148.

Mazard, S., Ostrowski, M., Partensky, F., and Scanlan, D.J. (2012a). Multi-locus sequence analysis, taxonomic resolution and biogeography of marine *Synechococcus*. Environmental Microbiology *14*, 372–386.

Mazard, S., Wilson, W.H., and Scanlan, D.J. (2012b). Dissecting the physiological response to phosphorus stress in marine *Synechococcus* isolates (Cyanophyceae). Journal of Phycology *48*, 94–105.

McDaniel, L.D., and Paul, J.H. (2005). Effect of nutrient addition and environmental factors on prophage induction in natural populations of marine *Synechococcus* species. Applied and Environmental Microbiology *71*, 842–850.

McDaniel, L.D., Houchin, L. a, Williamson, S.J., and Paul, J.H. (2002). Lysogeny in marine *Synechococcus*. Nature *415*, 496.

McDaniel, L.D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K.B., and Paul, J.H. (2010). High frequency of horizontal gene transfer in the oceans. Science *330*, 50.

McGrath, C.L., and Katz, L. a. (2004). Genome diversity in microbial eukaryotes. Trends in Ecology and Evolution *19*, 32–38.

Mendell, J.E., Clements, K.D., Choat, J.H., and Angert, E.R. (2008). Extreme polyploidy in a large bacterium. Proceedings of the National Academy of Sciences *105*, 6730–6734.

Meselson, M., and Yuan, R. (1968). DNA restriction enzyme from *E. coli*. Nature *217*, 1110–1114.

Michelsen, O., Hansen, F.G., Albrechtsen, B., and Jensen, P.R. (2010). The MG1363 and IL1403 laboratory strains of *Lactococcus lactis* and several dairy strains are diploid. Journal of Bacteriology *192*, 1058–1065.

Millard, A. (2009). Isolation of cyanophages from aquatic environments. In Bacteriophage: Methods and Protocols, A.M. Kropinski, and M.R.J. Clokie, eds. pp. 33–42.

Millard, A., and Mann, N.H. (2006). A temporal and spatial investigation of cyanophage abundance in the Gulf of Aqaba Red Sea. Journal of the Marine Biological Association of the United Kingdom *86*, 507–515.

Millard, A., Clokie, M.R.J., Shub, D. a, and Mann, N.H. (2004). Genetic organization of the *psbAD* region in phages infecting marine *Synechococcus* strains. Proceedings of the National Academy of Sciences *101*, 11007–11012.

Millard, A., Zwirglmaier, K., Downey, M.J., Mann, N.H., and Scanlan, D.J. (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*, 2370–2387.

Millard, A., Gierga, G., Clokie, M.R.J., Evans, D.J., Hess, W.R., and Scanlan, D.J. (2010). An antisense RNA in a lytic cyanophage links *psbA* to a gene encoding a homing endonuclease. The ISME Journal *4*, 1121–1135.

Miller, E.S., Kutter, E., Mosig, G., Kunisawa, T., Rüger, W., Arisaka, F., and Ru, W. (2003). Bacteriophage T4 genome. Microbiology and Molecular Biology Reviews: MMBR *67*, 86–156.

Mizoguchi, K., Morita, M., Fischer, C.R., Yoichi, M., Tanji, Y., and Unno, H. (2003). Coevolution of bacteriophage PP01 and *Escherichia coli* O157:H7 in continuous culture. Applied and Environmental Microbiology *69*, 170–176.

Mohapatra, B.R., and Fukami, K. (2004). Comparison of the numerical grazing response of two marine heterotrophic nanoflagellates fed with different bacteria. Journal of Sea Research *52*, 99–107.

Mojica, F.J.M., Díez-Villaseñor, C., García-Martínez, J., and Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. Microbiology *155*, 733–740.

Mojica, K.D. a, Huisman, J., Wilhelm, S.W., and Brussaard, C.P.D. (2015). Latitudinal variation in virus-induced mortality of phytoplankton across the North Atlantic Ocean. The ISME Journal 1–14.

Mojtahid, M., Zubkov, M. V., Hartmann, M., and Gooday, a. J. (2011). Grazing of intertidal benthic foraminifera on bacteria: Assessment using pulse-chase radiotracing. Journal of Experimental Marine Biology and Ecology *399*, 25–34.

Monger, B.C., Landry, M.R., and Brown, S.L. (1999). Feeding selection of heterotrophic marine nanoflagellates based on the surface hydrophobicity of their picoplankton prey. Limnology and Oceanography *44*, 1917–1927.

Montag, D., Riede, I., Eschbach, M.L., Degen, M., and Henning, U. (1987). Receptor-recognizing proteins of T-even type bacteriophages. Constant and hypervariable regions and an unusual case of evolution. Journal of Molecular

Biology *196*, 165–174.

Moore, L.R., Rocap, G., and Chisholm, S.W. (1998). Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. Nature *393*, 464–467.

Moore, L.R., Ostrowski, M., Scanlan, D.J., Feren, K., and Sweetsir, T. (2005). Ecotypic variation in phosphorus-acquisition mechanisms within marine picocyanobacteria. Aquatic Microbial Ecology *39*, 257–269.

Morel, A., Ahn, Y.-H., Partensky, F., Vaulot, D., and Claustre, H. (1993). *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. Journal of Plankton Research *51*, 617–649.

Mori, T., Binder, B.J., and Johnson, C.H. (1996). Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours. Proceedings of the National Academy of Sciences *93*, 10183–10188.

Morita, M., Tanji, Y., Mizoguchi, K., and Akitsu, T. (2002). Characterization of a virulent bacteriophage specific for *Escherichia coli* O157:H7 and analysis of its cellular receptor and two tail fiber genes. FEMS Microbiology Letters *211*.

Mota, R., Guimarães, R., Büttel, Z., Rossi, F., Colica, G., Silva, C.J., Santos, C., Gales, L., Zille, A., De Philippis, R., et al. (2013). Production and characterization of extracellular carbohydrate polymer from *Cyanothece* sp. CCY 0110. Carbohydrate Polymers *92*, 1408–1415.

Mühling, M., Fuller, N.J., Somerfield, P.J., Post, A.F., Wilson, W.H., Scanlan, D.J., Joint, I., and Mann, N.H. (2006). High resolution genetic diversity studies of marine *Synechococcus* isolates using *rpoC1* -based restriction fragment length polymorphism. Aquatic Microbial Ecology *45*, 263–275.

Murray, A.G., and Eldridge, P.M. (1994). Marine viral ecology: incorporation of bacteriophage into the microbial planktonic food web paradigm. Journal of Plankton Research *16*, 627–641.

Nakao, R., Ramstedt, M., Wai, S.N., and Uhlin, B.E. (2012). Enhanced biofilm formation by *Escherichia coli* LPS mutants defective in Hep biosynthesis. PLoS ONE *7*.

Nara, F., Matsuyama, S., Mizuno, T., and Mizushima, S. (1986). Molecular analysis of mutant *ompR* genes exhibiting different phenotypes as to osmoregulation of the *ompF* and *ompC* genes of *Escherichia coli*. Molecular and General Genetics *202*, 194–199.

Noble, R.T., and Fuhrman, J. a. (1998). Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. Aquatic Microbial Ecology *14*, 113–118.

Ohtani, N., Tomita, M., and Itaya, M. (2010). An extreme thermophile, *Thermus*

thermophilus, is a polyploid bacterium. Journal of Bacteriology *192*, 5499–5505.

Okada, K., Iida, T., Kita-Tsukamoto, K., and Honda, T. (2005). Vibrios commonly possess two chromosomes. Journal of Bacteriology *187*, 752–757.

Olson, R.J., Chisholm, S.W., Zettler, E.R., and Armbrust, E. V (1988). Analysis of *Synechococcus* pigment types in the sea using single and dual beam flow-cytometry. Deep Sea Research *35*, 425–440.

Olson, R.J., Chisholm, S.W., and Zettler, E.R. (1990). Pigments, size and distribution of *Synechococcus* in the North Atlantic and Pacific Oceans. Limnology and Oceanography *35*, 45–58.

Ong, L.J., and Glazer, a. N. (1991). Phycoerythrins of marine unicellular cyanobacteria: I. Bilin types and locations and energy transfer pathways in *Synechococcus* spp. phycoerythrins. Journal of Biological Chemistry *266*, 9515–9527.

Orskov, I., Orskov, F., Jann, B., and Jann, K. (1977). Serology, chemistry, and genetics of O and K antigens of *Escherichia coli*. Bacteriological Reviews *41*, 667–710.

Ortmann, a. C., Lawrence, J.E., and Suttle, C.A. (2002). Lysogeny and lyric viral production during a bloom of the cyanobacterium *Synechococcus* spp. Microbial Ecology *43*, 225–231.

Otto, S.P., and Whitton, J. (2000). Polyploid incidence and evolution. Annual Review of Genetics *34*, 401–437.

Palenik, B., and Haselkorn, R. (1992). Multiple evolutionary origins of prochlorophytes, the chlorophyll b-containing prokaryotes. Nature *355*, 265–267.

Palenik, B., Brahamsha, B., Larimer, F.W., Land, M., Hauser, L., Chain, P., Lamerdin, J., Regala, W., Allen, E.E., McCarren, J., et al. (2003). The genome of a motile marine *Synechococcus*. Nature *424*, 1037–1042.

Palenik, B., Ren, Q., Dupont, C.L., Myers, G.S., Heidelberg, J.F., Badger, J.H., Madupu, R., Nelson, W.C., Brinkac, L.M., Dodson, R.J., et al. (2006). Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. Proceedings of the National Academy of Sciences *103*, 13555–13559.

Parfrey, L.W., Lahr, D.J.G., and Katz, L.A. (2008). The dynamic nature of eukaryotic genomes. Molecular Biology and Evolution *25*, 787–794.

Parfrey, L.W., Lahr, D.J.G., and Katz, L.A. (2010). Dynamic genomes of eukaryotes and the maintenance of genomic integrity. Microbe *5*, 787–794.

Park, K.M., and So, J.S. (2000). Altered cell surface hydrophobicity of

lipopolysaccharide-deficient mutant of *Bradyrhizobium japonicum*. Journal of Microbiological Methods *41*, 219–226.

Partensky, F., Blanchot, J., and Vaulot, D. (1999). Differential distribuition and ecology of *Prochloococcus* and *Synechococcus* in oceanic waters: a review. Bulletin de l'Institut Oceanographique, Monaco 457–475.

Paterson, S., Vogwill, T., Buckling, A., Benmayor, R., Spiers, A.J., Thomson, N.R., Quail, M., Smith, F., Walker, D., Libberton, B., et al. (2010). Antagonistic coevolution accelerates molecular evolution. Nature *464*, 275–278.

Paul, J.H., Sullivan, M.B., Segall, A.M., and Rohwer, F. (2002). Marine phage genomics. Comparative Biochemistry and Physiology - B Biochemistry and Molecular Biology *133*, 463–476.

Pecoraro, V., Zerulla, K., Lange, C., and Soppa, J. (2011). Quantification of ploidy in proteobacteria revealed the existence of monoploid, (mero-)oligoploid and polyploid species. PLoS ONE *6*, e16392.

Pedruzzi, I., Rosenbusch, J.P., and Locher, K.P. (1998). Inactivation in vitro of the *Escherichia coli* outer membrane protein FhuA by a phage T5-encoded lipoprotein. FEMS Microbiology Letters *168*, 119–125.

Pereira, S.B., Zille, A., Micheletti, E., Moradas-Ferreira, P., De Philippis, R., and Tamagnini, P. (2009). Complexity of cyanobacterial exopolysaccharides: Composition, structures, inducing factors and putative genes involved in their biosynthesis and assembly. FEMS Microbiology Reviews *33*, 917–941.

Pernthaler, J. (2005). Predation on prokaryotes in the water column and its ecological implications. Nature Reviews Microbiology *3*, 537–546.

Perry, E.B., Barrick, J.E., and Bohannan, B.J.M. (2015). The molecular and genetic basis of repeatable coevolution between *Escherichia coli* and bacteriophage T3 in a laboratory microcosm. PLoS ONE *10*, e0130639.

Philosof, A., Battchikova, N., Aro, E.-M., and Béjà, O. (2011). Marine cyanophages: tinkering with the electron transport chain. The ISME Journal *5*, 1568–1570.

Pinchuk, G.E., Ammons, C., Culley, D.E., Li, S.M.W., McLean, J.S., Romine, M.F., Nealson, K.H., Fredrickson, J.K., and Beliaev, A.S. (2008). Utilization of DNA as a sole source of phosphorus, carbon, and energy by *Shewanella* spp.: ecological and physiological implications for dissimilatory metal reduction. Applied and Environmental Microbiology *74*, 1198–1208.

Pittera, J., Humily, F., Thorel, M., Grulois, D., Garczarek, L., and Six, C. (2014). Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. The ISME Journal 1–16.

Postgate, J.R., Kent, H.M., Robson, R.L., and Chesshyre, J. a (1984). The genomes of *Desulfovibrio gigas* and *D. vulgaris*. Journal of General Microbiology *130*,

1597–1601.

Proctor, L.M., and Fuhrman, J. a. (1990). Viral mortality of marine bacteria and cyanobacteria. Nature *343*, 60–62.

Putker, F., Bos, M.P., and Tommassen, J. (2015). Transport of lipopolysaccharide to the Gram-negative bacterial cell surface. FEMS Microbiology Reviews 1–18.

Puxty, R.J., Millard, A., Evans, D.J., and Scanlan, D.J. (2014). Shedding new light on viral photosynthesis. Photosynthesis Research.

Puxty, R.J., Perez-Sepulveda, B., Rihtman, B., Evans, D.J., Millard, A., and Scanlan, D.J. (2015). Spontaneous deletion of an 'ORFanage' region facilitates host adaptation in a 'photosynthetic' cyanophage. PLoS ONE *10*, e0132642.

Quail, M., Smith, M.E., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., and Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. BMC Genomics *13*, 1.

Raetz, C.R.H., and Whitfield, C. (2002). Lipopolysaccharide endotoxins. Annual Review of Biochemistry *71*, 635–700.

Raetz, C.R.H., Reynolds, C.M., Trent, M.S., and Bishop, R.E. (2007). Lipid A modification systems in gram-negative bacteria. Annual Review of Biochemistry *76*, 295–329.

Raetz, C.R.H., Guan, Z., Ingram, B.O., Six, D. a, Song, F., Wang, X., and Zhao, J. (2009). Discovery of new biosynthetic pathways: the lipid A story. Journal of Lipid Research S103–S108.

Rakhuba, D. V., Kolomiets, E.I., Szwajcer Dey, E., and Novik, G.I. (2010). Bacteriophage receptors, mechanisms of phage adsorption and penetration into host cell. Polish Journal of Microbiology *59*, 145–155.

Raymond, C.K., Sims, E.H., Kas, A., Spencer, D.H., Kutyavin, T. V, Ivey, R.G., Kaul, R., Clendenning, J.B., Maynard, V., Zhou, Y., et al. (2002). Genetic variation at the O-antigen biosynthetic locus in Pseudomonas aeruginosa. Journal of Bacteriology *184*, 3614–3622.

Raytcheva, D. a, Haase-Pettingell, C., Piret, J.M., and King, J. a (2011). Intracellular assembly of cyanophage Syn5 proceeds through a scaffold-containing procapsid. Journal of Virology *85*, 2406–2415.

Regué, M., Climent, N., Abitiu, N., Coderch, N., Merino, S., Izquierdo, L., Altarriba, M., and Tomás, J.M. (2001). Genetic characterization of the *Klebsiella pneumoniae waa* gene cluster, involved in core lipopolysaccharide biosynthesis. Journal of Bacteriology *183*, 3564–3573.

Riede, I., and Eschbach, M.-L. (1986). Evidence that TraT interacts with OmpA of Escherichia coli. FEBS Letters *205*, 241–245.

Riede, I., Degen, M., and Henning, U. (1985). The receptor specificity of bacteriophages can be determined by a tail fiber modifying protein. The EMBO Journal *4*, 2343–2346.

Rippka, R., Coursin, T., Hess, W., Lichtle, C., Scanlan, D.J., Palinska, K. a., Iteman, I., Partensky, F., Houmard, J., and Herdman, M. (2000). *Prochlorococcus* marinus Chisholm et al. 1992 subsp. pastoris subsp. nov. strain PCC 9511, the first axenic chlorophyll a2/b2-containing cyanobacterium (Oxyphotobacteria). International Journal of Systematic and Evolutionary Microbiology *50*, 1833–1847.

Rocap, G., Distel, D.L., Waterbury, J.B., and Chisholm, S.W. (2002). Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S-23S ribosomal DNA internal transcribed spacer sequences. Applied and Environmental Microbiology *68*, 1180–1191.

Rodriguez-Valera, F., Martin-Cuadrado, A.B., Rodriguez-Brito, B., Pasić, L., Thingstad, T.F., Rohwer, F., and Mira, A. (2009). Explaining microbial population genomics through phage predation. Nature Reviews Microbiology *7*, 828–836.

Roncero, C., and Casadaban, M.J. (1992). Genetic analysis of the genes involved in synthesis of the lipopolysaccharide core in *Escherichia coli* K-12: Three operons in the rfa locus. Journal of Bacteriology *174*, 3250–3260.

Ruiz, N., Gronenberg, L.S., Kahne, D., and Silhavy, T.J. (2008). Identification of two inner-membrane proteins required for the transport of lipopolysaccharide to the outer membrane of *Escherichia coli*. Proceedings of the National Academy of Sciences *105*, 5537–5542.

Saldías, M.S., Patel, K., Marolda, C.L., Bittner, M., Contreras, I., and Valvano, M. a. (2008). Distinct functional domains of the *Salmonella enterica* WbaP transferase that is involved in the initiation reaction for synthesis of the O antigen subunit. Microbiology *154*, 440–453.

Samuel, G., and Reeves, P. (2003). Biosynthesis of O-antigens: Genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly. Carbohydrate Research *338*, 2503–2519.

Sander, J.D., and Joung, J.K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. Nature Biotechnology *32*, 347–355.

Sanders, R., Henson, S. a., Koski, M., De La Rocha, C.L., Painter, S.C., Poulton, A.J., Riley, J., Salihoglu, B., Visser, A., Yool, A., et al. (2014). The biological carbon pump in the North Atlantic. Progress in Oceanography *129*, 200–218.

Scanlan, D.J. (2003). Physiological diversity and niche adaptation in marine *Synechococcus*. Advances in Microbial Physiology *47*, 1–64.

Scanlan, D.J. (2012). Marine picocyanobacteria. In Ecology of Cyanobacteria II: Their Diversity in Space and Time, B. a. Whitton, ed. (Springer Netherlands), pp. 297–298.

Scanlan, D.J., and West, N.J. (2002). Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. FEMS Microbiology Ecology *40*, 1–12.

Scanlan, D.J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W.R., Post, a F., Hagemann, M., Paulsen, I.T., and Partensky, F. (2009). Ecological genomics of marine picocyanobacteria. Microbiology and Molecular Biology Reviews: MMBR *73*, 249–299.

Scanlan, P.D., Hall, A.R., Blackshields, G., Friman, V.-P., Davis Jr, M.R., Goldberg, J.B., and Buckling, A. (2015). Coevolution with bacteriophages drives genome-wide host evolution and constrains the acquisition of abiotic-beneficial mutations. Molecular Biology and Evolution *32*, 1425–1435.

Schneider, D., Fuhrmann, E., Scholz, I., Hess, W.R., and Graumann, P.L. (2007). Fluorescence staining of live cyanobacterial cells suggest non-stringent chromosome segregation and absence of a connection between cytoplasmic and thylakoid membranes. BMC Cell Biology *8*, 39.

Sedlazeck, F.J., Rescheneder, P., and Von Haeseler, A. (2013). NextGenMap: Fast and accurate read mapping in highly polymorphic genomes. Bioinformatics *29*, 2790–2791.

Sekiguchi, H., Moriya, M., Nakayama, T., and Inouye, I. (2002). Vestigial chloroplasts in heterotrophic stramenopiles *Pteridomonas danica* and *Ciliophrys infusionum* (Dictyochophyceae). Protist *153*, 157–167.

Serrato, R. V. (2014). Lipopolysaccharides in diazotrophic bacteria. Frontiers in Cellular and Infection Microbiology *4*, 1–6.

Shan, J., Jia, Y., Clokie, M.R.J., and Mann, N.H. (2008). Infection by the 'photosynthetic' phage S-PM2 induces increased synthesis of phycoerythrin in *Synechococcus* sp. WH7803. FEMS Microbiology Letters *283*, 154–161.

Shannon, S.P., Chrzanowski, T.H., and Grover, J.P. (2007). Prey food quality affects flagellate ingestion rates. Microbial Ecology *53*, 66–73.

Sharon, I., Alperovitch, A., Rohwer, F., Haynes, M., Glaser, F., Atamna-Ismaeel, N., Pinter, R.Y., Partensky, F., Koonin, E. V, Wolf, Y.I., et al. (2009). Photosystem I gene cassettes are present in marine virus genomes. Nature *461*, 258–262.

Sherr, E.B., and Sherr, B.F. (2002). Significance of predation by protists in aquatic microbial food webs. Antonie van Leeuwenhoek, International Journal of General and Molecular Microbiology *81*, 293–308.

Shibata, A., Goto, Y., Saito, H., Kikuchi, T., Toda, T., and Taguchi, S. (2006). Comparison of SYBR Green I and SYBR Gold stains for enumerating bacteria and viruses by epifluorescence microscopy. Aquatic Microbial Ecology *43*, 223–231.

Siegele, D. a, Hu, J.C., and Gross, C. a (1988). Mutations in *rpoD*, the gene encoding the sigma 70 subunit of *Escherichia coli* RNA polymerase, that increase expression of the *lac* operon in the absence of CAP-cAMP. Journal of Molecular Biology *203*, 29–37.

Silverman, J.A., and Benson, S.A. (1987). Bacteriophage K20 requires both the OmpF porin and lipopolysaccharide for receptor function. Journal of Bacteriology *169*, 4830–4833.

Simkovsky, R., Daniels, E.F., Tang, K., Huynh, S.C., Golden, S.S., and Brahamsha, B. (2012). Impairment of O-antigen production confers resistance to grazing in a model amoeba-cyanobacterium predator-prey system. Proceedings of the National Academy of Sciences *109*, 16678–16683.

Simon, R.D. (1977). Macromolecular composition of spores from the filamentous cyanobacterium *Anabaena cylindrica*. Journal of Bacteriology *129*, 1154–1155.

Simon, R.D. (1980). DNA content of heterocysts and spores of the filamentous cyanobacterium *Anabaena variabilis*. FEMS Microbiology Letters *8*, 241–245.

Six, C., Thomas, J.-C., Garczarek, L., Ostrowski, M., Dufresne, A., Blot, N., Scanlan, D.J., and Partensky, F. (2007a). Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: a comparative genomics study. Genome Biology *8*, R259.

Six, C., Joubin, L., Partensky, F., Holtzendorff, J., and Garczarek, L. (2007b). UV-induced phycobilisome dismantling in the marine picocyanobacterium *Synechococcus* sp. WH8102. Photosynthesis Research *92*, 75–86.

Snyder, D.S., Brahamsha, B., Azadi, P., and Palenik, B. (2009). Structure of compositionally simple lipopolysaccharide from marine *Synechococcus*. Journal of Bacteriology *191*, 5499–5509.

Sohm, J. a, Ahlgren, N. a, Thomson, Z.J., Williams, C., Moffett, J.W., Saito, M. a, Webb, E. a, and Rocap, G. (2015). Co-occurring *Synechococcus* ecotypes occupy four major oceanic regimes defined by temperature, macronutrients and iron. The ISME Journal 1–13.

Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR - a widespread system that provides acquired resistance against phages in bacteria and archaea. Nature Reviews. Microbiology *6*, 181–186.

Spence, E. Mac (2010). Characterisation of cyanophage- resistance in marine

Synechococcus WH7803. PhD Thesis, University of Warwick.

Spencer, R. (1955). A marine bacteriophage. Nature *175*, 690–691.

Sperandeo, P., Dehò, G., and Polissi, A. (2009). The lipopolysaccharide transport system of Gram-negative bacteria. Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids *1791*, 594–602.

Stent, G.S. (1963). Molecular biology of bacterial viruses (San Francisco, W. H. Freeman).

Sternberg, N., and Austin, S. (1981). The maintenance of the P1 plasmid prophage. Plasmid *5*, 20–31.

Steward, G.F., Montiel, J.L., and Azam, F. (2000). Genome size distributions indicate variability and similarities among marine viral assemblages from diverse environments. Limnology and Oceanography *45*, 1697–1706.

Steward, G.F., Culley, A.I., Mueller, J. a, Wood-Charlson, E.M., Belcaid, M., and Poisson, G. (2013). Are we missing half of the viruses in the ocean? The ISME Journal *7*, 672–679.

Stoddard, L.I., Martiny, J.B.H., and Marston, M.F. (2007). Selection and characterization of cyanophage resistance in marine *Synechococcus* strains. Applied and Environmental Microbiology *73*, 5516–5522.

Sullivan, M.B., Huang, K.H., Ignacio-Espinoza, J.C., Berlin, A.M., Kelly, L., Weigele, P.R., DeFrancesco, A.S., Kern, S.E., Thompson, L.R., Young, S., et al. (2010). Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. Environmental Microbiology *12*, 3035–3056.

Sun, S., Chen, J., Li, W., Altintas, I., Lin, A., Peltier, S., Stocks, K., Allen, E.E., Ellisman, M., Grethe, J., et al. (2011). Community cyberinfrastructure for advanced microbial ecology research and analysis: The CAMERA resource. Nucleic Acids Research *39*, D546–D551.

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

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

Suttle, C.A. (2007). Marine viruses - major players in the global ecosystem. Nature Reviews Microbiology *5*, 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*, 99–109.

Suttle, C.A., and Chan, A.M. (1994). Dynamics and distribution of cyanophages and their effect on marine *Synechococcus* spp. Applied and Environmental Microbiology *60*, 3167–3174.

Suttle, C.A., Chan, A.M., and Cottrell, M.T. (1990). Infection of phytoplankton by viruses and reduction of primary productivity. Nature.

Taylor, B.L., and Zhulin, I.B. (1999). PAS Domains : Internal Sensors of Oxygen, Redox Potential and Light. Microbiology and Molecular Biology Reviews: MMBR *63*, 479–506.

Thingstad, T.F. (2000). Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnology and Oceanography *45*, 1320–1328.

Thingstad, T.F., and Lignell, R. (1997). Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. Aquatic Microbial Ecology *13*, 19–27.

Thompson, L.R., Zeng, Q., Kelly, L., Huang, K.H., Singer, a. U., Stubbe, J., and Chisholm, S.W. (2011). Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. Proceedings of the National Academy of Sciences *108*, E757–E764.

Tillich, U.M., Wolter, N., Franke, P., Dühring, U., and Frohme, M. (2014). Screening and genetic characterization of thermo-tolerant *Synechocystis* sp. PCC6803 strains created by adaptive evolution. BMC Biotechnology *14*, 66.

Ting, C.S., Rocap, G., King, J., and Chisholm, S.W. (2002). Cyanobacterial photosynthesis in the oceans: The origins and significance of divergent light-harvesting strategies. Trends in Microbiology *10*, 134–142.

Tobiason, D.M., and Seifert, H.S. (2006). The obligate human pathogen, *Neisseria gonorrhoeae*, is polyploid. PLoS Biology *4*, 1069–1078.

Tobiason, D.M., and Seifert, H.S. (2010). Genomic content of *Neisseria* species. Journal of Bacteriology *192*, 2160–2168.

Tock, M.R., and Dryden, D.T.F. (2005). The biology of restriction and anti-restriction. Current Opinion in Microbiology *8*, 466–472.

Toledo, G., Palenik, B., and Brahamsha, B. (1999). Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratios form a monophyletic group. Applied and Environmental Microbiology *65*, 5247–5251.

Torrella, F., and Morita, R.Y. (1979). Evidence by electron micrographs for a high incidence of bacteriophage particles in the waters of Yaquina Bay, Oregon: Ecological and taxonomical implications. Applied and Environmental Microbiology *37*, 774–778.

Traurig, M., and Misra, R. (1999). Identification of bacteriophage K20 binding regions of OmpF and lipopolysaccharide in *Escherichia coli* K-12. FEMS Microbiology Letters *181*, 101–108.

Traving, S.J., Clokie, M.R.J., and Middelboe, M. (2014). Increased acidification has a profound effect on the interactions between the cyanobacterium *Synechococcus* sp. WH7803 and its viruses. FEMS Microbiology Ecology *87*, 133–141.

Trent, M.S. (2004). Biosynthesis, transport, and modification of lipid A. Biochemistry and Cell Biology *82*, 71–86.

Tsukaya, H. (2013). Does ploidy level directly control cell size? Counterevidence from *Arabidopsis* genetics. PLoS ONE *8*, 1–7.

Urbach, E., Robertson, D.L., and Chisholm, S.W. (1992). Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. Nature *355*, 267–270.

Urbach, E., Scanlan, D.J., Distel, D.L., Waterbury, J.B., and Chisholm, S.W. (1998). Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (cyanobacteria). Journal of Molecular Evolution *46*, 188–201.

van Valen, L. (1973). A new evolutionary law. Evolutionary Theory *1*, 1–30.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahe, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., et al. (2015). Eukaryotic plankton diversity in the sunlit ocean. Science *348*, 1261605–1261605.

Vaulot, D., Marie, D., Olson, R.J., and Chisholm, S.W. (1995). Growth of *Prochlorococcus*, a photosynthetic prokaryote, in the equatorial Pacific Ocean. Science *268*, 1480–1482.

Veldhuis, M.J.W., Kraay, G.W., Van Bleijswijk, J.D.L., and Baars, M. a (1997). Seasonal and spatial variability in phytoplankton biomass, productivity and growth in the northwestern Indian Ocean. Deep Sea Research Part I *44*, 425–449.

Villar, E., Farrant, G.K., Follows, M., Garczarek, L., Speich, S., Audic, S., Bittner, L., Blanke, B., Brum, J.R., Brunet, C., et al. (2015). Environmental characteristics of Agulhas rings affect interocean plankton transport. Science *348*, 1261447–1261447.

Vincent, W.F., Bowman, J.P., Rankin, L.M., and McMeekin, T.A. (2000). Phylogenetic diversity of picocyanobacteria in Arctic and Antarctic ecosystems. Polar Microbiology Ecology 317–322.

Wang, K., and Chen, F. (2008). Prevalence of highly host-specific cyanophages in the estuarine environment. Environmental Microbiology *10*, 300–312.

Wang, X., and Quinn, P.J. (2010). Lipopolysaccharide: Biosynthetic pathway and structure modification. Progress in Lipid Research *49*, 97–107.

Waterbury, J.B., and Valois, F.W. (1993). Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. Applied and Environmental Microbiology *59*, 3393–3399.

Waterbury, J.B., Watson, S., Valois, F.W., and Franks, D. (1986). Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. In Photosynthetic Picoplankton, p. 71 – .

Webb, C.D., Graumann, P.L., Kahana, J. a., Teleman, A. a., Silver, P. a., and Losick, R. (1998). Use of time-lapse microscopy to visualize rapid movement of the replication origin region of the chromosome during the cell cycle in *Bacillus subtilis*. Molecular Microbiology *28*, 883–892.

Weigele, P.R., Pope, W.H., Pedulla, M.L., Houtz, J.M., Smith, A.L., Conway, J.F., King, J., Hatfull, G.F., Lawrence, J.G., and Hendrix, R.W. (2007). Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. Environmental Microbiology *9*, 1675–1695.

Weinbauer, M.G. (2004). Ecology of prokaryotic viruses. FEMS Microbiology Reviews *28*, 127–181.

Weinbauer, M.G., and Peduzzi, P. (1994). Frequency, size and distribution of bacteriophages in different marine bacterial morphotypes. Marine Ecology Progress Series *108*, 11–20.

Weinbauer, M.G., and Rassoulzadegan, F. (2004). Are viruses driving microbial diversification and diversity? Environmental Microbiology *6*, 1–11.

Weitz, J.S., Stock, C. a, Wilhelm, S.W., Bourouiba, L., Coleman, M.L., Buchan, A., Follows, M.J., Fuhrman, J. a., Jover, L.F., Lennon, J.T., et al. (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. The ISME Journal *9*, 1352–1364.

West, N.J., and Scanlan, D.J. (1999). Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. Applied and Environmental Microbiology *65*, 2585–2591.

Westra, E.R., and Brouns, S.J.J. (2012). The rise and fall of CRISPRs - dynamics of spacer acquisition and loss. Molecular Microbiology *85*, 1021–1025.

Whitfield, C. (2006). Biosynthesis and assembly of capsular polysaccharides. Annual Review of Biochemistry *75*, 39–68.

Whitfield, C., and Larue, K. (2008). Stop and go: regulation of chain length in the biosynthesis of bacterial polysaccharides. Nature Structural & Molecular Biology *15*, 121–123.

Wiggins, B. a., and Alexander, M. (1985). Minimum bacterial density for bacteriophage replication: implications for significance of bacteriophages in natural ecosystems. Applied and Environmental Microbiology *49*, 19–23.

Wilhelm, S.W., and Suttle, C.A. (1999). Virus and nutrient cycles in the sea. BioScience *49*, 781–788.

Wilson, W.H., Joint, I.R., Carr, N.G., and Mann, N.H. (1993). Isolation and molecular characterization of five marine cyanophages propagated on *Synechococcus* sp. strain WH7803. Applied and Environmental Microbiology *59*, 3736–3743.

Wilson, W.H., Carr, N.G., and Mann, N.H. (1996). The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium *Synechococcus* sp WH7803. Journal of Phycology *32*, 506–516.

Winget, D.M., and Wommack, K.E. (2009). Diel and daily fluctuations in virioplankton production in coastal ecosystems. Environmental Microbiology *11*, 2904–2914.

Winter, C., Bouvier, T., Weinbauer, M.G., and Thingstad, T.F. (2010). Trade-offs between competition and defense specialists among unicellular planktonic organisms: the 'killing the winner' hypothesis revisited. Microbiology and Molecular Biology Reviews: MMBR *74*, 42–57.

Wommack, K.E., and Colwell, R.R. (2000). Virioplankton: viruses in aquatic ecosystems. Microbiology and Molecular Biology Reviews: MMBR *64*, 69–114.

Wood, A.M., Lipsen, M., and Coble, P. (1999). Fluorescence-based characterization of phycoerythrin-containing cyanobacterial communities in the Arabian Sea during the Northeast and early Southwest Monsoon (1994-1995). Deep-Sea Research Part II: Topical Studies in Oceanography *46*, 1769–1790.

Woolhouse, M.E.J., and Webster, J.P. (2000). In search of the Red Queen. Parasitology Today *16*, 506–508.

Woolhouse, M.E.J., Webster, J.P., Domingo, E., Charlesworth, B., and Levin, B.R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. Nature Genetics *32*, 569–577.

Worden, A.Z., and Binder, B.J. (2003). Application of dilution experiments for measuring growth and mortality rates among *Prochlorococcus* and *Synechococcus* populations in oligotrophic environments. Aquatic Microbial Ecology *30*, 159–174.

Wu, T., McCandlish, A.C., Gronenberg, L.S., Chng, S.-S., Silhavy, T.J., and Kahne, D. (2006). Identification of a protein complex that assembles lipopolysaccharide in the outer membrane of *Escherichia coli*. Proceedings of the National Academy of Sciences *103*, 11754–11759.

Xu, X., Khudyakov, I., and Wolk, C.P. (1997). Lipopolysaccharide dependence of cyanophage sensitivity and aerobic nitrogen fixation in *Anabaena* sp. strain PCC 7120. Journal of Bacteriology *179*.

Zeng, Q., and Chisholm, S.W. (2012). Marine viruses exploit their host's two-component regulatory system in response to resource limitation. Current Biology *22*, 124–128.

Zerulla, K., and Soppa, J. (2014). Polyploidy in haloarchaea: Advantages for growth and survival. Frontiers in Microbiology *5*, 274.

Zerulla, K., Chimileski, S., Nather, D., Gophna, U., Papke, R.T., and Soppa, J. (2014). DNA as a phosphate storage polymer and the alternative advantages of polyploidy for growth or survival. PLoS ONE *9*, e94819.

Zhong, Y., Chen, F., Wilhelm, S.W., Poorvin, L., and Hodson, R.E. (2002). Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. Applied and Environmental Microbiology *68*, 1576–1584.

Zobell, C.E. (1946). Marine microbiology, a monograph on hydrobacteriology (Waltham, Mass., Chronica Botanica Company).

Zwirglmaier, K., Heywood, J.L., Chamberlain, K., Woodward, E.M.S., Zubkov, M. V., and Scanlan, D.J. (2007). Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. Environmental Microbiology *9*, 1278–1290.

Zwirglmaier, K., Jardillier, L., Ostrowski, M., Mazard, S., Garczarek, L., Vaulot, D., Not, F., Massana, R., Ulloa, O., and Scanlan, D.J. (2008). Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Environmental Microbiology *10*, 147–161.

Zwirglmaier, K., Spence, E.M., Zubkov, M. V., Scanlan, D.J., and Mann, N.H. (2009). Differential grazing of two heterotrophic nanoflagellates on marine *Synechococcus* strains. Environmental Microbiology *11*, 1767–1776.