Isolation and Characterisation of Novel Viruses Infecting Marine Phytoplankton

by

Karen Dawn Weynberg

A thesis submitted in partial fulfilment of the requirements for the degree of
Doctor of Philosophy in Biological Sciences

University of Warwick, Department of Biological Sciences

June 2009

Table of Contents

List of Figures and Tables

Acknowledgements

Many thanks must be expressed to a number of people who helped me during the time of this study. Firstly, I would like to express my gratitude to my supervisors Willie Wilson, Mike Allen and Dave Scanlan for their guidance, scientific input and support throughout my PhD. Special thanks to Mike who had to put up with me and my questions on a regular basis but always did so with a smile!

Thanks are due to the members of various incarnations of Team Willie and Team Virus: Susan, Claire, Jayme, Joaquin, Antonio, Julio, John W-A, Paul, and Charlie. Particular thanks to Declan for my initial induction to the world of algal virology and firing my enthusiasm for this field of research. Most especially, I would like to express my deep gratitude to Andrea Baker, who patiently guided me round the molecular lab and all the techniques needed to hunt for viruses.

Thanks also to my fellow students for chats, understanding and nudging me to take a break now and then: Nicole, Hannah, Helen, Kate, Bonnie, Emmer, James and Simon.

Many thanks and love to my parents whose support, love and encouragement have meant so much to me. Thanks also to Sheila, Tor, Jacko and Fred for your kind words, support, great family holidays and help with child care!

Finally, but most importantly, the biggest thanks must go to my wonderful husband, Allen, and our three beautiful children, Rebecca, Christopher and Finlay. Your unerring support, understanding, love, patience and encouragement have given me the determination to make my dreams reality. Thank you for seeing me through the tougher moments. Without your help, this would never have reached fruition. I love you all very much.

This thesis is dedicated to my family.

Declaration of Authorship

I, KAREN DAWN WEYNBERG declare that the thesis entitled

Isolation and Characterisation of Novel Viruses Infecting Marine Phytoplankton

and the work presented in it are my own.

- This work was done wholly or mainly while in candidature for a research degree at the University of Warwick;
- No part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution;
- Where I have consulted the published work of others, this is always clearly attributed;
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- None of this work has been published or submitted for publication prior to the beginning of the period of this study.

Signed: ……

Date: ……

Abstract

Viruses are the most abundant biological agents in the global marine environment. Through cellular lysis viruses influence many biogeochemical and ecological processes, including energy and nutrient cycling, host distribution and abundance, algal bloom control and genetic transfer. Nano- and picophytoplankton are ubiquitous in the world's oceans and are responsible for a high proportion of the annual global carbon fixation. However, relatively few viruses have been isolated and described that infect these key primary producers and little is known of their diversity, dynamics or propagation strategies. The aims of this study were to detect, isolate and characterise novel marine viruses that infect these important members of the phytoplankton assemblage.

Screening of seawater samples for viruses that infect a broad representation of nano- and picophytoplankton species was undertaken here. To enable this, a large culture collection of 106 phytoplankton species was established and used to screen seawater samples for viruses on a weekly basis over a two year period. A total of 12 novel viruses infecting the prasinophyte species' *Ostreococcus tauri* and *Micromonas pusilla* were isolated from seawater sampled in coastal waters of the Western English Channel. Viruses were purified by plaque purification or liquid serial dilution techniques. Characterisation of novel virus isolates included growth kinetics, visualisation using transmission electron microscopy, host range analysis and estimates of viral genome sizes using pulsed field gel electrophoresis. Phylogenetic analysis of these viruses was conducted based on the sequence of the conserved DNA polymerase gene. Genome sequencing of two of the viruses infecting *O. tauri* was completed and revealed many exciting features, including a suite of genes hitherto unreported, or with rare occurrence, in viruses. Evidence is presented for horizontal gene transfer between viruses isolated in this study and their hosts, as well as between other eukaryotic and bacterial sources. Functional characterisation of the viral genomes sequenced and described in this study will provide clearer insights into viral dynamics and evolutionary history.

Abbreviations

µl	microlitre
µm	micrometre
µM	micromole
Φ	bacteriophage
ACT	Artemis Comparison Tool
AFC	analytical flow cytometry
ASW	artificial seawater
bp	base pair
CDS	coding sequence
CTAB	cetyltrimethylammonium bromide
DNA	deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DOC	dissolved organic carbon
dsDNA	double stranded DNA
EDTA	ethylenediaminetetraacetic acid
HAB	Harmful Algal Bloom
l	litre
LB	Luria Bertani
MCP	Major Capsid Protein
MOI	Multiplicity of Infection
MilliQ/MQ	18.2Ω, high purity, reverse osmosed, UV treated, 0.2µm filtered water created using the Millipore Milli-Q system
NaCl	sodium chloride
NCLDV	nucleocytoplasmic large DNA viruses
nm	nanometre
ORF	open reading frame
PCC	Plymouth Culture Collection
PCR	polymerase chain rection
PEG	polyethylene glycol
pers. comm.	personal communication

PFGE	Pulsed Field Gel Electrophoresis
PFU	plaque forming unit
p.i.	post-infection
pmol	picomoles
POC	particulate organic carbon
pol	DNA polymerase gene
PPEs	Photosynthetic picoeukaryotes
RCC	Roscoff Culture Collection
RCF	relative centrifugal force
RNA	ribonucleic acid
rRNA	ribosomal RNA
rpm	revolutions per second
SOC	growth medium designed to ensure maximum transformation efficiency (*E.coli*)
sp.	Species
SYBR Green	fluorescent dye for staining nucleic acid
Taq	DNA polymerase from *Thermus aquaticus*
TEM	Transmission Electron Microscopy
TMV	Tobacco Mosaic Virus
tRNA	transfer RNA
unpubl.	unpublished
UV	Ultra Violet
VLP	Virus-Like Particle
X-Gal	5-bromo-4-chloro-3-indoyl β-D-galactopyranoside

Publications arising from this PhD study

Weynberg, K.D., Allen, M.J., Ashelford, K., Scanlan, D.J., and Wilson, W.H. (2009) From small hosts come big viruses: the complete genome of a second *Ostreococcus tauri* virus, OtV-1. *Environmental Microbiology.* In press.

Chapter 1

Introduction

1.1. General introduction to the viruses

1.1.1. Historical aspects of virology

The field of virology arose towards the end of the nineteenth century. In 1892, a filterable, submicroscopic element was identified as being linked to tobacco mosaic disease (Ivanovsky, 1892). Six years later, Beijerinck (1898) published similar findings and termed this new concept as filterable 'contagium vivum fluidum' (contagious living fluid). The first filterable agent from animals, the foot-and-mouth disease virus, was then described (Loeffler and Frosch, 1898) and in 1901 the first human virus that causes yellow fever was identified (Reed, 1901). The nature of viruses as liquid or particulate was then debated until the advent of a plaque assay (d'Herelle, 1917) and the first electron micrographs of tobacco mosaic virus (TMV) in the early 20th century (Kausche and Ruska, 1939). Thus, characterisation studies of TMV played a key role in the origins of virology. Virus particles were further 'visualised' by X-ray crystallography (Bernal and Fankuchen, 1941) and in 1956 the TMV genome was used to prove that infectious or genetic information could be stored in RNA molecules (Fraenkel-Conrat *et al.,* 1957). The concept of self-assembly with an infectious virus was pioneered by using TMV RNA and coat protein (Butler and Klug, 1971). Research efforts into viruses, including those that infect bacteria (bacteriophages) and animal viruses, during the latter half of the 20th century to the present day have made extraordinary contributions to the life sciences.

1.1.2. What are viruses?

Viruses are typically small particles usually measuring in the range 20 – 200 nm, although the recently discovered Mimivirus is estimated to be 400 nm in diameter (La Scola *et al.,* 2003). Viruses consist of genetic material (nucleic acid - either DNA or

RNA, single- or double-stranded) surrounded by a protein coat called a capsid. Some may also contain lipids (Alberts *et al.,* 2002). Viruses are incapable of independent growth and metabolism and therefore require the intracellular apparatus of their hosts to facilitate their replication (Alberts *et al.,* 2002). As non-motile entities, viruses contact their hosts by passive diffusion. Adsorption to the host occurs via attachment to external structures on the cell membrane and entry is gained.

Viruses that infect bacteria are termed bacteriophages or, more commonly, phages. This distinguishes them from viruses that infect eukaryotes. Phage T2 and phage T4 are just two examples of types of phage that specifically infect the bacterium *Escherichia coli.* More specific definitions include cyanophages describing phage that infect cyanobacteria, which are an important group of primary producers in the global aquatic environment (Castenholz and Waterbury, 1989).

1.1.3. Modes of infection

There are various modes of viral infection. These include lytic infection, lysogeny, pseudolysogeny and chronic infection (Dimmock *et al.,* 2001). The two principal modes of infection are lysis and lysogeny (Fig. 1). Bacteriophages may have a lytic or lysogenic cycle although a number of them are capable of carrying out both propagation strategies. Lytic phages such as T4, lyse bacterial cells causing the demise of the host, following viral replication and assembly of the virion progeny.

Lytic infection occurs when the virus attaches to the host cell, injects its nucleic acid and redirects the host replication machinery to produce progeny viruses. These new

viruses emerge from the cell and are released to begin the cycle again (Fig. 1.1). The propagation strategy of a temperate phage is outlined in Figure 1.1.

A lytic phage would follow the lytic cycle lifestyle only, unless a mutation arose which enabled it to switch to a lysogenic state. Lysogeny in phage occurs when, following injection, the phage nucleic acid forms a stable association with, and incorporates into, the host genome (Ackermann and DuBow, 1987). The phage genetic material is replicated along with that of the host. Lysogeny confers protection for the phage and immunity to the host cell from lytic infections by similar phage. It has been theorised that lysogeny dominates when the host abundance is too low (due to low resources or high viral lysis rates) to sustain the lytic pathway (Hennes *et al.,* 1995; Weinbauer *et al.,* 2003). Phage can persist in this lysogenic state for several generations until an environmental or biochemical 'trigger' event occurs, which causes the phage to switch to the lytic pathway. This induction to the lytic pathway may be caused by damage to the host DNA (e.g. UV light damage) or rapid host cell growth, due to optimal conditions. The latter could imply the virus has an increased chance of finding new host cells to infect. For photosynthetic hosts, lysogenic associations between cyanophage and filamentous cyanobacteria have been reviewed by Sime-Ngando and Colombet (2009) and Brüssow *et al.* (2004). The only well-documented lysogenic association with unicellular cyanobacteria is of a virus infecting the marine *Synechococcus* strain NKBG 042902 (Sode *et al.,* 1994). To date, lysogeny in eukaryotic algal viruses has only been described in macroalgal species (Muller *et al.,* 1998) (Table 1.1). Viruses infecting macroalgae only infect free-swimming host zoospore or gametes and the virus genome then integrates into the host genome (Fig. 1.2).

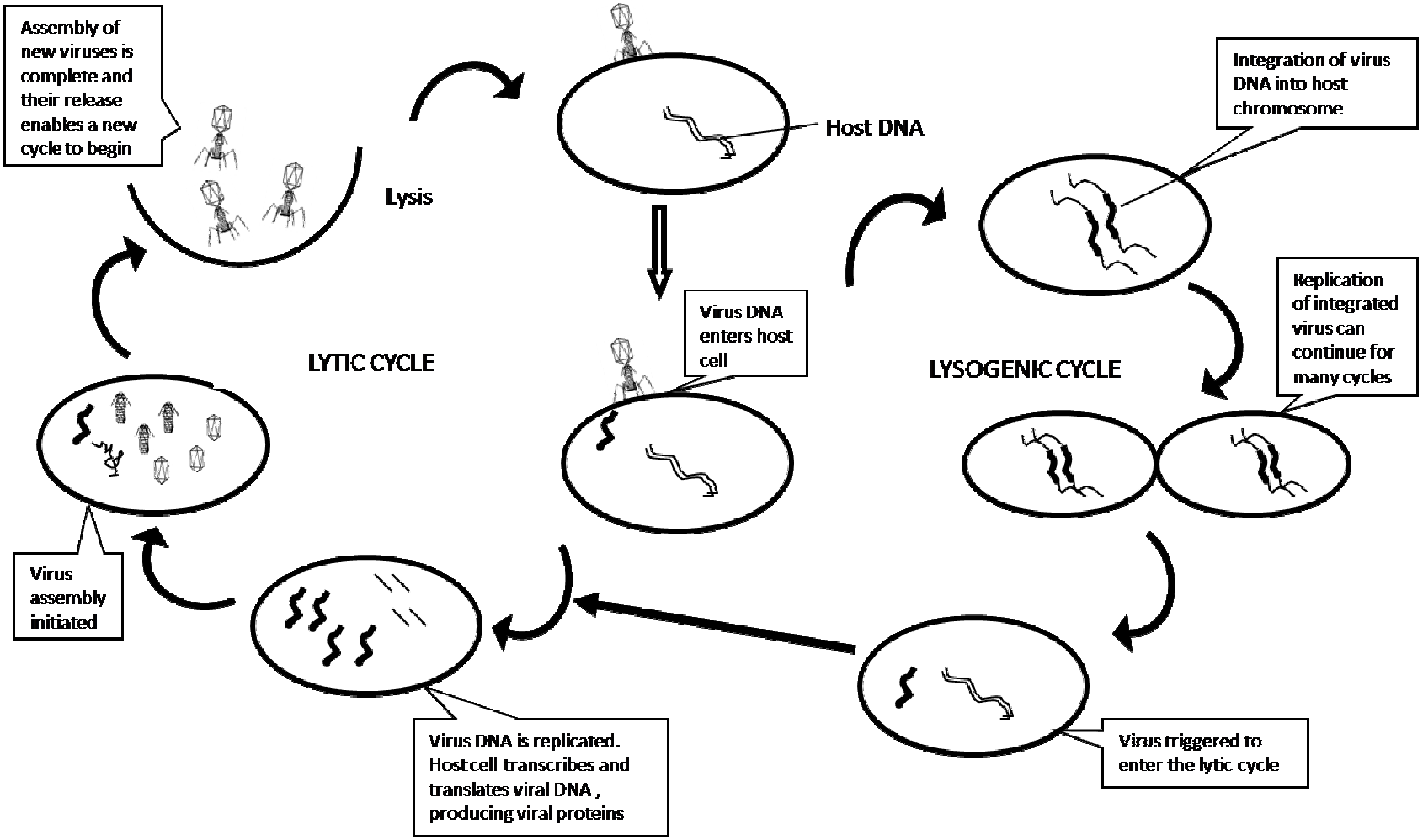


Figure 1.1. Life cycle of a temperate phage.

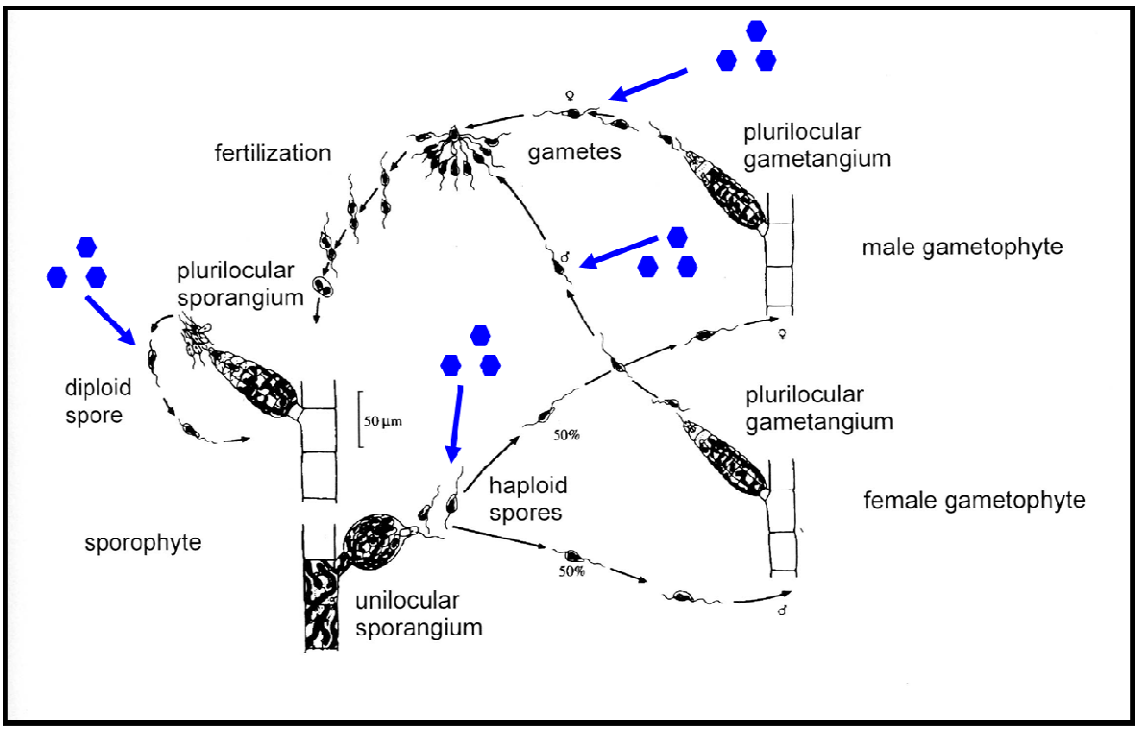


Figure 1.2. The life-cycle of filamentous brown ectocarpoid macroalgal species. Viruses (blue hexagonals) only infect free-swimming zoospore or gamete stages of hosts. The virus genome is integrated into the host genome and is inherited in a Mendelian manner. Virus particles are only formed in prospective gametangia or sporangia cells of the host. Adapted from Van den Hoek *et al.* (1995).

Persistent infections, such as pseudolysogeny, have been described for certain infections of host bacteria by bacteriophages (Miller and Ripp, 1998, 2002). Pseudolysogeny is defined as an unstable coexistence of a bacteriophage in a host bacterium without a constant inheritance of the phage genome (Ackermann and DuBow, 1987). This leads to both infected and phage sensitive progeny in the same culture. Chronic infection arises when the host cell releases the viral progeny by extrusion, leading to a slow rate of infection (Flint *et al.,* 2004).

To date, all viruses infecting microalgae have been described as following a lytic infection process only (Suttle 2007). A representation of the evolutionary relationships among virus taxa, e.g. using a phylogenetic tree, is not possible due to the polyphyletic nature of viruses (Moreira and López-García, 2009). A single gene is not shared by all viruses and hence from a molecular phylogeny perspective, viruses cannot be compared among each other. However, DNA polymerase gene fragments of viruses which infect microalgae and are members of the *Phycodnaviridae* and other large dsDNA viruses have been used to construct a phylogenetic tree (Fig. 1.3).

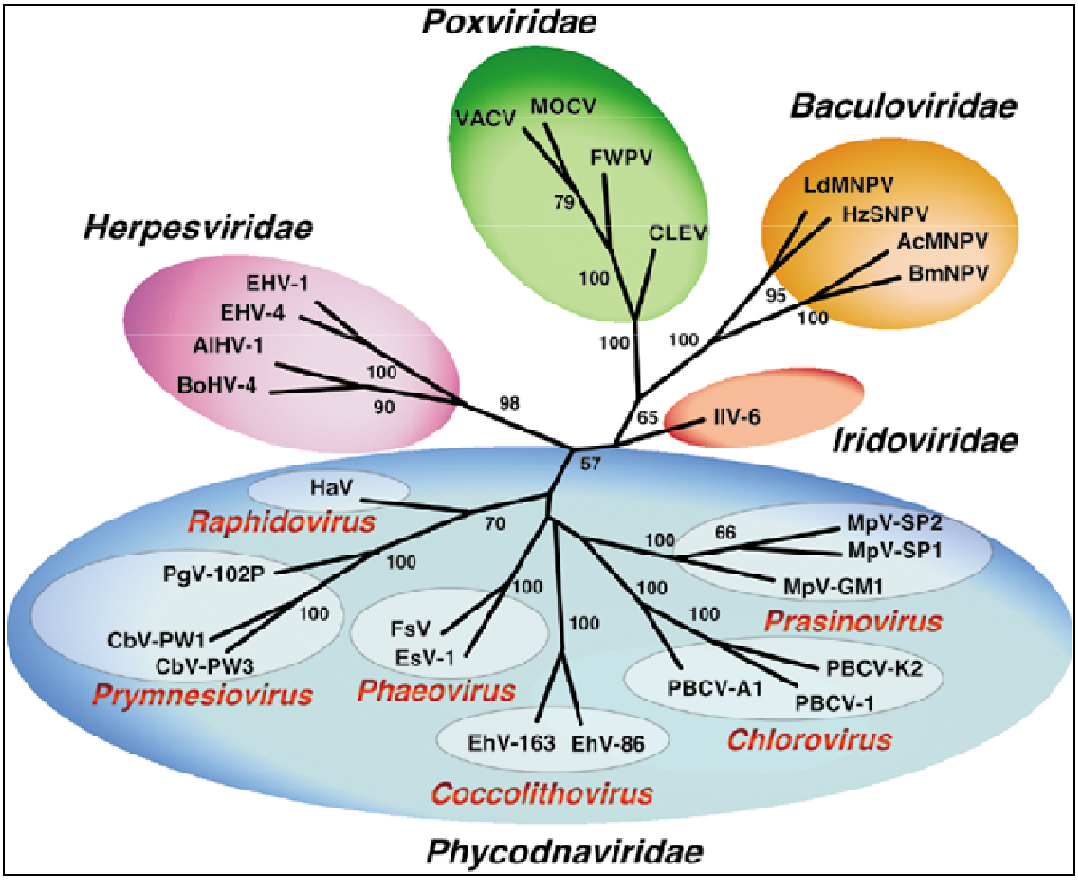


Figure 1.3. Phylogenetic tree of DNA polymerase gene fragments from members of the family *Phycodnaviridae* and other large dsDNA viruses generated using protein parsimony analysis of 100 bootstrapped data sets. Numbers at branches indicate the relative strengths of those branches. Adapted from Schroeder *et al.* (2002).

1.2. Viruses in the World's Oceans

1.2.1. Virus Abundance and Ecological Aspects

Viruses are now widely accepted to be ubiquitous and abundant components of marine ecosystems (Suttle 2007). Historically, the roles of microbial populations, and by association, viruses, were largely ignored in marine ecosystems. It has only been in recent decades that the true abundance and diversity of bacteria and viruses have been revealed (Hobbie *et al.,* 1977; Azam *et al.,* 1983; Bergh *et al.,* 1989). As the importance of microbial populations has been recognised, so has the potential significance of viruses as a major cause of mortality in prokaryotes and phytoplankton. The far reaching impacts of such discoveries have been duly recognised and have made marine virology a rapidly expanding field of study. Quantification techniques have led to estimates of viral abundances of tens of millions per millilitre in the world's oceans (Brussaard, 2004; Breitbart and Rohwer, 2005). Viral numbers typically exceed 10^8 ml^{-1} in coastal surface waters. This outnumbers total prokaryotic abundance by as much as 5 – 25 times (Suttle 2005).

Viral-induced lysis is a major cause of prokaryotic mortality in marine ecosystems and can equal grazing-induced mortality of bacteria (Wommack and Colwell, 2000; Sandaa, 2008). Studies show that bacterioplankton mortality due to viral infection is between 15 – 20 % day^{-1} (Wommack and Colwell, 2000). The lytic infection cycle has been well-documented in microalgae (Suttle, 2007). As described for phages, viral nucleic acid is injected into the microalgal host cell and transcription, translation and replication of progeny viruses begins. The host cell is eventually lysed to release progeny virions which are available to other host cells. Lysis of cells contributes significant organic matter to the surrounding environment and this recycling of material through viral lysis is termed the viral shunt (Wilhelm and Suttle, 1999).

Host-virus interactions have important consequences for the entire marine food web. Through the lysis of host cells, marine viruses influence many ecological and biogeochemical processes, including energy and nutrient cycling, algal bloom control and genetic transfer (Wommack and Colwell, 2000). Although the importance of viruses is largely recognised, many aspects of viral-host interactions are poorly understood. Developments in key techniques such as flow cytometry, epifluorescence microscopy, genomics and proteomics will help to advance the evaluation and quantification of viral impact in marine ecosystems.

Viruses can be viewed as dynamic members of marine ecosystems (Wilhelm and Suttle, 1999; Weinbauer and Rassoulzadegan, 2004). The major primary producers in marine ecosystems include the prokaryotes (cyanobacteria and prochlorophytes) and eukaryotic microalgae and macroalgae. All of these groups have been shown to be susceptible to infection by viruses (Sandaa, 2008). This has challenged the accepted views of the marine pelagic food web by highlighting the crucial role played by viruses influencing population dynamics at all trophic levels (Thingstad *et al.,* 1993). Virus-induced cell lysis has major implications for the flow of energy and material in marine pelagic food webs. Viruses are potentially critical factors in the structure and function of marine food webs (Wilhelm and Suttle, 1999). There is, however, only a limited understanding of the occurrence and distribution of viruses in marine ecosystems and of the dynamics between viruses and host communities in the natural environment.

1.2.2. Impact of marine viruses at the population level

The specificity of viral infection indicates that viruses may selectively affect the occurrence and distribution of host strains within communities of plankton (Suttle, 2000; Short and Suttle, 2003). Viral infection rates are expected to rise as an increased

host population density leads to higher contact rates. As the most common host will be infected, it will be this population that is directly impacted upon. This concept is known as 'killing the winner' (Thingstad and Lignell, 1997). The presence of viruses thus ensures the coexistence of competing species and prevents a particular dominant host becoming established.

Viruses have been documented, and some isolated, as the major factor in the decline of certain phytoplankton blooms. The raphidophyte *Heterosigma akashiwo* forms harmful blooms that can be responsible for large fish kills. Viruses infecting this alga have been cultured (Nagasaki and Yamaguchi, 1997) and seen to have a role in termination of blooms of the species (Tarutani *et al.,* 2000). Viral infection has also been implicated as a major factor in the termination of blooms of the coccolithophorid *Emiliana huxleyi* (Schroeder *et al.,* 2002; Wilson *et al.,* 2002a; 2003). Vast blooms of this unicellular alga are known to disintegrate suddenly and viral lysis has been cited as a causative factor. Cloud-forming dimethyl sulphide (DMS) is released into the atmosphere during these population crashes (Holligan *et al.,* 1993). Viral activity therefore plays a crucial role in the production of DMS, a chemical known to influence climate change (Charlson *et al.,* 1987).

Viruses infecting the pelagophyte *Aureococcus anophagefferens*, which produces devastating brown tides, have also been reported (Milligan and Cosper, 1994; Garry *et al.,* 1998; Gastrich *et al.,* 1998; 2002; 2004; Gobler *et al.,* 2007). Viruses have been isolated for other phytoplankton species including *Phaeocystis pouchetti* (Jacobsen *et al.,* 1996); the novel shellfish-killing dinoflagellate *Heterocapsa circularisquama,* which forms red tides (Tarutani *et al.,* 2001); and the harmful bloom-causing

Phaeocystis globosa, (Bratbak *et al.,* 1998b; Baudoux and Brussaard, 2005; Baudoux *et al.,* 2006; Brussaard *et al.,* 2007).

1.2.3. Viruses and the marine pelagic food web.

The discovery that viruses infect all components of the food web has resulted in an increasingly complex picture of the food web structure (Thingstad *et al.,* 1993). Viral infection has direct implications for the host organism and this will in turn indirectly influence the food web as a whole (Wommack and Colwell, 2000).

The huge importance of microbes in aquatic marine ecosystems has been recognised only in recent decades (Pomeroy, 1974; Azam *et al.,* 1983; Sherr and Sherr, 1991). The discovery that bacterial abundance is typically as high as 10^9 l^{-1} (Hobbie *et al.,* 1977) helped elevate the significance of microbial populations in biological oceanography. The classic view of trophic interactions had been of a simple food chain in which primary production was attributed mainly to phytoplankton, which were grazed upon by herbivores. Larger zooplankton fed on the herbivorous grazers and were, in turn, consumed by a top predator, e.g. fish. Prior to improved isolation and quantification techniques, estimates of bacterial abundance were low (Fuhrman, 1999). Prokaryotes, therefore, were not included in this food chain and so remained largely ignored until about 25 years ago.

Photosynthetic prokaryotes are abundant and form a significant proportion of living carbon in the oceans. Cyanobacteria can represent a considerable percentage of the total bacteria in the euphotic zone [for example, *Prochlorococcus* can be 25% of total bacteria (Partensky *et al.,* 1999)] and are believed to fix 20 – 80 % of carbon in the marine environment (Wilhelm and Suttle, 1999). Heterotrophic prokaryotes are hugely

important in terms of total living carbon and material processed in the oceans. The combined biomass of heterotrophic and photosynthetic prokaryotes accounts for more than 90% of the total organic carbon in the marine environment (Weinbauer, 2004).

It is now accepted that bacterioplankton are responsible for a significant proportion (10 – 50%) of primary production in the oceans (Azam *et al.,* 1983). Heterotrophic bacteria, which utilize the dissolved products of photosynthesis, recycle the primary production lost to higher trophic levels back into biomass (Fig. 1.4). This, in turn, makes a larger proportion of primary production available to the marine food web. This is known as the microbial loop (Azam et al 1983). Carbon derived from photosynthesis is recycled several times through the food web (Weinbauer, 2004).

In marine food webs, two sources of carbon exist – Dissolved Organic Carbon (DOC) and Particulate Organic Carbon (POC). The two sources of carbon have different fates. As bacteria are the largest consumers of DOC, this carbon is recycled mainly through the microbial loop. In contrast, significant amounts of POC (including bacteria and other plankton) are available to higher trophic levels by grazing (Weinbauer, 2004). The latter route eventually links the microbial web to the classic food web (Fig. 1.4.). Virus-mediated cell lysis alters these budgets of carbon by diverting carbon from the POC pool to the DOC pool (Bratbak *et al.,* 1992; 1998a; Wilhelm and Suttle, 1999). Viruses are therefore directly responsible for the death of primary producers but are indirectly responsible for the recycling of material that fuels ongoing production. It is estimated that between 6% and 26% of organic carbon produced by photosynthesis ends in the DOC pool due to viral lysis (Weinbauer, 2004).

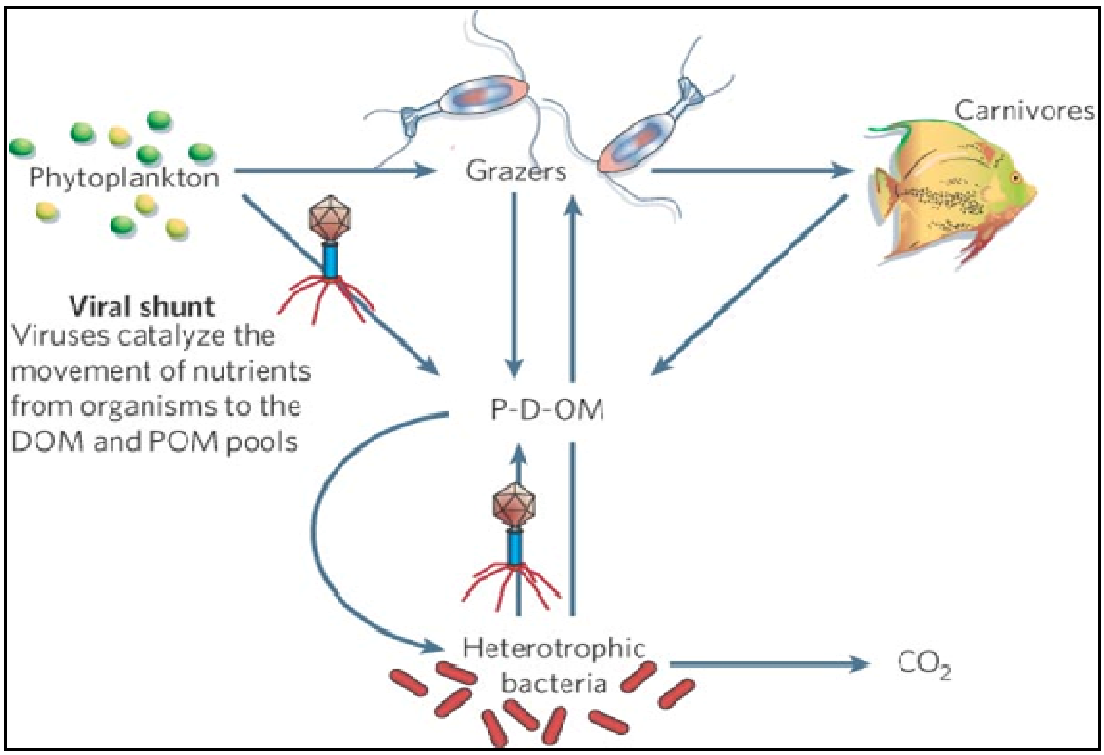


Figure. 1.4. Pelagic food chain model and virus-mediated nutrient flow. Through cellular lysis, viruses divert the flow of carbon and nutrients from phytoplankton and bacteria to higher trophic levels and shunt the flux to the pool of particulate and dissolved organic matter (P-D-OM). Adapted from Suttle (2007).

Viral lysis therefore diverts carbon, fixed as bacterial and phytoplankton biomass, away from mesozooplankton consumers into the DOC pool. This 'viral shunt' makes available liberated carbon and elements, such as nitrogen, phosphorus, sulphur, and iron, for uptake by the microbial food web (Gobler *et al.,* 1997). This primary result of viral cell lysis has been demonstrated to stimulate production of non-infected bacteria, which may not otherwise grow due to a lack of suitable substrates and competition (Brussaard *et al.,* 1996a; Middelboe *et al.,* 1996; Bratbak *et al.,* 1998a). As much as 50% of bacteria mortality can be attributed to viral lysis, leading to the subsequent release of nutrients (Gobler *et al.,* 1997; Weinbauer, 2004). As a result, it is predicted that the overall level of bacterial production and respiration rate increases by as much as a third (compared to a model with no mortality attributed to viral infection). Cell lysis

therefore drives the food web towards a more regenerative system (Wilhelm and Suttle, 1999).

Viruses will hold a significant role in oligotrophic systems, where nutrients such as nitrogen are scarce (Gobler *et al.,* 1997). The cyanobacterial species *Trichodesmium* represent an important component of oligotrophic waters, as they are nitrogen-fixing organisms. Potential grazers of *Trichodesmium* are deterred due to its production of harmful toxins (Hawser *et al.,* 1992). Lysis, due to viral infection, may instead release the fixed nitrogen in the prokaryote, making it available for uptake by surrounding microorganisms (Hewson *et al.,* 2004). Viral lysis is also a significant pathway for phosphate recycling, particularly in phosphate-limited systems (Weinbauer, 2004). Viral lysis products turnover relatively rapidly, particularly in oligotrophic, P-limited environments (Noble and Fuhrman, 1999). The biogeochemical net effect of viral activity is therefore diversion of organic matter from transfer to higher trophic levels through grazing (classic food web) towards microbe-mediated recycling processes via viral lysis (Suttle, 2007). This enhanced activity of the microbial food web also results in the regeneration of potential algal growth-limiting nutrients, such as P and N, and affects competition between and succession of the remaining phytoplankton species (Gobler *et al.,* 1997).

Viruses can exert indirect influences on all levels of the food web by directly controlling the structure of microbial communities (Hennes *et al.,* 1995; Noble and Fuhrman, 1999; Middelboe and Lyck, 2002). Viral activity that maintains nutrients in surface waters, via the viral shunt, has the effect of increasing levels of biomass and productivity in the entire marine pelagic system. Viral lysis induces changes in the relative importance of functional groups in the marine food web (Weinbauer, 2004). It has been postulated

that viral lysis, by altering the overall production to community respiration ratio, could be a significant modulating factor of the total metabolism in the pelagic ecosystem (Eissler *et al.,* 2003). Debate continues regarding the directional pressures exerted within food webs (Verity *et al.,* 2002). It is commonly believed that production and biomass at each trophic level are determined either by the available nutrients ('bottom-up' control) or by the effect of predators ('top-down control').

A dynamic combination of both bottom-up and top-down controls has been proposed in modelling the structure of food webs (Verity and Smetacek, 1996). Both top-down and bottom-up controls were seen to be regulatory factors in shaping bacterial abundance and production (Gasol *et al.,* 2002; Tanaka and Rassoulzadegan, 2004). Bacterial growth rate is controlled by substrate availability, i.e. bottom-up control, whilst bacterial biomass is influenced by predation, i.e. top-down control (Thingstad *et al.,* 1997). Viruses can directly impact on species at all trophic levels. Fish, for example, are susceptible to infection by a range of viruses from the iridoviruses to the rhabdoviruses. Viruses could therefore be viewed as a side-in control.

Research has been conducted on the role of the microbial food web in conjunction with the classic food web (Mousseau *et al.,* 1998; Pitta *et al.,* 1998; Tzaras *et al.,* 1999; Rodriguez *et al.,* 2000; Shinada *et al.,* 2001). However, the picture, when viewed as a whole, is a dynamic and extremely complex one. To fully understand the spatial and temporal occurrence of different functional groups or species, the mechanisms structuring the pelagic marine food web need to be characterised. One of the greatest challenges to marine ecology is to determine how different food web processes, most of all viral activity, influence ecosystem structure and functioning (Pace and Cole, 1996). At present, no definitive all-encompassing studies exist to ascertain the outcomes of

viral infection at all levels of the marine food web. Advances in methods to assess both indirect and direct effects in natural systems are needed. Mathematical modelling incorporating viral influences throughout, and validated against actual data, is required to gain a clear insight into the processes involved.

1.2.4. Virus-mediated horizontal gene transfer

The role of viruses in mediating genetic exchange among microorganisms is known to occur (Jiang and Paul, 1998) but details of this phenomenon in the marine environment are uncertain and there are several gaps in our knowledge (Fuhrman, 1999). Transduction is the viral-mediated horizontal gene transfer from one host (the donor) to the next host (the recipient). Transduction occurs mainly in a restricted host range, although reports exist of transfer across a wide host range (Chiura, 1997). Transformation can also occur when viral lysis causes the host cell to release DNA, which is subsequently taken up and incorporated by another microorganism (Fuhrman, 1999). Both transduction and transformation potentially give rise to genetic variety in microbial populations. This has major implications for the population dynamics, genetic structure and evolution of marine populations (Weinbauer, 2004; Weinbauer and Rassoulzadegan, 2004).

Viruses can incorporate host DNA into their own genomes. A recent discovery has found a phage (S-PM2) that infects the cyanobacterium *Synechococcus* contains two genes which are usually found in the host (Mann *et al.,* 2003). The genes encode the D1 and D2 proteins that are key components of one of the photosynthetic reaction centres (photosystem II, PS II). These centres can become damaged in high light intensity leading to photo-inhibition. The expression of the genes encoding D1 and D2 proteins would allow the repair cycle to take place in PS II, after the host cell has shut it

down due to photo-inhibition. Photosynthesis can then continue, ensuring enough energy is available for the host cell to continue replication of the virus. It is thought the S-PM2 virus acquired the genes horizontally from its *Synechococcus* host. A similar survival strategy has been reported for a virus that infects the green alga *Chlorella* (Seaton *et al.,* 1996). The virus is known to enhance the mechanism used by the host to avoid photo-inhibition. Viral-encoded photosynthesis genes are now thought to be common in cyanophages (Mann *et al.,* 2003; Lindell *et al.,* 2004; Millard *et al.,* 2004; Sullivan *et al.,* 2006). Extensive sequencing of these core photosystem reaction-centre genes show that cyanophages themselves act as genetic reservoirs for their hosts, generating diversity even at the level of these globally distributed genes (Zeidner *et al.,* 2005; Sullivan *et al.,* 2006).

1.2.5. Use of '-Omic' Techniques

Sequencing of shotgun libraries from environmental microbial communities, or metagenomics, has emerged as a powerful tool that enables the discovery of entirely novel groups of microbes, regardless of their ability to be cultured in the laboratory (Rondon *et al.,* 2000). Metagenomic data from coastal waters and sediments has revealed incredibly high genetic richness in marine viral communities. In addition, these results suggest that the majority of environmental viruses are still uncharacterised, since 65-95 % of marine viral metagenomic sequences are not similar to previously described sequences (Breitbart *et al.,* 2004a; 2005; Angly *et al.,* 2006; Culley *et al.,* 2006). Furthermore, metagenomic analysis provides important insights into biogeographical distribution and community structure (i.e. the number of genotypes and relative abundances).

The use of such culture independent techniques is widening, as the so-called 'omic' approach is developed (Allen and Wilson, 2008). Techniques include microarrays, metagenomics, proteomics and the newly emerging tools of metaproteomics and metatranscriptomics. It is anticipated the data generated from such fields of study will yield more information on the function of genes and metabolic pathways and, therefore, enlighten what is known about viral infection strategies and dynamics.

1.3. Phytoplankton hosts of viral infection

Phytoplankton are tiny, free-floating plants found in the world's oceans. Through the harnessing of light energy, these microscopic organisms form the foundation of marine ecosystems and act as one of the engines of planetary control, exerting a significant influence on several global environmental factors. Photosynthetic activity by phytoplankton (microalgae) contributes to almost 50% of the annual global carbon-based primary productivity (Field *et al.,* 1998). Phytoplankton are therefore of crucial importance in the global carbon cycle and, also, climate (Charlson *et al.,* 1987), as well as playing additional central roles in regulating marine food webs and biogeochemical cycling (Falkowski *et al.,* 1998). Microalgal phytoplankton can be sub-divided according to size and 'type'. Photosynthetic picoeukaryotes (PPEs) are 0.2 µm – 2 µm in diameter, and nanoeukaryotic phytoplankton are in the size range 2 µm - 20 µm in diameter. These organisms are important members of the photosynthetic assemblage in the world's oceans. Their significant biomass and high productivity indicate they play a major role in illuminated pelagic marine habitats, including oligotrophic open-ocean and coastal environments (Blanchot and Rodier, 1996; Landry *et al.,* 1996; Diez *et al.,* 2001; Massana *et al.,* 2002; 2004; 2008). The taxonomic classification of these phytoplankton species is shown in Figure 1.5.

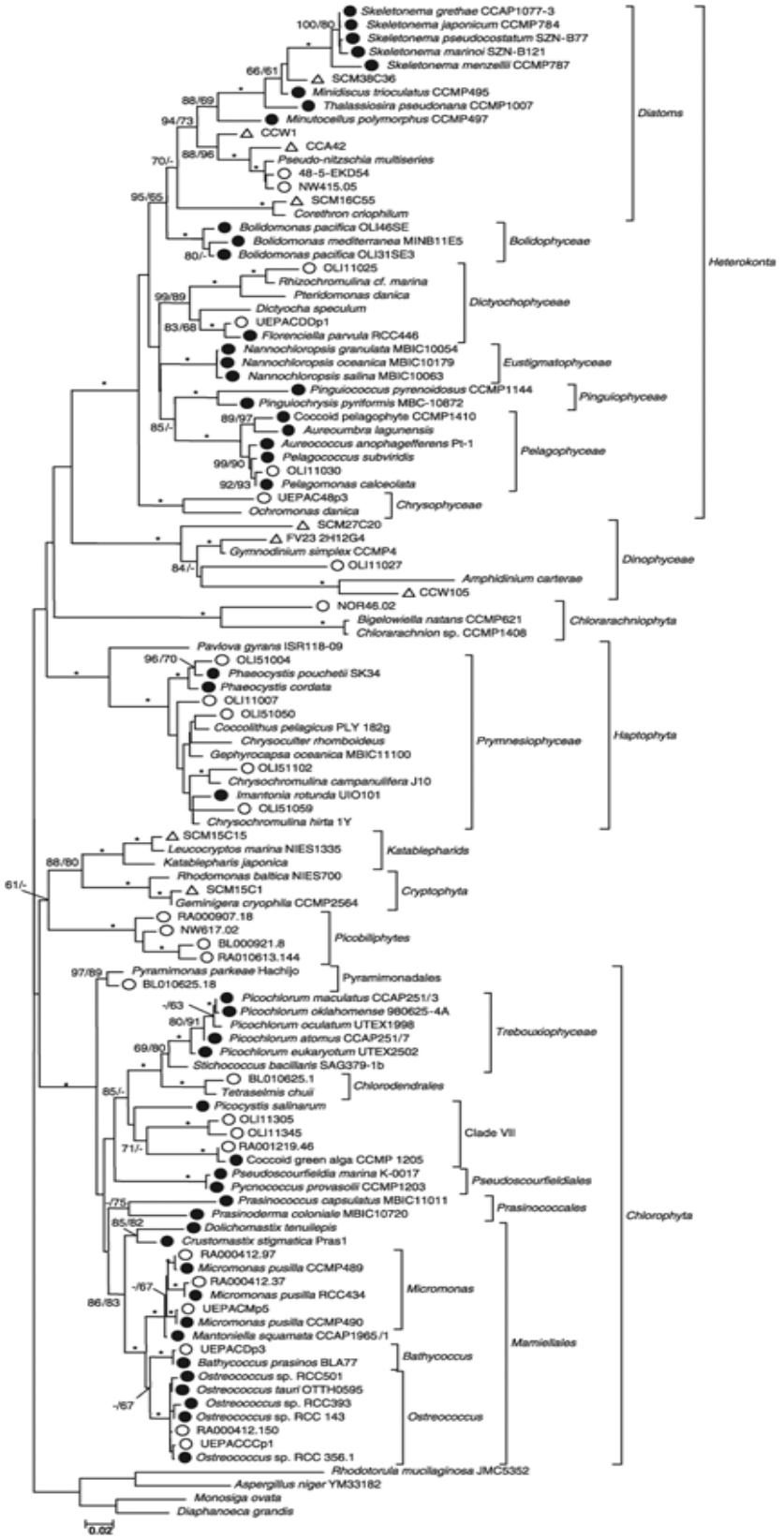


Figure 1.5. Phylogenetic tree, based on 18SrRNA gene sequences, representing described small phytoplankton species held in culture (closed circles) and environmental sequences (open circles). Numbers at nodes represent bootstrap values corresponding to neighbour joining and maximum parsimony respectively. The scale bar indicates 0.02 substitutions per site. Adapted from Vaulot *et al.* (2008).

The importance of eukaryotic picophytoplankton (PPEs) has been demonstrated in several studies. Their role as primary producers and their contribution to biomass (Tremblay *et al.,* 2009) and productivity in the euphotic zone of oceanic oligotrophic waters (Li, 1994) and also in coastal waters (Joint *et al.,* 1986; Not *et al.,* 2004) has been well-documented. The prokaryotic component of the picoplankton is represented by two genera, *Synechococcus* and *Prochlorococcus* (Partensky *et al.,* 1999; Scanlan, 2003), whilst the PPEs are much more diverse belonging to a variety of divisions (classes). To date, approximately 40 species belonging to nine classes (*Chlorophyceae, Prasinophyceae, Trebouxiophyceae, Prymnesiophyceae, Bacillariophyceae, Pinguiophyceae, Pelagophyceae, Bolidophyceae* and *Eustigmatophyceae*) have been formally described (Vaulot *et al.,* 2008) (Fig. 1.4). The molecular phylogenetic diversity among these organisms has been demonstrated through phylogenetic analysis of sequences, particularly 18S and 16S rRNA and the gene encoding D1 photosystem protein *(psbA)* genes, derived from natural samples (Massana *et al.,* 2002; Zhu *et al.,* 2005; Fuller *et al.,* 2006a). Initial reports of such broad diversity included studies conducted in the equatorial Pacific Ocean (Moon-van der Staay *et al.,* 2001), the Antarctic Polar Front (Lopez-Garcia *et al.,* 2001a), coastal waters of the Pacific Ocean (Worden, 2006), the Mediterranean, Wedell, and Scotia Seas as well as from the North Atlantic Ocean (Diez *et al.,* 2001) and the Arabian Sea (Fuller *et al.,* 2006a).

The importance of prasinophytes has been highlighted as one of the key picoplankton groups particularly in European marine coastal waters (Zingone *et al.,* 1999; Guillou *et al.,* 2004; Massana *et al.,* 2004; Not *et al.,* 2004; Romari and Vaulot, 2004). The prasinophytes are composed of at least seven clades, including the *Mamiellales* and *Prasinococcales* (Guillou *et al.,* 2004). In coastal waters, the order *Mamiellales* and in particular members of the class *Prasinophyceae*, most notably the three species

Micromonas pusilla (Butcher, 1952), *Bathycoccus prasinos* (Eikrem and Throndsen, 1990) and *Ostreococcus tauri* (Chretiennotdinet *et al.,* 1995), have been revealed as of great importance. The *Prasinophyceae* is considered as the most primitive in the green lineage and gave rise to all other green classes including land plants (Sym and Pienaar, 1993; Lewis and McCourt, 2004). Sequencing of the genomes of the smallest free-living eukaryote *O. tauri* (Derelle *et al.,* 2006), *O. lucimarinus* (Palenik *et al.,* 2007), and two genomes of *Micromonas* species (Worden *et al.,* 2009) are allowing important insights into the ecological adaptations of these important organisms and are helping to reveal clues as to the features of ancestral algae and their subsequent evolution.

1.4. Algal virology

1.4.1. Historical aspects of algal virology

The first cyanophage that infected a filamentous freshwater cyanobacterium had been isolated in 1963 (Safferman and Morris, 1963) and for the next 25 years most research focussed on basic cyanophage biology (Sherman and Brown, 1978; Martin and Benson, 1988). Awareness of the presence of virus-like particles (VLPs) in species of eukaryotic algae, based predominantly on microscopic observations, can be traced back to reports in the early 1970s. The earliest descriptions were mainly based on incidental observations of VLPs that were apparent in electron micrographs, e.g. Manton & Leadbeater (1974) described VLPs in *Chrysochromulina mantoniae*. One of the first marine algal viruses to be isolated and characterised was a virus that infects the marine unicellular alga *Micromonas pusilla* (Mayer and Taylor, 1979). However, no major investigations began in this area until the early 1990s, when high concentrations of viruses were reported in aquatic environments (Bergh *et al.,* 1989). During the 1980s, the most significant discovery was a group of viruses reported to infect freshwater, eukaryotic, symbiotic *Chlorella*-like green algae, thus named the chloroviruses (Meints

et al., 1981; Van Etten *et al.,* 1982; 1985). Much of the eukaryotic algal virus research conducted in the 1980s and 1990s focussed on the chloroviruses (Van Etten *et al.,* 1981; Van Etten *et al.,* 1982) leading to the establishment of the *Phycodnaviridae* (Van Etten and Ghabrial, 1991). Initial studies led to the descriptions of viruses infecting ca. 20 species of marine phytoplankton (Van Etten *et al.,* 1991) and the list continues to increase, now with a number of dinoflagellates (Tarutani *et al.,* 2001; Onji *et al.,* 2003; Tomaru *et al.,* 2004) and diatoms (Nagasaki *et al.,* 2004; Bettarel *et al.,* 2005) included.

1.4.2. DNA viruses

1.4.2.1. Nuclear-Cytoplasmic Large dsDNA Viruses (NCLDVs)

With the advent of the genomic era, enabling the generation of extensive genome sequence data, establishment of the evolutionary relationships within and between families of DNA viruses has begun (Shackleton and Holmes, 2004). Studies of whole genome comparisons have enabled the construction of evolutionary relationships between virus families and such studies have identified a group of large dsDNA viruses that are likely to have shared a common ancestor (Iyer *et al.,* 2001). The nuclear-cytoplasmic large dsDNA viruses (NCLDVs) groups is composed of at least five families that replicate in the nucleus and/or cytoplasm of eukaryotic cells: *Poxviridae; Iridoviridae; Asfarviridae; Phycodnaviridae;* and *Mimiviridae* (Fig. 1.6) (Iyer *et al.,* 2001; 2006).

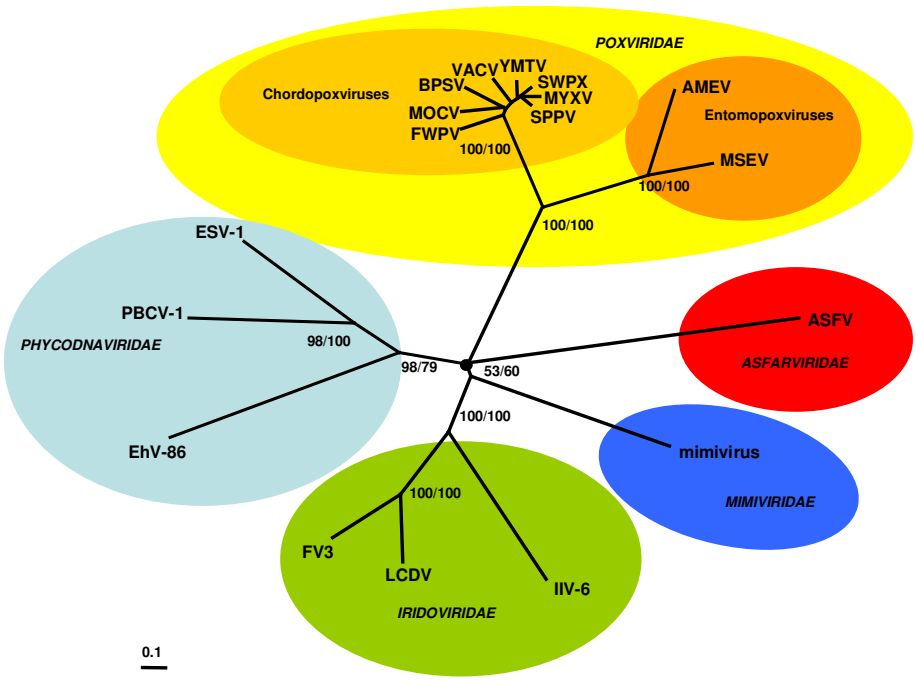


Figure 1.6. - Phylogenetic inference tree based on a distance matrix algorithm between the concatenated conserved domains from A18-like helicase, D6R-like helicase, A32-like ATPase, D5-like ATPase, DNA polymerase, thiol-oxidoreductase, and the two largest RNA polymerase subunits from members of the NCLDV group. Numbers at nodes indicate bootstrap values retrieved from 100 replicates for both the Neighbour-joining and parsimony analyses. The bar depicts 0.1 fixed mutations per 10 amino acids. Redrawn from Allen *et al.* 2006b.

Members of the NCLDVs share the following main factors in common; they all have large dsDNA genomes; secondly, as their name suggests, replication of these viruses most likely occurs in the nucleus and/or cytoplasm of the host cell; and finally, they share an ancient evolutionary heritage, with a common ancestor, which probably had a nuclear/cytoplasmic life cycle (Iyer *et al.,* 2001; 2006). Comparative analysis show that the NCLDVs share a core set of conserved genes, which implies these diverse viruses originated from a common ancestor (Iyer *et al.,* 2001). Evolutionary reconstructions,

based on the parsimony principle, assigned 31 genes to this hypothetical ancestral virus, whose products are largely involved in viral genome replication and transcription, as well as virion production. Nine core genes are shared by genomes from all family members (Group I) and a further 22 are found in at least three of the four families (Groups II and III) (Iyer *et al.,* 2001). The common NCLDV ancestor is believed to have had both nuclear and cytoplasmic phases of its life cycle (Iyer *et al.,* 2001). The loss and gain of lineage-specific genes within the NCLDV families is thought to have played a part in the diverse characteristics of the present virus members.

Poxviruses, asfarviruses, and iridoviruses encode their own transcription and replication machinery. Poxviruses replicate solely in the cytoplasm whilst asfarviruses and iridoviruses start replication in the nucleus and complete it in the cytoplasm. The exact location(s) of replication for the highly diverse members of the *Phycodnaviridae* family is less clear. The evolutionary history of the NCLDVs has become more complex with the recent sequencing of the Mimivirus (Raoult *et al.,* 2004), which has been placed on a branch diverging prior to the divergence of the *Phycodnaviridae.*

1.4.2.2. Phycodnaviruses

Most viruses that have been identified as infecting marine phytoplankton have been assigned taxonomically to the highly diverse family, the *Phycodnaviridae.* The *Phycodnaviridae* family is comprised of large double stranded (ds) DNA viruses infecting eukaryotic algae, both freshwater and marine (Van Etten *et al.,* 2002; Wilson *et al.,* 2005a). Members of the family are morphologically similar, exhibiting icosahedral symmetry, and include some of the largest known viruses, with genome sizes in the range of 156 to 560 kb and particle sizes of 100 – 220 nm (Wilson *et al.,* 2005a; Dunigan *et al.,* 2006). The *Phycodnaviridae* are currently comprised of the

Chloroviruses, Coccolithoviruses, Prasinoviruses, Prymnesioviruses, Phaeoviruses and *Raphidoviruses* (Wilson *et al.,* 2005a) (Table 1.1). These viruses are genetically diverse (Dunigan *et al.,* 2006), although they have been classified within the NCLDVs (Iyer *et al.,* 2001; Raoult *et al.,* 2004; Allen *et al.,* 2006b). Details of the characteristics of the phycodnaviruses isolated to date are summarised in Table 1.1.

Table 1.1. Characteristics of the eukaryotic algal viruses isolated and characterised to date. Adapted from Brussaard and Martinez (2008).

Host algae	Virus code	Nucleic acid type	Particle size (nm)	Genome size (kbp)	Location of replication	Burst size	Lysogenic/lytic Latent period, h	References
***Bacillariophyceae* (diatoms)**								
Rhizosolenia setigera	RsRNAV	ssRNA	32	11.2	Cytoplasm	1000-3000	Lytic, <24h	Nagasaki *et al.,* 2004
Chaetoceros cf. gracilis	CspNIV	ND	25	ND	Nucleus	ND	Lytic, <24h	Bettarel *et al.,* 2005
Chaetoceros salsugineume	CsNIV	ssDNA	38	6	Nucleus	325	Lytic, 12-24h	Nagasaki *et al.,* 2005b
Chaetoceros debilis	CdebDNAV	SsDNA	30	ND	Cytoplasm	55	Lytic, 12-24h	Tomaru *et al.,* 2008
***Chlorophyceae* (green algae)**								
Chlorella-like green alga	PBCV	dsDNA	185	330-380	Cytoplasm	200-350	Lytic, 6-8 h	Meints *et al.,* 1981, Van Etten *et al.,* 1982
***Chrysophyceae* (golden algae)**								
Aureococcus anophagefferens	AaV	dsDNA	150	ND	Cytoplasm	>500	Lytic, ND	Milligan and Cosper 1994 Gastrich *et al.,* 1998
***Dinophyceae* (dinoflagellates)**								
Heterocapsa circularisquama	HcV	dsDNA	200	350	Cytoplasm	>1,300	Lytic, 24 h	Tarutani *et al.,* 2001
Heterocapsa circularisquama	HcRNAV	ssRNA	30	4.4	Cytoplasm	3.4×10^3	ND	Tomaru *et al.,* 2004
Prasinophyceae								
Micromonas pusilla	MpV	dsDNA	115	190-210	Cytoplasm	72	Lytic, 7-14h	Waters and Chan 1982
Micromonas pusilla	MpV	dsDNA	115	190-220	Cytoplasm	-	Lytic, 4-8h, 8-12h, 12-16h	J Martinez Martinez unpubl. data
Micromonas pusilla	MpRV	dsRNA	90	25.6	Cytoplasm	ND	Lytic, 36	Brussaard *et al.,* 2004
Pyramimonas orientalis	PoV	dsDNA	200	560	Cytoplasm	800-1000	Lytic, 14-19h	Sandaa *et al.,* 2001

Table 1.1 contd.

Host algae	Virus code	Nucleic acid type	Particle size (nm)	Genome size (kbp)	Location of replication	Burst size	Lysogenic/lytic Latent period, h	References
***Phaeophyceae* (brown algae)**								
Ectocarpus siliculosus	EsV	dsDNA	140	335	Nucleus	2×10^6	lysogenic	Muller *et al.,* 1990
Feldmannia simplex	FsV	dsDNA	120-150	220	Nucleus	$1\text{-}5 \times 10^6$	lysogenic	Muller and Stache 1992
Feldmannia sp.	FsV	dsDNA	150	170	Nucleus	ND	lysogenic	Henry and Meints 1992
Hincksia hincksiae	HincV	dsDNA	140-170	ND	Nucleus	1.2×10^6	lysogenic	Parodi and Muller 1994
Piyaella littoralis	PlitV	dsDNA	130-170	280	Nucleus	2.5×10^5	lysogenic	Maier *et al.,* 1998
Myriotrichia clavaeformis	MclaV	dsDNA	170-180	ND	Nucleus	ND	lysogenic	Wolf *et al.,* 2000
Prymnesiophyceae								
Emiliania huxleyi	EhV	dsDNA	170	415	Cytoplasm	400-1000	Lytic, 12-14	Castberg *et al.,* 2002
Phaeocystis globosa	PgV-Group I	dsDNA	150	466	Cytoplasm	ND	Lytic, 12 or 16	Baudoux and Brussaard, 2005
Phaeocystis globosa	PgV-Group II	dsDNA	100	177	Cytoplasm	ND	Lytic, 10	Baudoux and Brussaard, 2005
Phaeocystis globosa	PgV	dsDNA	98	176	ND	ND	Lytic, ND	Wilson *et al.,* 2006
Phaeocystic pouchetti	PpV	dsDNA	130-160	485	Cytoplasm	350-600	Lytic, 12-18	Jacobsen *et al.,* 1996
Chrysochromulina brevifilum	CbV	dsDNA	145-170	ND	Cytoplasm	>320	Lytic, ND	Suttle and Chan 1995
Chrysochromulina ericina	CeV	dsDNA	160	510	Cytoplasm	1800-4100	Lytic, 14-19	Sandaa *et al.,* 2001
Raphidophyceae								
Heterosigma akashiwo	HaV	dsDNA	200	ND	Cytoplasm	770	Lytic, ND	Nagasaki *et al.,*1997
Heterosigma akashiwo	OiS1	dsDNA	ND	180	Cytoplasm	ND	Lytic, 17	Lawrence *et al.,* 2006
Heterosigma akashiwo	HaRNAV	ssRNA	25	0.91	Cytoplasm	ND	Lytic, 35	Tai *et al.,* 2003
Heterosigma akashiwo	HaNIV	ND	30	ND	Nucleus	1×10^5	Lytic, ND	Lawrence *et al.,* 2001

ND = Not Determined

1.4.2.2.1. Structure and morphology of phycodnaviruses

All members of the *Phycodnaviridae* have similar structural morphology and this could be indicative of a shared common ancestry. Although only a small number have been studied in great depth, the emerging characteristics suggest a broadly conserved structure of virions, consistent with the need for a simple, easily assembled virus particle (Bamford *et al.,* 2002). General features of the virion structure are they are large layered structures between 100 - 220 nm in diameter with a dsDNA-protein core surrounded by a lipid bilayer and an icosahedral capsid. Most of the detailed picture of the structure of these particles has arisen from extensive studies, involving cryo-electron microscopy and 3-D image reconstruction of the chlorovirus, PBCV-1 (Yan *et al.,* 2000; Simpson *et al.,* 2003). The largest capsid of members of the *Phycodnaviridae* is possessed by *Phaeocystis pouchetti* virus, PpV01 (Yan *et al.,* 2005). The nature of the genetic material and the morphology of members of the *Phycodnaviridae* are the two common factors, which group this highly diverse family of viruses together.

1.4.2.2.2. Genome structure and features

Genome structure of the currently available sequenced phycodnaviruses varies greatly. The chlorovirus, PBCV-1 is a linear, 330 kb dsDNA structure with covalently closed hairpin ends (Rohozinski *et al.,* 1989). The phaeovirus, EsV-1, has a linear dsDNA genome with near perfect inverted repeats at each end allowing for circularisation of the genome (Delaroque *et al.,* 2001). Interestingly, studies conducted before EsV-1 was sequenced concluded that this particular virus had a circular genome (Lanka *et al.,* 1993). The inverted repeats are proposed to anneal with each other to produce a cruciform structure that in effect circularises the genome (Delaroque *et al.,* 2001). The coccolithovirus, EhV-86, potentially has both

linear and circular phases, as the termini possess a random A/T single nucleotide overhang (Wilson *et al.,* 2005b; Allen *et al.,* 2006c).

Virus genomes are typically compact to maximise replication efficiency and phycodnaviruses are no exception, with approximately one gene per 900 to 1000 bp of genomic sequence (based on predicted coding sequences (CDS) of the chloroviruses and EhV-86) (Li *et al.,* 1997; Wilson *et al.,* 2005a; Fitzgerald *et al.,* 2007a, b, c). However, the 336 kb EsV-1 genome has only 231 protein encoding genes, which represents one gene every 1,450 bp. Some repetitive DNA does exist in the genomes characterised to date. Both EsV-1 and PBCV-1 contain approximately 2 kb inverted repeats adjacent to the terminal ends (Strasser *et al.,* 1991; Delaroque *et al.,* 2001). Tandem repeats are also located throughout the EsV-1 genome and comprise about 12% of the entire genome (Delaroque *et al.,* 2001). A similar percentage of the *Feldmannia sp*. Virus (FsV) genome also consists of repetitive DNA (Lee *et al.,* 1995). Such repetitive regions in the genomes of phycodnaviruses may be involved in recombination between viruses that facilitates the exchange of genetic information between themselves and their hosts (Allen *et al.,* 2006a, b).

1.4.2.2.3. Propagation strategies

Phycodnaviruses display a variety of propagation strategies. Some phycodnaviruses are host-specific and only infect single isolates or species of algae. Chloroviruses, for example, only attach to cell walls of certain unicellular, eukaryotic, *Chlorella*-like green algae e.g. PBCV-1 and NY-2A (Meints *et al.,* 1984; Reisser and Kapaun, 1991; Van Etten, 2003). Virus attachment is probably at a unique vertex (Onimatsu *et al.,* 2006) and is immediately followed

by dissolution of the host wall at the point of adsorption. Host cell wall degradation is followed by entry of the viral DNA and associated proteins (glycoproteins and several phosphoproteins) into the cell, with the empty capsid remaining on the host cell surface. A rapid depolarization of the host membrane is initiated by this process (postulated to be via a virus encoded potassium channel sited in the virus internal membrane) (Frohns *et al.,* 2006). Depolarization may ensure infection by further viruses is prevented (Mehmel *et al.,* 2003; Frohns *et al.,* 2006). Host chromosomal DNA begins to degrade during the immediate-early phase of infection (Agarkova *et al.,* 2006). Viral DNA replication begins 60 to 90 minutes after infection (Van Etten *et al.,* 1984) and transcription of late genes follows (Schuster *et al.,* 1986). Two to four hours post-infection, progeny virions are assembled in the cytoplasm of the host. Infectious virions can be detected inside the cell approximately 30-40 minutes prior to virus release, which occurs through cell lysis. An estimated 1000 particles are released per host cell, of which approximately 30% are infectious.

Infection of the prasinoviruses occurs when virions adhere to the wall-less host cell surface, followed by fusion of adjacent host and particle surfaces. Empty particles remain on the cell surface following the release of core contents. An eclipse period of approximately 3 hours follows the attachment stage (Waters and Chan, 1982). The virus growth cycle is complete after approximately 14 hours. During the replication cycle, particles appear in the cytoplasm and are associated with the production of cytoplasmic fibrils (~ 5 - 8 nm in diameter) and clusters of membrane-bound vesicles that are absent in healthy cells. Particles are released into the medium via localized ruptures (which often appear at several locations on the same cell) in the cell membrane.

Coccolithoviruses utilise a different propagation strategy to those discussed thus far and it has been proposed to be reflective of a lifestyle possibly exhibited by an ancestral phycodnavirus (Allen *et al.,* 2006b). Despite the presence of a calcium carbonate outer shell on the algal host, *Emiliania huxleyi,* virus adsorption is rapid and does not appear to be thwarted by the external lith structures. Extensively studied phycodnaviruses, such as PBCV-1 and EsV-1, are completely dependent on the host transcriptional machinery for their propagation. However, coccolithoviruses are unique within this family of viruses, in that they possess several genes encoding RNA polymerase subunits (Wilson *et al.,* 2005b). The presence of these genes suggests a possible partial, or whole, independence from the host nucleus. Transcription of a distinct localised 100 kb segment, containing a unique promoter region, of the virus genome begins immediately after initial infection (Allen *et al.,* 2006d). Genes associated with this unique promoter element are transcribed in the first hour after infection but they are classed as ORFans (Fischer and Eisenberg, 1999), i.e. no match exists for these genes in the databases.

A second transcriptional phase begins 1 – 2 hours post-infection when the remainder of the genome is expressed. Viral RNA polymerase components are expressed in this second phase and as replication may no longer be nuclear dependent, transcription may then move to the cytoplasm. In contrast to the other phycodnavirus release strategies, where virus progeny are seen to accumulate in the cytoplasm, coccolithovirus virions are released gradually, in a controlled manner (Mackinder *et al.,* in press). Possible explanations for this include the involvement of the host lith export pathway to enable a steady release of viral progeny. Another hypothesis is that the coccolithoviruses exploit the virally encoded sphingolipid metabolic pathway

to influence the sphingolipids in the host cell wall or membrane to delay apoptosis and prolong the length of viral infection (Han *et al.,* 2006).

Less is known about the replication of prymnesioviruses and raphidoviruses. Virus formation is observed in the cytoplasm and the nucleus remains intact and separate from the viroplasm that consists of a fibrillar matrix. Ultimately, viral production results in the disruption of organelles, lysis of the cell and release of the virus particles (Nagasaki *et al.*, 2004; Brussaard *et al.*, 2007).

There are often difficulties in cultivation of susceptible hosts in the laboratory, which hinders research into marine viral diversity. An additional drawback is the lack of a single genetic element shared by all viruses (Rohwer and Edwards, 2002). However, viruses from certain taxonomic groups share conserved genes that can be used to study diversity within those groups. For example, some studies of the diversity of cyanophages have been based on sequences of structural proteins (Fuller *et al.,* 1998; Zhong *et al.,* 2002; Mühling *et al.,* 2005; Short and Suttle, 2005) and diversity of algal viruses has been mainly studied by sequencing DNA *pol* genes (Chen and Suttle, 1995a, b). A few other genetic markers are available, such as those based on the major capsid protein (MCP) gene of viruses that infect *E. huxleyi* (Schroeder *et al.,* 2002).

1.4.2.2.4. Chloroviruses

Unlike most members of the *Phycodnaviridae,* which infect marine algae, viruses of the *Chlorovirus* genus infect freshwater algae. The isolation and characterisation of viruses that infect a *Chlorella*-like species led to the formation of the *Phycodnaviridae* family (Meints *et al.,* 1986; Van Etten, 1995; Van Etten and

Meints, 1999). Chloroviruses are large, plaque-forming dsDNA viruses that infect particular *Chlorella*-like unicellular green algae (Van Etten, 1995). Members of the genus *Chlorella* are small, unicellular, non-motile, asexual green algae with a universal distribution. Although *Chlorella* species are largely free-living, many species do have symbiotic relationships with species, representing a number of different classes from the animal kingdom, e.g. *Rhizopodia, Ciliata* and *Hydrozoa* (Reisser, 1993). The chloroviruses described to date infect endosymbiont *Chlorella* hosts, also known as zoochlorellae. Zoochlorellae are associated with protists, such as *Paramecium bursaria* and the coelenterate *Hydra viridis* and the heliozoon *Acanthocystis turfacea* (Kawakami and Kawakami, 1978; Meints *et al.,* 1981; Van Etten *et al.,* 1983a; Bubeck and Pfitzner, 2005). When the zoochlorellae are in the paramecium, they are resistant to infection from viruses. This is due to protection afforded to them by enclosure within individual host derived vacuoles, which results in exclusion of the viruses. It is possible however, to grow many of the zoochlorellae strains, such as the paramecium associated *Chlorella* isolates NC64A and Pbi, in the laboratory, independent from their symbiotic hosts.

Chloroviruses are ubiquitous and have been isolated from freshwater locations worldwide (Yamada *et al.,* 2006). Virus titers in native waters are typically in the range of 1 to 100 plaque-forming units (PFU) per ml, although titers as high as 100,000 PFU per ml of native water have been reported. Titers are seen to fluctuate seasonally with the greatest yields occurring in the spring. The most extensively studied chlorovirus, which was also the first phycodnavirus to be completely sequenced, is PBCV-1, which infects *Chlorella* NC64A (Li *et al.,* 1997; Van Etten, 2003; Yamada *et al.,* 2006).

1.4.2.2.5. Coccolithoviruses

Coccolithophores are unicellular calcifying marine algae with a global distribution (Brown and Yoder, 1994). The coccolithophorid *Emiliania huxleyi* (Haptophyta) is an important representative of this group of algae, both historically and at present, in relation to sediment formation, primary productivity and climate change. Vast blooms of *E. huxleyi* exert an enormous influence on global carbon and sulphur cycles, and they have been shown to affect the oceanic carbon pump and influence climate (Malin *et al.,* 1998). These great blooms of *E. huxleyi* are often seen to decline swiftly, with subsequent deposition of calcite to the seabed and release of the cloud-forming dimethyl sulphide (DMS) to the atmosphere (Malin and Kirst, 1997). The cause of the demise of blooms of *E. huxleyi* is now widely accepted to be attributed to a large degree to the action of viruses (Bratbak *et al.,* 1993; 1996; Brussaard *et al.,* 1996b; Castberg, 2002; Jacquet *et al.,* 2002; Schroeder *et al.,* 2002; Wilson *et al.,* 2002b).

Schroeder *et al.,* (2002) found that viruses isolated from *E. huxleyi,* based on phylogenetic analysis of the DNA polymerase gene, belong to a new genus within the *Phycodnaviridae* family. The name of the new genus was proposed to be *Coccolithovirus*. Differences between members of the genus were determined by host range analysis and sequence analysis of a gene fragment encoding part of the putative Major Capsid Protein (MCP). A coccolithovirus, EhV-86, that infects *E. huxleyi* has had its complete genome sequenced (Wilson *et al.,* 2005a; b). The virus, EhV-86 is a large virus about 170 to 175 nm in diameter with a dsDNA genome of length 407,339 bp. There are 25 core genes common to the NCLDVs present in EhV-86. This coccolithovirus encodes six RNA polymerase subunits, a unique feature

among members of the *Phycodnaviridae.* There is a 100 kb region of the EhV-86 genome that appears to have been incorporated from an unknown source. Such factors point to a unique evolutionary heritage and propagation strategy of this virus.

A second coccolithovirus, EhV-163, has had approximately 80% of its genome sequenced, equating to more than 200 full length CDSs (Allen *et al.,* 2006c). Both EhV-86 and EhV-163 have nearly identical sized genomes (Allen *et al.,* 2007). Comparative analysis of the 2 genomes revealed only 20 of the 202 fully sequenced CDSs in both genomes are completely identical at the nucleotide level and 17 at the protein level (Allen *et al.,* 2006c). These conserved CDSs are located across the genome although a selection tends to cluster in 4 defined genomic regions. The remaining EhV-163 CDSs have 95% to 99% identity with EhV-86 homologs. A genomic rearrangement is seen to occur in EhV-163. A putative phosphate permease, *ehv117* that is found in EhV-86 has been partially deleted and instead a putative endonuclease (*ehv117A*) is encoded in EhV-163 (Allen *et al.,* 2006c). Theoretically, this replacement still enables an increase in phosphate availability to the virus but via different mechanisms, i.e. one by increasing the uptake of phosphate and the other through the degradation of host nucleic acid.

A coccolithovirus microarray has been used to aid the annotation of the EhV-86 genome, study the infection process at the transcriptional level and assess the genomic diversity within the *Coccolithoviridae* (Allen and Wilson, 2006). Microarray analysis has offered an alternative to phylogeny-derived pattern for virus evolution, circumventing the need for sequencing. The genomic content of 12 coccolithoviruses, archived in the Plymouth Virus Collection (PVC), was screened

and revealed large variations between isolates (Allen and Wilson, 2006; Allen *et al.,* 2007). Of the 425 CDSs screened, 71 were either absent from certain viral strains or were highly variable in at least 1 virus isolate. All core NCLDV genes identified in EhV-86 were found in all PVC isolates analysed. Although genomic content varies greatly, the overall genome size is similar between coccolithoviruses (Allen *et al.,* 2007).

1.4.2.2.6. Phaeoviruses

Filamentous marine brown macroalgae assigned to the taxonomic order *Ectocarpales* (class *Phaeophyceae*) occur as members of benthic communities in near-shore coastal environments of all the world's oceans (Van den Hoek *et al.,* 1995; Muller *et al.,* 1998). The *Ectocarpales* contribute significantly to biofouling and frequently grow as epiphytes in mariculture (Van den Hoek *et al.,* 1995). Recently, the *Ectocarpales* species, *Ectocarpus siliculosus,* has been proposed as a model for studies on brown algal genetics and genomics (Peters *et al.,* 2004).

The *Ectocarpales* order is comprised of five families, based on molecular phylogenetic analysis of nuclear and ribosomal and RUBISCO sequence data (Peters and Ramirez, 2001). Viral infection has been reported in three of the five families, namely the *Ectocarpaceae*, the *Chordariaceae* and the *Acinetosporaceae* (Van Etten and Meints, 1999). To date, a total of eight viruses infecting the *Ectocarpales* have been isolated and described (Muller *et al.,* 1998; Van Etten and Meints, 1999). These eight viruses are assigned to the genus *Phaeovirus* within the *Phycodnaviridae* family (Van Etten and Meints, 1999). However, of the eight phaeoviruses, only three, namely *Ectocarpus siliculosus* virus (EsV-1), *Feldmannia* species virus (FsV-1) and

Feldmannia irregularis virus (FirrV-1), have been characterised and described in detail (Van Etten *et al.,* 2002).

Members of the genus *Phaeovirus* infect marine brown algal species and share icosahedral morphologies with internal lipid membranes and large, complex, dsDNA genomes (Muller *et al.,* 1998). All presently described phaeoviruses infect free-swimming, wall-less gametes or spores (Muller *et al.,* 1998) (Fig. 1.2). To date, infections of microalgae by phycodnaviruses have been described as lytic infections. However, macroalgae infected by phaeoviruses undergo a lysogenic infection. Viral DNA becomes integrated into the host genome and is subsequently transmitted, via mitosis, through all cell generations of the developing host (Muller *et al.,* 1998). The viral genome persists as a lysogenic infection in vegetative cells and infected algae show no apparent growth or developmental defects, except for partial or total inhibition of reproduction (Muller *et al.,* 1998). The viral genome is only expressed in cells of the reproductive organs, sporangia and gametangia (collectively termed zoidangia) (Fig. 1.2). Within these structures, the viral genome is replicated to the extent the cellular organs disintegrate and are replaced with densely packed viral particles (Muller *et al.,* 1998).

Environmental stimuli, such as temperature and light, which cause lysis of infected zoidangia also induce the release of functional zooids, thus producing synchronous release of virus particles and their potential host cells (Muller *et al.,* 1998). Infected ectocarpoid hosts may revert to the production of functional zoidangia without virus particles. The viral genome is still present in the plant and will be passed on to a proportion of the progeny, which will display symptoms when mature. Vertical

transmission of the virus can thus occur by the release of infected zooids (Muller *et al.,* 1998).

Previous molecular studies, using PCR, have indicated that in natural populations of *Ectocarpus*, viral infection may be in excess of 50% (Sengco *et al.,* 1996; Muller *et al.,* 2000). PCR employed in such studies has been designed to amplify regions of the gene encoding for the major capsid glycoprotein (Brautigam *et al.,* 1995). A lack of genetic sequence information for virus strains infecting a broad range of ectocarpoid hosts of different genera has made it impossible to estimate viral infection in natural populations of the *Ectocarpales*. Following the completion of the sequencing of the EsV-1 genome (Delaroque *et al.,* 2001), the FirrV–1 genome (Delaroque *et al.,* 2003), and the FsV–1 genome (Schroeder *et al.,* 2009), it is now feasible to design novel primers for phaeoviral genes, which may yield more information.

1.4.2.2.7. Prasinoviruses

Photosynthetic picoeukaryotes (PPEs), are small (less than 3 μm in size) photosynthetic eukaryotic cells. They are ubiquitous and highly abundant members of the global ocean phytoplankton assemblage and this, coupled to their ability to fix large quantities of carbon (Li, 1994; Worden *et al.,* 2004), make the PPEs important contributors to biomass production in the global oceans. A high diversity of PPEs is now well recognised (Diez *et al.,* 2001; Lopez-Garcia *et al.,* 2001b; Moon-van der Staay *et al.,* 2001; Stoeck and Epstein, 2003; Romari and Vaulot, 2004; Fuller *et al.,* 2006b; b; Worden, 2006) but from the many classes which comprise the PPEs, the class *Prasinophyceae* often dominates this group of microalgae. Representatives of the *Prasinophyceae* include the genus *Ostreococcus*, to date the smallest known free-

living eukaryote (Chretiennotdinet *et al.,* 1995) and *Micromonas*, a dominant member particularly in coastal regions (Not *et al.,* 2004).

One report exists of a transient bloom of *Ostreococcus* off the eastern coast of the USA, in which the presence of VLPs was noted (O'Kelly *et al.,* 2003). However, the VLPs reported in that study were not investigated further. More recently, as this thesis was being prepared for submission, another group of researchers reported the complete genome of a virus, OtV5 that infects *Ostreococcus tauri* (Derelle *et al.,* 2008).

Despite their high productivity and significant abundance, very little research has been conducted to characterise the viruses that infect these important microalgae. This is in spite of the fact that *Micromonas pusilla* was the first phytoplankton species that Mayer and Taylor (1979) isolated and characterised a virus for (which became the first virus to be assigned to the family *Phycodnaviridae*.). *M. pusilla* is a naked highly motile unicellular flagellate of very small size (1 -3 µm in diameter). It is widely distributed in both coastal and oceanic marine waters, and relatively abundant (Cottrell and Suttle, 1991; Sahlsten, 1998; Not *et al.,* 2005). Viruses isolated that infect *M. pusilla* are on the whole characterised as polyhedral, large (100 - 135 nm in diameter) with dsDNA genomes. In addition, a dsRNA virus has been isolated and described that also infects *M. pusilla* (Brussaard *et al.,* 2004) and has been assigned to the *Reoviridae*. Subsequent research has focussed mainly on the ecological aspect of these viruses and their genetic diversity (Cottrell and Suttle, 1991; 1995; Sahlsten, 1998; Sahlsten and Karlson, 1998; Zingone *et al.,* 1999; 2006).

The prasinophyte *Pyramimonas orientalis* has a worldwide distribution and most commonly occurs in low numbers, with no formation of blooms ever reported (Daugbjerg and Moestrup, 1993). A virus, PoV-01B, was isolated and described that infects *P. orientalis* (Sandaa *et al.,* 2001). The virus has a particle size of ca. 220 nm and a genome size of 560 kb, which is among the largest virus sizes ever reported. However, the virus shows variable properties in culture (e.g. variable lysis and burst size) and infectivity is lost easily (Larsen *et al.,* 2008).

Despite their host's recognised importance in the global oceans, no other reports of prasinoviruses exist at present. The viruses that infect prasinophytes are the most poorly described and characterised of the six *Phycodnaviridae* genera.

1.4.2.2.8. Prymnesioviruses

Prymnesiophytes are distributed globally and are often associated with large bloom formations. The *Prymnesiophyceae* is a member of the division Haptophyta and members of which are known as haptophytes. *Phaeocystis* is a member of the *Prymnesiophyceae* with a global distribution, which exists in unicellular form or as colonies, the latter often leading to the formation of dense blooms, mainly in temperate, coastal waters of Northern Europe and polar seas (Lancelot *et al.,* 1987). Their success has been attributed to their ability to form gel-like colonies in a matrix form (Schoemann *et al.,* 2005). Among the six species of *Phaeocycstis*, three have been reported as blooming species and of these, two species, *P. globosa* and *P. pouchetti*, have had strain-specific viruses that infect them isolated and described (Jacobsen *et al.,* 1996; Baudoux and Brussaard, 2005).

Phaeocystis plays a key role in biogeochemical cycling, particularly in carbon and sulphur cycling, which in turn strongly influences global climate (Smith *et al.,* 1991; Liss *et al.,* 1994; DiTullio *et al.,* 2000). Viruses of *Phaeocystis* species can infect during periods of rapid growth, resulting in the sudden crash and demise of the bloom. This has significant ecological consequences as the bacterial community composition utilises the newly released nutrients that result following viral cellular lysis (Baudoux and Brussaard, 2005). Intense blooms of *Phaeocystis* have been termed Harmful Algal Blooms (HABs), as they can result in anoxia and also produce a thick, mucilaginous foam, which deposits onto beaches as the bloom declines (Lancelot *et al.,* 1987).

Species of the prymnesiophyte genus *Chrysochromulina* are ubiquitous in the marine environment and have been found to make up between 2.5% and 65% of the total number of nanophytoplankton cells (Estep and MacIntyre, 1989; Hajdu *et al.,* 1996). The genus was first described in 1939 and today ca. 56 species have been described, with the aid of electron microscopy. *Chrysochromulina* rarely blooms but it can be lethal to fish stocks when it does (Dundas *et al.,* 1989). The earliest reports of VLPs in *Chrysochromulina* species were given following observations in thin sections by Manton & Leadbeater (1974). To date, a prymnesiovirus, *Chrysochromulina ericina* virus-01B (CeV-01B) has been isolated and partially sequenced that infects *Chrysochromulina ericina,* (Sandaa *et al.,* 2001) and the genome size has been reported as large at ca. 510 kb (Sandaa *et al.,* 2001). Primers have been designed which target the DNA pol gene and also the Major Capsid Protein (MCP) gene in CeV-01B (Larsen *et al.,* 2008).

1.4.2.2.9. Raphidoviruses

The class Raphidophyceae includes marine microalgae species, which are important in terms of the blooms they can often form (Figueroa and Rengefors, 2006). Certain raphidophyte species can cause toxic red tides (Khan *et al.,* 1997), which are responsible for fish kills and are particularly devastating to the aquaculture industry (Hiroishi *et al.,* 2005). *Heterosigma akashiwo* (Raphidophyceae) is a noxious microalga that forms dense blooms in coastal and subarctic regions of the global oceans and can impact on fisheries, leading to mortality in caged fish species such as salmon and sea bass (Honjo, 1993; Khan *et al.,* 1997). It has been reported that approximately 5% of *H. akashiwo* cells collected in Hiroshima Bay, Japan, during a bloom contained large VLPs intracellularly (Nagasaki *et al.,* 2005b) with the number of cells harbouring the VLPs increasing at the point of bloom termination, indicating a central role of virus infection in bloom control. The sole member of the genus *Raphidovirus* infects the red-tide forming species *Heterosigma akashiwo* and hence is named *Heterosigma akashiwo* virus (HaV). HaV was initially isolated and characterised in 1997 (Nagasaki and Yamaguchi, 1997). It has an icosahedral virion, approximately 202 nm in diameter and a dsDNA genome approximately 294 kb long (Nagasaki and Yamaguchi, 1997; Nagasaki *et al.,* 2005b).

It has also been observed that the susceptibility of *H. akashiwo* to infection by HaV differs among clonal strains (Nagasaki and Yamaguchi, 1998; Nagasaki *et al.,* 1999; Tarutani *et al.,* 2006). Further host strain analysis between *H. akashiwo* and HaV clones, suggest strain-dependent infection could play an important role in determining clonal composition and therefore allow intraspecies diversity in natural *H. akashiwo* populations (Tomaru *et al.,* 2004; Tarutani *et al.,* 2006). This diversity

could, in turn, promote the success of *H. akashiwo* as a ubiquitous bloom-forming species in coastal bodies of water.

Additionally, it is interesting to note, that since the isolation of HaV, several other types of viruses have been isolated which also infect this raphidophyte host. These include *H. akashiwo* nuclear inclusion virus (HaNIV), which forms paracrystalline arrays in the host's nucleus (Lawrence *et al.,* 2001); *H. akashiwo* RNA virus (HaRNAV), which has a positive-sense RNA genome and is seen to form crystalline arrays in the host cytoplasm (Tai *et al.,* 2003; Lang *et al.,* 2004); and a partially characterised dsDNA virus, OIs1, which consists of two different morphotypes that replicate simultaneously inside the host cell (Lawrence *et al.,* 2002; 2006). These findings would indicate the relationship between *H. akashiwo* and its infecting viruses is complex.

1.4.3. RNA viruses

1.4.3.1. Viruses infecting dinoflagellates and diatoms

In addition to the *Phycodnaviridae* family, whose members possess dsDNA genomes, ssRNA viruses have been reported that infect two toxic harmful algal bloom species, *Heterosigma akashiwo* (Raphidophyceae) (Tai *et al.,* 2003) and *Heterocapsa circularisquama* (Dinophyceae) (Nagasaki *et al.,* 2005b). The recently completed genome sequence of the ssRNA HaV led to the creation of the Marnaviridae family (Lang *et al.,* 2004). Also, a dsRNA virus of the family Reoviridae has been detected that infects *M. pusilla* (*Prasinophyceae*) (Brussaard *et al.,* 2004).

One of the most abundant and diverse groups of phytoplankton are the dinoflagellates (Dinophyceae), which are considered to contribute significantly as net primary producers and decomposers (Graham and Wilcox, 2000). There exists a large number of dinoflagellate species including various harmful bloom-forming species such as *Karenia mikimotoi* and *Heterocapsa ciruclarisquama* and toxic species such as *Alexandrium tamarense* and *Karenia brevis.* *H. circularisquama* is one of the most intensively studied species (Horiguchi, 1997). It forms dense blooms that can kill bivalves and can be responsible for mass mortalities in species such as the pearl oyster *Pinctada fucata martensii* and the mussel, *Mytilus edulis.* The viral impact on dinoflagellates was paid very little attention until recently. The first and the second cultured dinoflagellate-infecting viruses that were characterised were HcV (Tarutani *et al.,* 2001) and HcRNAV (Tomaru *et al.,* 2004), both of which infect *H. circularisquama.* Both are icosahedral virions with HcV being 180 – 210 nm long and HcRNAV ca. 30 nm long. HcV has a dsDNA genome ca. 356 kb in length and propagates within the host cytoplasm. It is a putative member of the *Phycodnaviridae,* although phylogenetic analysis on its DNA polymerase amino acid sequence has yet to be conducted. HcRNAV harbours a ssRNA genome ca. 4.4 kb long and replication also occurs in the host cytoplasm (Tomaru *et al.,* 2004). Research into the dynamics of the virus-host relationships and the ecology of these viruses is now ongoing (Nagasaki *et al.,* 2006).

Diatoms are important components of the marine biological community and food web and are the most abundant group of photosynthetic organisms with a universal distribution. The genus *Chaetoceros* is highly diverse with more than 400 species, which are considered to include the key primary photosynthetic producers forming

the base of the marine food chain (Rines and Hargraves, 1987). The first reports of viral infection in diatoms emerged following isolation of a ssRNA virus (RsRNAV) which produced a lytic infection in the pen-shaped diatom species *Rhizolena setigera* (Nagasaki *et al.,* 2004). A second diatom-infecting virus was described (CsNIV) which caused a lytic infection in *Chaetoceros salsugineum* (Nagasaki *et al.,* 2005b). The genome of this particular virus is somewhat distinctive, as it consists of a covalently closed circular ssDNA (6 kb long) that is partly double-stranded (997 bases). A third diatom virus, CspNIV that infects *C*. cf. *gracilis* has been reported (Bettarel *et al.,* 2005), although no data exist concerning the genome of this virus. Finally, a ssRNA virus, CtenRNAV has been isolated and characterised that infects *C. tenuissimus* (Shirai *et al.,* 2008).

1.4.4. Genome sequencing

Full genome sequences have been completed for members of four *Phycodnaviridae* genera: six chloroviruses – PBCV-1, NY-2A, AR158, MT325, FR483 and ATCV-1 (Li *et al.,* 1997; Fitzgerald *et al.,* 2007a; b; c); three phaeoviruses (Delaroque *et al.,* 2001; Delaroque *et al.,* 2003; Schroeder *et al.,* 2009); the coccolithovirus EhV-86 (Wilson *et al.,* 2005b; Allen *et al.,* 2006c); the prasinovirus, OtV-5 (Derelle *et al.,* 2008). Additional phycodnavirus genomes are undergoing sequencing but the data are unavailable at present. General features of the genomes of members of the *Phycodnaviridae* sequenced to date are summarised in Table 1.2.

Table 1.2. Characteristics of *Phycodnaviridae* genomes sequenced to date

Virus	Genome size (bp)	G + C%	No. of putative genes	Reference
Chlorovirus				
PBCV-1	330,744	40.0	366	Li *et al.,* 1997
NY-2A	368,683	40.7	404	Fitzgerald *et al.,* 2007
AR158	344,690	40.7	360	Fitzgerald *et al.,* 2007
MT325	314,335	45.3	331	Fitzgerald *et al.,* 2007
FR483	321,240	44.6	335	Fitzgerald *et al.,* 2007
ATCV-1	288,047	49.4	329	Fitzgerald *et al.,* 2007
Phaeovirus				
EsV-1	335,593	51.7	231	Delaroque *et al.,* 2001
FiV-1	191,667		156	Delaroque *et al.,* 2003
FsV-1	154,641	53.06	150	Schroeder *et al.,* 2009
Coccolithovirus				
EhV-86	407,339	40.2	472	Wilson *et al.,* 2005b
Prasinovirus				
OtV5	186,234	45.0	268	Derelle *et al.,* 2008

Genes for DNA methylation have been found in genomes of the chloroviruses, OtV5 and FsV-1. Integration genes have been found in the phaeoviruses, which are utilised for the latent phase of infection of these viruses (Delaroque *et al.,* 2003; Schroeder *et al.,* 2009). Genes have been found in all phycodnavirus genomes involved in DNA replication, protein synthesis, nucleotide synthesis, capsid proteins, transcription, and signal transduction. Unusual features include the presence of inteins in certain genes of some genomes e.g. two of the chlorovirus NY-2A genes, namely the large subunit of ribonucleotide reductase and a helicase encode inteins (Fitzgerald *et al.,* 2007b). The coccolithovirus, EhV-86 encodes a variety of putative sphingolipid pathway genes, previously undiscovered in a virus genome (Wilson *et al.,* 2005b). These genes may play a role in apoptosis. The EhV-86 genome is also the only member of the *Phycodnaviridae* to date to encode its own RNA polymerase raising the possibility of this virus diverging from the ancestral *Phycodnaviridae* earlier than the other members of this diverse family (Wilson *et al.,* 2005b).

1.5. Aims of this study

This study aims to isolate and characterise novel viruses that infect marine nano- and picophytoplankton species, thus enabling new host-virus model systems to be described and investigated. This should allow a greater insight to be gained into the dynamics and interactions of such systems. Despite their importance as key primary producers, there is a paucity of information about the viruses that infect these phytoplankton species. The characterisation and genetic analysis of novel viruses will enable molecular tools to be designed and employed to study the host-virus relationship in the natural environment. Targeting of key viral genes, e.g. the DNA polymerase gene or major capsid gene, will aid the design of novel diagnostic markers for use in the analysis of natural seawater samples. Community composition and structure, of both hosts and viruses, on a temporal and spatial scale can then be studied. Molecular data will allow the evolutionary history of algal viruses to be analysed, in addition to the influence of viruses on host evolution.

Analysis of the wider genetic content of novel virus strains, possibly whole genome sequencing, is also an aim of this study. Genomic information can allow isolated viruses to be screened for exploitable traits. Viruses are the most common biological agents in the marine environment and may represent the largest unexploited biotechnological resource on the planet. Newly isolated viruses could be screened for exploitation potential, e.g. novel enzymes and proteins encoded for by viruses; or where a lysogenic life cycle has been identified, a potential role for recombinant viruses as transfection vectors.

The main aim of the first experimental chapter was to detect, isolate and characterise a range of viruses that infect marine phytoplankton, in particular nanophytoplankton (3 to 20 µm in diameter) and picophytoplankton (<3 µm in diameter). Relatively few viruses have been isolated and described for members of the *Prasinophyceae* (picoeukaryotes) and *Prymnesiophyceae* (nanoeukaryotes) (Brussaard 2004b), despite their importance as key primary producers in the world's oceans. Seawater samples from a station in the Western English Channel were used to screen for viruses that infect species of the phytoplankton size fractions described. Techniques including TEM and analytical flow cytometry were employed to establish that the demise of a culture was the result of viral infection, prior to clonal virus isolation by plaque assay or liquid assay.

The second experimental chapter aimed to characterise the viruses isolated in chapter 3. Basic characterisation techniques included both traditional and molecular methods including TEM, analytical flow cytometry and pulsed field gel electrophoresis. The characterisation of novel viruses infecting key picophytoplankton groups (and especially their associated genome sequences) may help provide new genetic markers to allow assessment of the number of infected cells *in situ,* as this is still mostly unknown. This will in turn help to inform modellers of more accurate estimates of the importance of the viral shunt.

The third experimental chapter aimed to describe the complete sequence and annotation of one of the viruses isolated and characterised in chapters 3 and 4. This virus was found to infect the prasinophyte species, *Ostreococcus tauri* and specifically a high-light adapted strain of this species.

The fourth experimental chapter aimed to describe the complete sequence and annotation of another of the viruses isolated and characterised in chapters 3 and 4, namely one that was seen to infect and lyse a low-light adapted strain of *O. tauri.* As this is the third virus to be described that infects *O. tauri,* a comparative analysis between all three genomes sequenced to date was undertaken. Fundamental questions regarding the origins of these viruses and co-evolution with their hosts can be addressed by analysis of the virus genomes. Genome sequencing can provide insights into the evolutionary history of characterised viruses and particularly the origins of specific genes in the genome and the occurrence of horizontal gene transfer. As viruses are the most common biological agents in the marine environment they may represent the largest unexploited biotechnological resource on the planet. The isolation and characterisation of novel marine viruses is a priority research area in marine biodiversity, genomics and exploitation. Genome sequencing of newly isolated viruses will allow subsequent screening for exploitation potential, e.g. novel enzymes and proteins encoded for by viruses

Future work and a summary of the work conducted here are outlined in the final discussion chapter.

Chapter 2

Materials and Methods

2.1. Materials

Unless otherwise stated in this chapter all water used throughout the thesis was sterile double distilled MilliQ water.

2.1.1. Strains used in this study

2.1.1.1. Algal strains

The phytoplankton strains used in this study were obtained from the University of Warwick, courtesy of Professor David Scanlan, and the Plymouth Culture Collection (Marine Biological Association, UK; http://www.mba.ac.uk/culturecollection.php) and are shown in Table 2.1.

Table 2.1. Strains of phytoplankton used in this study are listed with their Roscoff Culture Collection (RCC) code and Plymouth Culture Collection (PCC) code.

Division	Class	Species	PCC Code	RCC code
Chlorophyta	*Prasinophyceae*	*Ostreococcus tauri*	-	745
Chlorophyta	*Prasinophyceae*	*Ostreococcus tauri*	-	501
Chlorophyta	*Prasinophyceae*	*Ostreococcus tauri*	-	143
Chlorophyta	*Prasinophyceae*	*Ostreococcus tauri*	-	393
Chlorophyta	*Prasinophyceae*	*Ostreococcus tauri*	-	356
Chlorophyta	*Prasinophyceae*	*Micromonas pusilla*	-	434
Bacillariophyta	*Coscinodiscophyceae*	*Corethron hystrix*	615	-
Bacillariophyta	*Coscinodiscophyceae*	*Odontella sinesis*	624	-
Bacillariophyta	*Coscinodiscophyceae*	*Stephanopyxis palmeriana*	625	-
Bacillariophyta	*Coscinodiscophyceae*	*Coscinodiscus wailesii*	695	-
Bacillariophyta	*Bacillariophyceae*	*Asterionellopsis glacialis*	607	-
Bacillariophyta	*Bacillariophyceae*	*Bacillaria paxillifera*	665	-
Chlorophyta	*Chlorophyceae*	*Dunaliella primolecta*	81	-
Chlorophyta	*Chlorophyceae*	*Dunaliella minuta*	430	-
Chlorophyta	*Chlorophyceae*	*Dunaliella tertiolecta*	83	-
Chlorophyta	*Chlorophyceae*	*Chlorella stigmatophora*	85	-
Chlorophyta	*Chlorophyceae*	*Chlorella salina*	309	-
Chlorophyta	*Chlorophyceae*	*Nannochloris atomus*	583	-
Chlorophyta	*Chlorophyceae*	*Chlamydomonas concordia*	491	-
Chlorophyta	*Chlorophyceae*	*Chlamydomonas reginae*	399	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis tetrathele*	272	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis striata*	443	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis impellucida*	429	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis convolutae*	372	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis suecica*	305	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis marina*	308	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis marina?*	570	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis rubens*	456	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis sp.*	315	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis sp.*	511a	-
Chlorophyta	*Prasinophyceae*	*Tetraselmis sp.*	512	-

Table 2.1. contd.

Division	Class	Species	PCC Code	RCC code
Chlorophyta	*Prasinophyceae*	*Tetraselmis sp.*	513	-
Chlorophyta	*Prasinophyceae*	*Pyramimonas amylifera*	246	-
Chlorophyta	*Prasinophyceae*	*Pyramimonas obovata*	280	-
Chlorophyta	*Prasinophyceae*	*Pyramimonas urceolata*	299	-
Chlorophyta	*Prasinophyceae*	*Pyramimonas parkeae*	492	-
Chromophyta	*Chrysophyceae*	*Ochromonas sp.*	2	-
Chromophyta	*Chrysophyceae*	*Ochromonas sp.*	504	-
Chromophyta	*Dictyochophyceae*	*Pseudopedinella sp.*	361	-
Chromophyta	*Dictyochophyceae*	*Pseudopedinella sp.*	438	-
Chromophyta	*Raphidophyceae*	*Heterosigma akashiwo*	12a	-
Chromophyta	*Raphidophyceae*	*Heterosigma akashiwo*	239	-
Chromophyta	*Raphidophyceae*	*Heterosigma akashiwo*	461	-
Chromophyta	*Eustimatophyceae*	*Nannochloropsis oculata*	663	-
Chromophyta	*Eustimatophyceae*	*Nannochloropsis salina*	591	-
Cryptophyta	*Cryptophyceae*	*Cryptomonas calceiformis*	175	-
Cryptophyta	*Cryptophyceae*	*Cryptomonas calceiformis*	412	-
Cryptophyta	*Cryptophyceae*	*Cryptomonas maculata*	174	-
Cryptophyta	*Cryptophyceae*	*Cryptomonas appendiculata*	23	-
Cryptophyta	*Cryptophyceae*	*Hemiselmis virescens*	157	-
Cryptophyta	*Cryptophyceae*	*Hemiselmis brunnescens*	14	-
Cryptophyta	*Cryptophyceae*	*Hemiselmis sp.*	631	-
Dinophyta	*Dinophyceae*	*Alexandrium tamrense*	173	-
Dinophyta	*Dinophyceae*	*Heterocapsa triquetra*	169	-
Dinophyta	*Dinophyceae*	*Amphidinium carterae*	127	-
Dinophyta	*Dinophyceae*	*Gymnodinium veneficum*	103	-
Dinophyta	*Dinophyceae*	*Gymnodinium simplex*	368	-
Euglenophyta	*Euglenophyceae*	*Eutreptiella gymnastica*	462	-
Haptophyta	*Prymnesiophyceae*	*Isochrysis galbana*	565	-
Haptophyta	*Prymnesiophyceae*	*Isochrysis aff. galbana*	8	-
Haptophyta	*Prymnesiophyceae*	*Isochrysis sp.*	240	-
Haptophyta	*Prymnesiophyceae*	*Isochrysis sp.*	506a	-
Haptophyta	*Prymnesiophyceae*	*Isochrysis sp.*	562	-
Haptophyta	*Prymnesiophyceae*	*Imantonia rotunda*	133	-
Haptophyta	*Prymnesiophyceae*	*Imantonia rotunda?*	577	-
Haptophyta	*Prymnesiophyceae*	*Ochrosphaera neapolitania*	162	-
Haptophyta	*Prymnesiophyceae*	*Ochrosphaera?*	341	-
Haptophyta	*Prymnesiophyceae*	*Ochrosphaera?*	434(5)	-
Haptophyta	*Prymnesiophyceae*	*Coccolithus pelagicus*	667	-
Haptophyta	*Prymnesiophyceae*	*Pleurochrysis elongata*	535	-
Haptophyta	*Prymnesiophyceae*	*Pleurochrysis carterae*	156	-
Haptophyta	*Prymnesiophyceae*	*Pleurochrysis carterae*	17	-
Haptophyta	*Prymnesiophyceae*	*Pleurochrysis carterae?*	378(1)	-
Haptophyta	*Prymnesiophyceae*	*Pleurochrysis sp.*	568	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium parvum*	94	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium patelliferum*	527	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium patelliferum*	527a	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium patelliferum*	527b	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium sp.*	569	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium sp.*	593	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium sp.*	594	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium sp.*	596	-
Haptophyta	*Prymnesiophyceae*	*Prymnesium sp.*	598	-
Haptophyta	*Prymnesiophyceae*	*Chrysotila stipitata*	377	-
Haptophyta	*Prymnesiophyceae*	*Chrysotila stipitata*	432	-
Haptophyta	*Prymnesiophyceae*	*Chrysotila lamellosa*	475	-
Haptophyta	*Prymnesiophyceae*	*Chrysotila sp.*	509	-
Haptophyta	*Prymnesiophyceae*	*Chrysotila sp.*	510a	-
Haptophyta	*Prymnesiophyceae*	*Chrysotila sp.?*	257	-

Table 2.1. contd.

Division	Class	Species	PCC Code	RCC code
Haptophyta	*Prymnesiophyceae*	*Hymenomonas globosa*	536	-
Haptophyta	*Prymnesiophyceae*	*Phaeocystis globosa*	64	-
Haptophyta	*Prymnesiophyceae*	*Phaeocystis globosa*	147	-
Haptophyta	*Prymnesiophyceae*	*Phaeocystis globosa*	540	-
Haptophyta	*Prymnesiophyceae*	*Geophyrocapsa oceanica*	572	-
Haptophyta	*Prymnesiophyceae*	*Dicrateria inornata*	564	-
Haptophyta	*Prymnesiophyceae*	*Chrysochromulina sp.*	307	-
Haptophyta	*Prymnesiophyceae*	*Chrysochromulina strobilus*	43a	-
Haptophyta	*Prymnesiophyceae*	*Chrysochromulina aff. polylepis*	200	-
Haptophyta	*Prymnesiophyceae*	*Pavlova salina*	154	-
Haptophyta	*Prymnesiophyceae*	*Pavlova aff. lutheri*	554	-
Haptophyta	*Prymnesiophyceae*	*Pavlova lutheri*	75	
Haptophyta	*Prymnesiophyceae*	*Pavlova pinguis*	471	-
Haptophyta	*Prymnesiophyceae*	*Pavlova sp.*	484	-
Haptophyta	*Prymnesiophyceae*	*Pavlova aff. salina*	486	-
Haptophyta	*Prymnesiophyceae*	*Pavlova gyrans*	93	-
Haptophyta	*Prymnesiophyceae*	*Pavlova virescens*	515	-
Haptophyta	*Prymnesiophyceae*	*Apistonema sp.*	508	-
Haptophyta	*Prymnesiophyceae*	*Cricosphaera?*	351	-
Haptophyta	*Prymnesiophyceae*	*Unidentified prymnesiophyte*	168	-
Heterokontophyta	*Pelagophyceae*	*Pelagomonas calceolata*	-	100
Rhodophyta	*Rhodophyceae*	*Rhodosarus marinus*	472	-

2.1.1.2. Cyanobacteria strains

The cyanobacteria strains used in this study were obtained from the University of Warwick, courtesy of Dr Katrin Zwirglmaier and are shown in Table 2.2.

Table 2.2. Cyanobacteria strains used in this study

Phylum	Order	Genus	Culture code
Cyanobacteria	Chroococcales	*Synechococcus*	WH7803
Cyanobacteria	Chroococcales	*Synechococcus*	WH8102
Cyanobacteria	Chroococcales	*Synechococcus*	RS9917
Cyanobacteria	Prochlorales	*Prochlorococcus*	MIT9313
Cyanobacteria	Prochlorales	*Prochlorococcus*	PCC9511

2.1.2. Chemicals and Molecular Biology Kits

Biochemical kits for molecular experiments were routinely employed throughout this study. General laboratory chemicals were obtained from Sigma-Aldrich (Poole, UK), Promega (Southampton, UK), or Fisher-Scientific (Leicester, UK), and were of analytical grade or higher. The enzymes *Taq* polymerase, RNase, DNase and Proteinase K were all obtained from Sigma-Aldrich. ExoSapIT was obtained from GE Healthcare Amersham Biosciences (Buckinghamshire, UK). The QiaPrep mini prep kit and QiaEx gel extraction kit were obtained from Qiagen Ltd (West Sussex, UK). Original TA cloning kits were obtained from Promega and *Escherichia coli* DH5α sub-cloning efficiency chemically competent cells were obtained from Invitrogen (Paisley, UK). ABI PRISM BigDye version 3.1 Terminator Cycle sequencing Ready Reaction Kit for DNA sequencing was supplied by Applied Biosystems. Oligonucleotides were acquired from Eurofins MWG Operon (Ebersburg, Germany). Ultra pure water used to make up media and buffers was derived from a Synergy 185 MilliQ water purification system (Millipore, Watford, UK). Sterilisation was performed by autoclaving for 20 minutes at 120 °C, 15 lb per square inch.

2.1.3. Equipment

Thermal cycling was performed in the MJ-Research PTC-100 thermal cycler. Pulsed field gel electrophoresis was carried out using the Bio-Rad C.H.E.F-DR II system. Electron microscopy was undertaken at the University of Plymouth (UK) electron microscopy facility, using a JEOL 1200 EX microscope. Analytical flow cytometry was performed using the Becton Dickenson FACScan flow cytometer (Becton-Dickenson, CA). Automated sequencing was carried out on the ABI Prism 3100 Genetic Analyser (Applied Biosystems). Centrifugation was either undertaken in the

bench top Eppendorf 5810 centrifuge using the A-4-62 rotor, or in a Beckman-Coulter Avanti J-26 XP centrifuge using the JS-5.3 rotor. Ultracentrifugation was undertaken in a Beckman-Coulter Optima L-100 XP or L8-M ultracentrifuge using the SW40ti swing out rotor. Filtration of seawater samples was performed using a Millipore tripod filter holder and a 50 kDa Molecular Weight Cut-Off (MWCO) tangential flow unit (Sartorius, Epsom, UK). Ultrafiltration was performed on the Quixstand system acquired from GE Healthcare Amersham Biosciences.

2.1.4. Media for the growth of phytoplankton

Media were prepared using ultra pure water obtained from a MilliQ water system.

All culture strains from the PCC and *Ostreococcus tauri* strain RCC 501 (Table 2.1) were grown in the growth medium f/2 (Guillard, 1975) using autoclaved seawater (95%). The remaining *Ostreococcus tauri* strains and *Micromonas pusilla* strain were grown in Keller (K) medium (Keller *et al.,* 1987). The cyanobacteria strains used for the host range analysis (Table 2.2) were grown in artificial seawater (ASW) and PCR S11. Composition of growth media used for phytoplankton are given in Table 2.3. All stock solutions added were filter sterilised (0.22 µm, syringe filter, VWR, Lutterworth, UK) and stored at 4 °C. All media components for each recipe (except vitamins) were added to 950 ml of filtered natural seawater and the final volume was brought up to 1 litre with dH_2O and autoclaved. The required amount of vitamins was added post-autoclaving, when the media had cooled to room temperature.

Table 2.3. Media components for the culture of algal strains grown in this study.

Medium	Component	Stock solution ($g \cdot l^{-1}$ dH_2O)	Quantity per litre of media	Concentration in final medium (M)
K	$NaNO_3$	75.0	1 ml	8.82×10^{-4}
	NH_4Cl	2.67	1 ml	5.00×10^{-5}
	Na_2β-glycerophosphate	2.16	1 ml	1.00×10^{-5}
	$H_2SeO_3 \cdot 9H_2O$	0.00129	1 ml	1.00×10^{-8}
	Tris-base(pH 7.2)	121.1	1 ml	1.00×10^{-3}
	Trace metals solution*		1 ml	-
	consisting of:	1° stock solution		
	$Na_2EDTA \cdot 2H_2O$	-		1.0×10^{-5}
	Fe-Na-EDTA $\cdot 3H_2O$	-		1.17×10^{-5}
	$FeCl_3 \cdot 6H_2O$	-		1.17×10^{-5}
	$MnCl_2 \cdot 4H_2O$	-		9.00×10^{-5}
	$ZnSO_4 \cdot 7H_2O$	23.00		8.00×10^{-8}
	$CoSO_4 \cdot 7H_2O$	14.05		5.00×10^{-8}
	$Na_2MoO_4 \cdot 2H_2O$	7.26		3.00×10^{-8}
	$CuSO_4 \cdot 5H_2O$	2.50		1.00×10^{-8}
	Vitamins solution#	-	0.5 ml	-
	consisting of:	1° stock solution		
	Thiamine · HCl (vitamin $B_{1)}$	200 mg		2.96×10^{-7}
	Biotin (vitamin H)	0.1		2.05×10^{-9}
	Cyanocobalamin (vitamin $B_{12)}$	1.0		3.69×10^{-10}
f/2	$NaNO_3$	75.0	1 ml	8.82×10^{-4}
	$NaH_2PO_4 \cdot H_2O$	5.0	1 ml	3.62×10^{-5}
	Trace metal solution*	-	1 ml	-
	consisting of:	1° stock solution		
	$FeCl_3 \cdot 6H_2O$	1.3	-	
	$Na_2EDTA \cdot 2H_2O$	8.7	-	
	$CuSO_4 \cdot 5H_2O$	980mg/100ml dH_2O	1.0 ml	
	$Na_2MoO_4 \cdot 2H_2O$	630mg/100ml dH_2O	1.0 ml	
	$ZnSO_4 \cdot 7H_2O$	2.2g/100ml dH_2O	1.0 ml	
	$CoCl_2 \cdot 6H_2O$	1.0g/100ml dH_2O	1.0 ml	
	$MnCl_2 \cdot 4H_2O$	18.0g/100ml dH_2O	1.0 ml	
	Vitamin solution#		0.5 ml	-
	See recipe for K medium			
ASW	NaCl	25.0	1.0 ml	4.2×10^{-1}
	KCl	0.72	1.0 ml	9.39×10^{-3}
	$MgCl_2 \cdot 6H_2O$	3.5	1.0 ml	5.46×10^{-2}
	$CaCl_2 \cdot 2H_2O$	1.4	1.0 ml	1.05×10^{-2}
	$NaHCO_3$	1.1	1.0 ml	5.0×10^{-3}
	$NaNO_3$	0.75		9.4×10^{-5}
	Vitamin solution#		1.0 ml	-
	See recipe for K medium			

Table 2.3. contd.

Medium	Component	Stock solution ($g \cdot L^{-1}$ dH_2O)	Quantity per litre of media	Concentration in final medium (M)
PCR S11	Na-PO_4 (pH 7.5)	-		5.0×10^{-5}
	$(NH_4)_2.SO_4$	-		4.0×10^{-4}
	HEPES-NaOH (pH7.5)	-		1.0×10^{-3}
	Na_2-EDTA/$FeCl_3$	-		8.0×10^{-5}
	Trace metal 'Gaffron + Se' consisting of:	-	0.5 ml	-
	H_3BO_3	0.185		1.5×10^{-4}
	$MnSO_4.H_2O$	0.101		3.0×10^{-5}
	$Na_2WO_4.2H_2O$	0.002		3.0×10^{-7}
	$(NH_4)_6Mo_70_{24}.4H_2O$	0.004		1.45×10^{-7}
	KBr	0.0072		3.0×10^{-6}
	KI	0.0049		1.5×10^{-6}
	$ZnSO_4.7H_2O$	0.0173		3.0×10^{-6}
	$Cd(NO_3).4H_2O$	0.009		1.5×10^{-6}
	$Co(NO_3)_2.6H_2O$	0.0087		1.5×10^{-6}
	$CuSO_4.5H_2O$	0.0075		1.5×10^{-6}
	$NiCl_2.6H_2O$	0.0071		1.5×10^{-6}
	$Cr(NO_3)_3.9H_2O$	0.0024		0.3×10^{-7}
	$VOSO_4.5H_2O$	0.0015		0.3×10^{-7}
	$KAl(SO_4)_2.12H_2O$	0.028		3.0×10^{-6}
	SeO_2	0.0033		1.5×10^{-6}

* Into 950 ml of dH_2O the EDTA was dissolved and then the components added. The final volume was brought to 1 litre, using dH_2O.

Into 950 ml of dH_2O, the thiamine HCl was dissolved. 1 ml of the primary stocks were added and the final volume brought to 1 litre with dH_2O. The vitamin stock was then filter-sterilised and frozen at -20 °C.

2.1.5. Buffers and media for molecular experiments

2.1.5.1. Universal buffers

The compositions of universal buffers commonly used throughout this study are detailed in Table 2.4.

Table 2.4. Composition of commonly used buffer solutions in this study.

Solution	Component	Quantity for 1 litre
50 × TAE	Tris-HCl	242 g
	Acetic Acid	57.1 ml
	0.5 M EDTA (pH 8.0)*	100 ml
	H_2O to 1 L	
10 × TBE	Tris-HCl	108 g
	Boric acid	55 g
	0.5M EDTA (pH 8.0)	40 ml
TE-buffer	1 M Tris-HCl (pH 8.0)	10 ml
	0.5M EDTA (pH 8.0)*	2 ml
	H_2O	988 ml

*0.5M EDTA – 186.1 g EDTA disodium salt added to 800 ml of dH_2O. 20 g NaOH pellets added whilst stirring to adjust pH to 8.0. Autoclave.

2.1.5.2. DNA Extraction Buffers

All buffers were stored at 4 °C before use.

2.1.5.2.1. Lysis buffers for the extraction of viral DNA

For the extraction of viral DNA, lysis buffer (0.05 M EDTA; pH 8, 1% (w/v) SDS) was freshly prepared using stock solutions prior to each extraction (Table 2.5).

Table 2.5. Components for viral lysis buffer

Stock Solutions	Volume required for 10 ml lysis buffer (ml)
0.5 M EDTA; pH 8.0	1.0
10% (w/v) SDS	1.0
dH_2O	8.0

2.1.5.2.2. CTAB buffer for the extraction of DNA

CTAB (hexadecyltrimethylammonium bromide) buffer (2% (w/v) CTAB, 1.4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl) was prepared for the extraction of DNA to a final volume of 100 ml (Table 2.6.).

Table 2.6. Components for CTAB buffer for the extraction of DNA

Component/Stock Solution	Requirements for 100 ml CTAB Buffer
CTAB	2 g
NaCl	8.18 g
0.5 M EDTA (pH 8.0)	4 ml
1 M Tris-HCl	10 ml

2.1.5.2.3. Pulsed Field Gel Electrophoresis (PFGE) Lysis Buffer

PFGE Lysis buffer (1% SDS (w/v), 100 mM EDTA, 0.5 mg ml^{-1} Proteinase K) was made up using stock solutions to a total volume of 15 ml (Table 2.7).

Table 2.7. Components for PFGE Lysis Buffer

Component	Requirements for 15 ml PFGE Lysis buffer
10% SDS	1.5 ml
0.5 M EDTA (pH 8.0)	3 ml
Proteinase K	0.0075 g

2.1.5.2.4. Cloning media and solutions

LB Medium (1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) was prepared for the growth and culture of *E. coli* for cloning experiments (Table 2.8). SOC medium (2% (w/v) tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 20 mM glucose) was prepared for transformation experiments (Table 2.8). Stock solutions of KCl, $MgCl_2$ and glucose were prepared separately. A 250 mM stock of KCl was prepared by adding 1.86 g to a total volume of 100 ml water. The 1 M stock of $MgCl_2$ was prepared by adding 20.33 g to a total volume of 100 ml water. A 2M stock solution of glucose was prepared by adding 36 g to a total volume of 100 ml water and by adjusting the pH to 7 using 1M NaOH.

Table 2.8. Composition of growth media for cloning experiments

Medium	Composition (per litre)		Supplier
Luria-Bertani broth (Lennox pre-mixed form)	10 g	Tryptone	Invitrogen (Paisley, UK)
	5 g	Yeast extract	
	10 g	NaCl	
Luria-Bertani agar	10 g	Tryptone	Invitrogen (Paisley, UK)
	5 g	Yeast extract	
	10 g	NaCl	Fisher-Scientific (Leicester, UK)
	15 g	Agar	
SOC	20 g	Tryptone	Invitrogen (Paisley, UK)
	5 g	Yeast extract	
	0.5 g	NaCl	
	10 ml	1 M $MgCl_2$*	
	10 ml	1 M $MgSO_4$*	
	2 ml	20% (w/v) Glucose*	

* Pre-prepared sterile stocks

2.1.5.3. Loading Dye for Standard Agarose Gel Electrophoresis

6 × agarose gel loading dye (0.25% (w/v) bromophenol blue, 40% (w/v) sucrose, 20 mM EDTA) was made up in 100 ml volumes of water. Typically 1 µl 6 × loading dye was added to 10 µl of the sample to be electrophoresed.

2.1.6. Oligonucleotides

Oligonucleotides were acquired from Eurofins MWG Operon (Ebersburg, Germany), unless otherwise stated. Each primer was supplied at 50-nmol scale of synthesis with no additional modifications. On receipt, each oligonucleotide was resuspended in sterile ultra pure H_2O to a 10 × stock concentration of 100 pmol μl^{-1} and stored at -20°C. Oligonucleotides were stored as 10 × stock solution and working solutions were made up from this stock, thus preventing the need for freeze-thawing of the stock concentrations, which can cause degradation. These 10 × stock solutions were diluted ten-fold in H_2O to give a 10 pmol μl^{-1} final working solution, for use in PCR reactions. M13 oligomers were used to amplify DNA fragments ligated into pGEM®-T Easy Vectors (Promega, Southampton, UK). PCR primers used in this study were acquired from the literature (Table 2.9) or alternatively were designed as a result of this study (Appendix I).

Table 2.9. PCR primers used in this study.

Primer name	Sequence (5'–3')	Reference/Supplier
AVS-1	GA(A/G)GGGCACGT(T/C)TGA(T/C)GC	Chen & Suttle 1996
AVS-2	GCGC(A/G)TAC(G/T)(T/C)TT(T/C)TT(G/C)(A/T)(A/G)TA	Chen & Suttle 1996
POL	(G/C)(A/T)(A/G)TCGT(A/G)TCCC(A/G)TA	Chen & Suttle 1996
M13F	GTA AAA CGA CGC CCA G	Invitrogen
M13 R	CAG GAA ACA GCT ATG AC	Invitrogen
EUK328F	ACC TGG TTG ATC CTG CCA G	A. Kirkham, pers.comm.
EUK329R	TGA TCC TTC YGC AGG TTC AC	A. Kirkham, pers.comm.

2.2. Methods

2.2.1. Maintenance of phytoplankton cultures

Stock cultures of PCC isolates (Table 2.1) were grown at 15 °C, under a 16:8 light:dark cycle at irradiance between 10 and 50 µmol m^{-2} s^{-1}(measured using a Traceable® Dual-Range light meter (Fisher-Scientific). Stock cultures of *O. tauri* and *M. pusilla* were grown at 20 °C, under a 16:8 light:dark cycle, at irradiance between 4.0 and 100 µmol m^{-2} s^{-1} in a Sanyo versatile environmental chamber model MLR-350 (Sanyo, Japan) and cyanobacteria cultures (Table 2.2) for host range analysis at 22 °C, under constant light conditions, at irradiance between 2.5 and 7 µmol m^{-2} s^{-1} in a Gallenkamp cooled incubator (Sanyo Gallenkamp, Loughborough, UK). All cultures were maintained in high optical quality culture flasks (Nunc, Hereford, UK) and sub-cultured with fresh media (Table 2.3) every 10-14 days.

2.2.2. Sample collection

Seawater sampling was conducted between October 2006 and September 2007 at the sampling station L4 in the Western English Channel on a weekly basis, prevailing conditions permitting. The co-ordinates of the L4 sampling station are 50°15'N, 04°13'W. Seawater samples were collected at 0-5 metres with Niskin bottles mounted on a conductivity-temperature-density (CTD) profiler.

2.2.3. Virus samples

2.2.3.1. Samples for virus isolation

20 l of seawater was collected from the L4 sampling station for the isolation of algal viruses (Fig. 2.1.i) and was pre-filtered through low-protein binding GF/F 1.6 µm filters with 142 mm diameter (Whatman) (Fig. 2.1.iii). The pre-filtered 20 l of seawater was processed further as follows (Fig. 2.1) : a) 50 ml passed through a 0.2

µm filter (Fig. 2.1.iv); b) 50 ml passed through a 0.45 µm filter (Fig. 2.1.v); c) 1 l 0.45 µm filtered and enriched with media nutrients (Fig. 2.1.vi); d) the remaining 18 l of the 1.6 µm filtrate was concentrated to a final volume of approximately 20 ml by spiral cartridge ultrafiltration (300,000 Da molecular weight cut-off; Quixstand™ GE Healthcare), to concentrate the virus fraction (Fig. 2.1.vii). All filtrates and concentrated water were stored at 4 °C to maintain viability. In addition, 2 ml of concentrate was stored at -80 °C for DNA extraction at a later date (Fig. 2.1.ix).

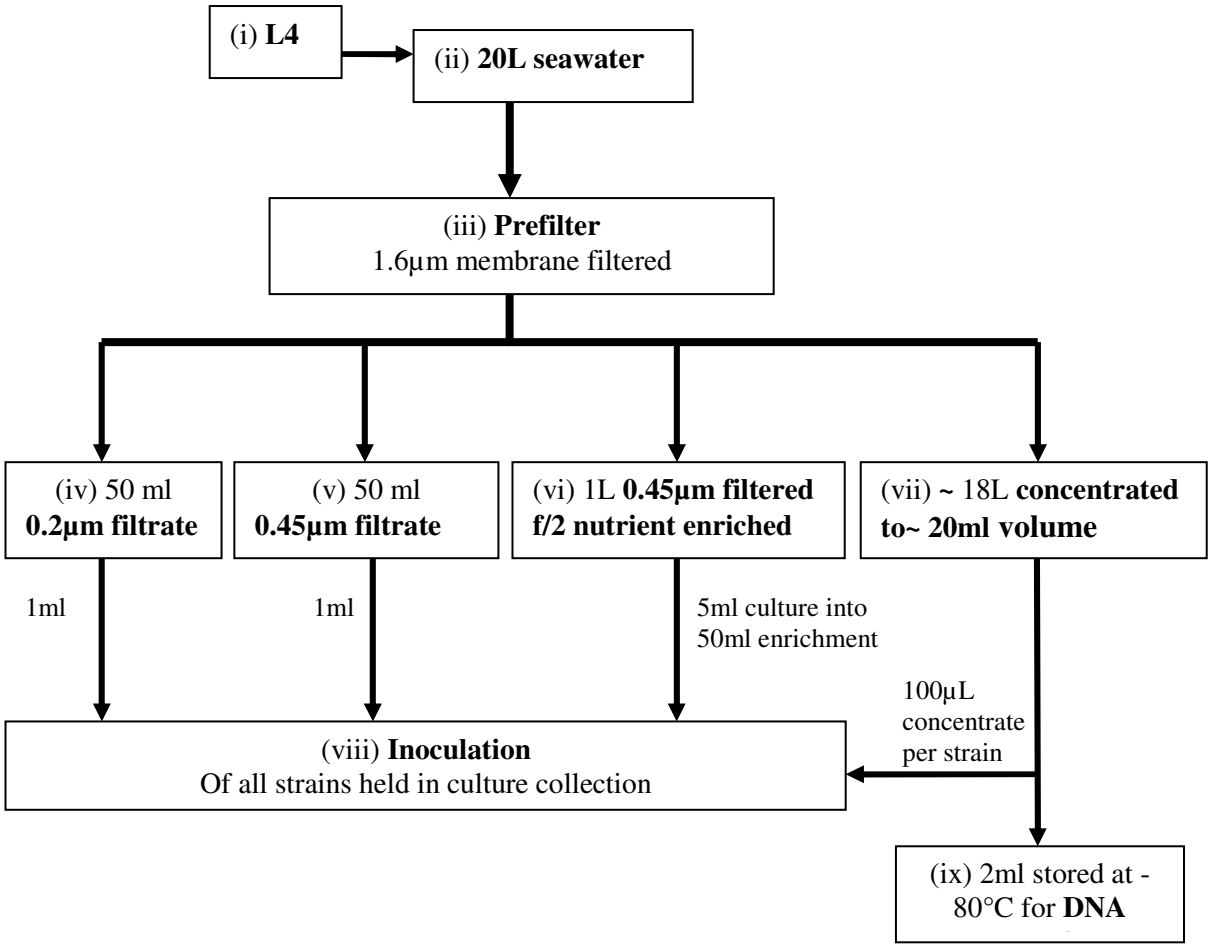


Figure 2.1. Flow diagram to represent the processing of L4 seawater for viral screening of phytoplankton strains held in culture.

2.2.4. Virus isolation

2.2.4.1. Isolation of new viruses

50 ml volumes of individual exponentially growing cultures (growth measured by analytical flow cytometry, section 2.2.7) from Table 2.1 were inoculated with either 1 ml of 0.2 µm filtered seawater; 1 ml 0.45 µm filtered seawater; 100 µl of concentrated seawater; or, alternatively, 5 ml of algal culture was added to 50 ml of enriched media (Fig. 2.1.viii). 50 ml volumes of each phytoplankton species were inoculated with 1 ml autoclaved filtered seawater to serve as negative controls. The experiments were carried out under the culturing conditions described in Section 2.2.1. The cultures were inspected on a daily basis for visible signs of lysis, indicated by a lack of growth or clearing of the culture and comparison with uninoculated control cultures. If lysis was observed, 1 ml of lysate was fixed for flow cytometric analysis (Section 2.2). The remainder of the lysed culture was passed through a 0.45 µm filter (Gelman) and the filtrate was used to inoculate a fresh exponentially growing culture of the sample strain.

2.2.4.2. Clonal isolation

Clonal isolates of viruses were obtained preferentially by plaque assay purification or, if this was not possible, by serial dilution in liquid cultures. Samples from the scheme outlined in Figure 2.1 were used for virus isolation.

2.2.4.2.1. Plaque assays

The *Ostreococcus tauri* virus strain OtV-1 isolated in this study was purified by plaque assay, as described by Schroeder *et al.,* (2002), in order to obtain clonal virus stocks. 50 ml volumes of exponentially growing cultures of *O. tauri* strain OTH 95 (Table 2.1) were concentrated by centrifugation at 5,000 × *g* in an Eppendorf 5810

benchtop centrifuge A-4-62 rotor for 5 minutes at 15 °C. The supernatant was removed and the algal pellet was re-suspended in the overlying medium to give a final volume of approximately 1 ml. Tenfold dilutions of virus inocula were prepared in K medium (Table 2.3) and 100 µl of the appropriate dilutions were added to 1 ml of suspended host cells and gently mixed. The virus/host suspensions were incubated for 3 hours under normal culture conditions (section 2.2.1) in the light to allow for virus adsorption.

3 ml of 0.4% (w/v) molten electrophoresis grade agarose K medium mixture (Table 2.3) was added to each virus/host cell suspension and was mixed thoroughly before pouring onto a solid 1.5% (w/v) agarose/seawater/K media nutrients plate. Control plates contained algal cells and K medium only. Solid medium was allowed to set at room temperature and plates were then inverted, placed in plastic, transparent bags and incubated at 20 °C under the growth conditions outlined in section 2.2.1. Plates were monitored daily until clear areas of non-growth were visible. Plaques were picked by stabbing the tip of a sterile Pasteur pipette into the plaque, which was then transferred into 0.5 ml sterile media (Table 2.3) and used for subsequent inoculations after overnight incubation at 4 °C. Agarose and any remaining host were removed from the samples by centrifugation at 16,000 × *g* for 2 minutes in the Eppendorf microfuge 5415 D. The supernatant was transferred to a sterile Eppendorf tube and stored at 4°C in the dark until further use. In total, three rounds of plaque purification assay were performed for each virus isolate.

2.2.4.2.2. Liquid assays

All of the MpV isolates and OtV isolates, except OtV-1, were purified by serial dilution to extinction (Cottrell and Suttle, 1991), in order to obtain clonal virus

stocks. Virus lysate was obtained by adding 100 µl of concentrated seawater to exponentially growing host cultures. Once clearing of the host culture was observed, the lysate was passed through a 0.2 µm PVDF filter (Durapore, Millipore).

Filtered lysate was used to infect exponentially growing phytoplankton cells in a 9-step series of 10-fold dilution steps from 1 to 10^{-8}. Cell lysis was recorded as the appearance of a virus group and a decline in fluorescence using an analytical flow cytometer (Section 2.2.7). The algal lysate was filtered on a 0.2 µm PVDF filter (Durapore, Millipore) and the procedure repeated a further two times. Filtered algal lysate was then diluted in an 8-step series of 10-fold dilution steps. Aliquots (100 µl) of each dilution were added to 3 wells of a 24-well assay plate, each well containing 2 ml of exponentially growing culture. The assay plate was incubated as described for host cultures (section 2.2.1) and monitored for discolouration over 14 days. One of the wells in the most diluted series showing lysis was propagated in a flask containing host strain culture. After lysis, the content of the flask was filtered and re-incubated in an assay plate following the same procedure described above for two further rounds. The clonal virus sample obtained was filtered and stored at 4 °C in the dark.

2.2.5. Virus Purification & Concentration

Virus concentration was carried out by PEG precipitation and subsequent purification by caesium chloride gradient ultracentrifugation.

2.2.5.1. PEG Precipitation

58.4g l^{-1} NaCl was added to the viral lysate and dissolved by gently mixing on ice for 1 hour. Cellular debris was removed by centrifugation for 10 minutes at 6,000 × *g* in

the Eppendorf 5810 centrifuge. The supernatant was transferred to 50 ml volume sterile Falcon tubes before 10% (w/v) PEG 6000 (PEG; Sigma-Aldrich, Poole, UK) was added to the lysate and dissolved gently at room temperature before incubating at 4 °C for at least 2 hours. Viruses were recovered from the lysate by centrifugation for 30 minutes at 6,000 × *g* in the Eppendorf 5810 centrifuge. The supernatant was removed and the pellet was re-suspended in approximately 12 ml autoclaved filtered seawater.

2.2.5.2. Caesium Chloride Gradient purification

Caesium chloride was dissolved into 3 ml aliquots of concentrated virus to give final densities of 1.0 g cm^{-3} (0.03 g in 3 ml virus concentrate), 1.2 g cm^{-3} (0.8319 g in 3 ml virus concentrate), 1.4 g cm^{-3} (1.644 g in 3 ml virus concentrate) and 1.6 g cm^{-3} (2.4495 g in 3 ml virus concentrate). The CsCl-containing virus suspensions were used to prepare gradients in clean 12 ml volume Beckman ultracentrifuge tubes. The densest solution (1.6 g cm^{-3}) was added to the bottom of the tube and was overlaid with the less dense solutions in descending order. The gradients were centrifuged at 100,000 × *g* for 2 hours at 20 °C in a SW40ti swing out rotor in a Beckman L8-M Ultracentrifuge. Deceleration was programmed to setting 7 in order not to disturb the gradient following centrifugation. Virus bands could be observed within the gradient and were collected by piercing the side of the tube just beneath the band with a needle attached to a syringe. Virus bands were drawn out from the tube and were transferred to dialysis tubing. Dialysis proceeded overnight at 4 °C against 1 litre sterile filtered seawater. Purified viruses were collected and stored at 4 °C.

2.2.6. Host range determination

For host range determination, fresh working solutions of virus lysates were obtained from stock lysates stored at 4 °C. Briefly, 100 µl of each lysate was added to 35 ml of exponentially growing cultures of the host strain originally used for their isolation. Once clearing of the host culture was observed, the lysate was passed through a 0.2 µm syringe filter (Gelman) and the filtrate containing virus was stored at 4 °C. The host range for each virus strain was determined by adding 500 µl of each virus lysate to 50 ml of several exponentially growing phytoplankton cultures. The inoculated cultures were occasionally agitated to encourage virus adsorption. Growth of the host cultures was monitored daily over a 14 day period. Lysis was determined by colour comparison to 20 ml of uninfected cultures and by AFC (Section 2.2.7.). Cultures that did not lyse after 14 days following viral inoculation were considered to be non-susceptible to infection by the particular virus strain. The experiment was repeated twice more in order to verify the results.

2.2.7. Analytical flow cytometry (AFC)

The presence of viruses in lysed cultures was confirmed by analytical flow cytometry. AFC analyses were performed on a FACScan flow cytometer (Becton Dickinson, New Jersey, USA) equipped with an air-cooled laser providing 15 mW at 488 nm and with standard filter set-up. Typically, algal counts were obtained from fresh samples at high flow rate (average 97 µl/minutes). The trigger was set on red fluorescence and the samples were run on the flow cytometer for 120 seconds.

Virus enumeration was performed on fixed samples, following the protocol described by Brussaard (2004a). 1 ml samples were fixed with 20 µl of 25 % glutaraldehyde (0.5% final concentration) for 30 minutes at 4 °C, then snap frozen in liquid nitrogen and stored at -80 °C until analysed. Prior to analysis, the samples were thawed at

37°C, diluted in TE buffer (Table 2.4) and stained with SYBR Green I (Molecular Probes Inc., Eugene, USA) in a water bath at 80 °C for 10 minutes. The commercial stock solution of SYBR Green I was diluted with TE buffer to a final dilution 1.0×10^{-4}, and kept in the dark. A fresh working stock of SYBR Green was made by adding 10 µl of 1:10 50% stock ml^{-1} to 990 µl sterile TE buffer. 10 µl of diluted sample was added to 940 µl TE-buffer and 50 µl SYBR Green working stock. This was incubated in the dark at 80 °C for 10 minutes then cooled at room temperature for 5 minutes. Virus samples were run at low flow rate (average 35 µl/minutes), for 120 seconds at an event rate between 200 and 900 per second.

The discrimination of virus and bacteria groups was based on groups observed in scatter plots of SSC signal versus green fluorescence. Fluorescent latex microsphere beads (Molecular Probes Inc., Eugene, USA) with a diameter of 0.95 µm were added to all samples as an internal reference to calculate the flow rate and normalise the signal of the cytometer. Measurements of side scatter and green fluorescence were made using CellQuest software (Becton Dickinson) with log-amplification on a four-decade scale. Listmode files were further analysed using the software program Win MDi 2.8 (Joseph Trotter, Scripps, San Diego, CA, USA – freely available from http://facs.scripps.edu/).

2.2.8. Transmission Electron Microscopy (TEM)

The presence of Virus-Like Particles (VLPs) in lysates and the morphology of VLPs were confirmed by TEM. To visualize viruses and virus-like particles, grids and thin sections were prepared and examined under a Transmission Electron Microscope (TEM).

2.2.8.1. TEM fluid sample grid preparation

1 ml concentrated viral lysate samples were fixed in 20 µl electron microscopy grade (50%) glutaraldehyde (final concentration 0.5%). 10 µl volumes of fixed concentrated viral lysate were dropped on to formvar-coated copper grids and allowed to adsorb for 10 minutes. The grids were blotted on filter paper and stained with 2% (w/v) uranyl acetate for 30 seconds. Grids were blotted again and rinsed with H_2O. Excess liquid was removed with filter paper. Grids were allowed to air dry before examination. Samples were analysed with a JEOL 1200 CX TEM operated at 160Kv at magnifications ranging from ×800 to ×600,000. Photographic images were taken at magnifications between ×5,000 and ×50,000.

2.2.8.2. Thin-sectioning

2.2.8.2.1. TEM thin section preparation

Samples of fixed virus-infected cells were spun at 13,000 × *g* in an Eppendorf 5415 D microfuge for 1 minute and the supernatant discarded. The fixed cells were then washed with 0.2 µm filtered autoclaved seawater, spun for 1 minute and the supernatant discarded. This wash step was repeated twice. 500 µl of 2% molten agar in dH_2O was added to the pellet and allowed to set. Excess agar was trimmed from the samples. Samples were transferred into a 4% osmium tetraoxide (OsO_4) solution and rotated for approximately 1 hour at room temperature, followed by 2 washes in 0.22 µm filtered seawater. Series dehydration was carried out using 30%, 50%, 70% and 90% (v/v) ethanol, for 15 minutes each, then using 100% ethanol overnight at room temperature. The dehydrated samples were embedded in EPON resin embedding resin (TAAB Laboratory Equipment Ltd., Spurr's Resin) in pre-shaped moulds and dried for 48 hours at 35 ºC. Dried blocks were cut to size and shaped into sections of 70 - 80 nm (500-700 nm) thickness on a Reichert-Jung

Ultracut Microtome. The sections were placed on Formvar-coated grids and dried overnight at 35 °C before staining. Thin sections were stained as outlined in Section 2.2.8.1.

2.2.9. Molecular Biology Techniques

2.2.9.1. Host DNA Extraction

Lysis buffer (0.5 % SDS, 20 µg ml^{-1} Proteinase K) for the extraction of algal host DNA was prepared fresh for each extraction using a stock solution of 10% SDS and was made up to 10 ml with water. 100 ml of algal cells were concentrated by centrifugation in 50 ml volumes for 15 minutes in the bench top Eppendorf 5810 centrifuge using the A-4-62 rotor, at 5,000 × *g*. The supernatant was removed and the pelleted cells were re-suspended in 500 µl lysis buffer and incubated at 55 °C for 1 hour.

80 µl 5 M NaCl and 150 µl CTAB buffer (Table 2.6), pre-warmed to 60 °C, were added to the sample. This was mixed thoroughly by inversion and incubated at 60 °C for 15 minutes. Proteins were removed by extraction in an equal volume of chloroform:isoamylalcohol (24:1). DNA precipitation was carried out in 0.6 × volume isopropanol at room temperature for 1 hour. DNA was recovered by centrifugation for 30 minutes at 16,000 × *g* in the Eppendorf 5415 D microfuge. The supernatant was removed and discarded and the pellet was washed in 200 µl 70% (v/v) ethanol. The supernatant was again discarded and the DNA pellet was air dried and re-suspended in 30 µl TE buffer (Table 2.4).

2.2.9.2. Viral DNA Extraction

Viral DNA was extracted from purified viral lysate (section 2.2.5.). 10% (v/v) lysis buffer (Table 2.5) was added to 1 ml virus lysate and was incubated at room temperature for 15 minutes. One round of phenol extraction was followed by at least one round of 1:1 (v/v) phenol:chloroform extraction until no debris could be detected at the phenol:chloroform/sample interface. The final aqueous layer was retained and the proteins were extracted by adding 0.5 × volume 7.5 M ammonium acetate. The suspension was incubated at room temperature for 30 minutes before centrifugation at 16,000 × *g* for 15 minutes in the Eppendorf 5415 D microfuge. The supernatant was retained and DNA was extracted by ethanol precipitation overnight at 4 °C. The DNA pellet was washed in 200 µl 70% ethanol and was subsequently allowed to air dry before re-suspension in up to 50 µl TE (Table 2.4).

2.2.9.3. Nucleic acid quantification

Genomic DNA concentrations were determined by absorbance at 260 nm using a BioPhotometer spectrophotometer (Eppendorf, Cambridge, UK). 10 µl of sample was diluted in 1 ml of ultra pure MilliQ water prior to reading.

2.2.9.4. Amplification of DNA fragments

2.2.9.4.1. Polymerase Chain Reaction (PCR)

PCR was frequently employed in this project on a range of templates and using a range of conditions. PCR primers used in this study were acquired from the literature (Table 2.9) or alternatively were designed as a result of this study (Appendix I).

2.2.9.4.2. Typical Reaction Conditions

All PCR reactions were conducted in a PTC-Peltier Thermal Cycler (MJ Research). Typical PCR reactions were carried out in 50 µl volumes containing 1 × *Taq* polymerase buffer, 2 mM $MgCl_2$, 0.025 mM each deoxynucleosidetriphosphate (dNTPs), 20 pmol forward primer, 20 pmol reverse primer, 1 unit *Taq* polymerase and approximately 10 - 100 ng template DNA (alternatively 1 µl virus lysate) and molecular biology grade water (SigmaAldrich). PCR reactions typically commenced with an initial denaturation step at 95 °C for 5 minutes, followed by 30 cycles of denaturation at 95 °C for 30 seconds, annealing at 50 °C for 1 minute, extension at 74 °C for 1 minute, then a final extension step at 74 °C for 5 minutes, unless otherwise stated in results chapters. Finally the reactions were cooled down to 15°C. PCR products were verified by agarose gel electrophoresis of 5 µl aliquots (Section 2.2.9.4.4.). For increasing amplification specificity, nested (two stages) PCR reactions were performed using 2.5 µl of the first stage PCR products as template.

2.2.9.4.3. Genome amplification

Multiple displacement amplification reactions were used to amplify low concentrations of genomic DNA. This proportionally reduced the concentration of PCR-inhibitors. Reactions were performed using the GenomiPhi DNA amplification kit (Amersham Biosciences, Piscataway, NJ) in 20 µl reactions following the manufacturer's instructions. Briefly, 1 µl template DNA in 9 µl sample buffer (50 mM Tris-HCl, pH8.2), 0.5 mM EDTA) containing random hexamers was denatured at 95 °C for 3 minutes and placed on ice. Buffer (9 µl) containing dNTPs and 1 µl enzyme mix containing Φ29 DNA polymerase were added to the 10 µl of denatured DNA template-random hexamers solution and incubated at 30 °C for 16 hours. A final incubation at 65 °C for 10 minutes inactivated the Φ29 DNA polymerase.

2.2.9.4.4. Agarose Gel Electrophoresis

Ultrapure agarose (SeaKem GTM) at a concentration of 1.2% (w/v) was dissolved in 1 × TBE buffer (Table 2.5) by heating to boiling point in a microwave oven. The solution was cooled down to approximately 60 C and ethidium bromide was added to a final concentration of 0.5 ug/ml. The solution was then poured into a gel casting tray and allowed to solidify at room temperature. Samples were loaded in the gel with 1 × bromophenol blue loading buffer (Section 2.1.5.3). Electrophoresis was performed in 1 × TBE buffer at 90 – 120 volts for the desired duration (according to the degree of separation required). Band sizes were estimated using 100 bp or 1 kb size markers (Promega) run alongside the samples. Gels were subsequently visualised using short wave UV light on a UV transilluminator, and photographed with a Gel Doc 2000 system (Bio-Rad). If the DNA was to be extracted from the gel, long wave UV light was used for a minimal exposure time. Fragments were excised from the gel using a sterile scalpel and were extracted from the agarose using the QiaEx II gel extraction kit according to the manufacturer's instructions.

2.2.9.5. Cloning

Cloning was undertaken using biochemical kits detailed in Section 2.1.2 according to manufacturer's instructions unless otherwise stated.

2.2.9.5.1. Ligations

Ligations were performed using the pGEM-T cloning kit (Promega, Southampton, UK) according to the manufacturer's instructions. Two units of T4 DNA ligase (Promega, Southampton, UK) were used per reaction. Ligations were performed at 16^{o}C for 12 hours.

2.2.9.5.2. Transformation of *Escherichia coli* cells with plasmid DNA

50 µl aliquots of *Escherichia coli* DH5α chemically competent cells (Invitrogen, Paisley, UK) were thawed on ice and 1 µl of plasmid DNA was added to the cell suspension. After incubation on ice for 30 minutes, the cells were heated at 42 C for 45 seconds and returned to the ice for a further 2 minutes. To facilitate the expression of antibiotic proteins, 450 µl SOC media was added and incubated for 30 minutes – 1 hour at 37 C, shaking at 225 rpm. The cell suspension was plated out onto LB agar plates containing ampicillin (Sigma-Aldrich, Poole, UK) at a final concentration of 50 µg ml^{-1} and X-gal (5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside; Sigma-Aldrich, Poole, UK) to a final concentration of 20 µg ml^{-1} and incubated at 37 C overnight.

The resultant *E. coli* colonies were screened by PCR (colony PCRs). This involved picking and suspending aliquots of individual *E. coli* colonies in 50 µl H_2O and boiling for 5 minutes. The resultant cell mix was then vortexed and centrifuged and the supernatant was used in a second PCR to re-amplify the cloned inserts.

2.2.9.6. Pulsed Field Gel Electrophoresis (PFGE)

PFGE was used to separate full length viral genomic DNA from lysates of host cells, identify genome size and purify viral DNA. PFGE analysis was performed according to the protocol described by Larsen et al (2001). Virus lysates were concentrated by PEG precipitation (section 2.2.5.1). Plugs were cast by mixing 50 µl of the resulting viral concentrates and 50 µl 1.5% (w/v) InCert agarose (FMC, Rockland, Maine) and dispensing into plug molds. After the gel had solidified, the plugs were pushed gently out of the moulds and placed in 2 ml lysis buffer (Table 2.7). The plugs were

incubated in the dark at 30 °C overnight. The lysis buffer was decanted and the plugs were washed three times, for 30 minutes each time, in 10 ml TE buffer (pH 8.0) for 1 hour. If immediate analysis was not to be conducted, virus agarose plugs were stored at 4 °C in TE buffer. The virus plugs were loaded into the wells of a 1% SeaKem GTG agarose (FMA, Rockland, Maine) gel in 1 × TBE buffer and sealed with an overlay of molten 1% agarose. Samples were electrophoresed using a Bio-Rad DR-II CHEF Cell (Bio-Rad, UK) electrophoresis unit operating at a voltage of 6.0 V cm^{-1} with pulse ramps from 5.5s to 35.5 seconds at 14 °C for 21 hours in 1× TBE tank buffer. A *Hin*dIII digest phage lambda concatemer (Sigma-Aldrich, UK) and a *Sacchromyces cerevisiae* concatemer (BioRad, Richmond, CA) were used as molecular weight markers. Following electrophoresis, the gels were stained for 30 minutes with ethidium bromide and DNA bands were visualised on a UV transilluminator (Syngene, Cambridge, UK).

2.2.9.7. Automated DNA Sequencing

DNA sequencing was carried out using the primers detailed in Section 2.1.6. If the fragment to be sequenced was a PCR product, it was firstly incubated with ExoSap-IT (GE Healthcare), according to the manufacturer's instructions, in order to eliminate any unincorporated dNTPs or primers. Sequencing reactions were carried out using the ABI Big-dye version 3.1 sequencing kit, following the manufacturer's instructions. Sequencing reactions were carried out at $1/16^{th}$ dilution using approximately 70 ng template DNA. Typically, sequencing reactions were incubated at 96 °C for 1 minute followed by 25 cycles of 96 °C for 10 seconds, 50 °C for 5 seconds and 60 °C for 4 minutes. Reactions were stored for a maximum of 48 hours at 4 °C.

Dye terminators were removed from the sequencing reactions by an ethanol/EDTA precipitation reaction as follows: 5 µl 125 mM EDTA and 60 µl 100% ethanol were added to each sample and gently mixed. The mixture was then incubated at room temperature for 15 minutes in the dark. The tubes were then centrifuged at 3,000 × *g* for 30 minutes at 4 °C. The supernatant was discarded and the pellet was washed in 60 µl 70% ethanol at 1650 × *g* for 15 minutes. Finally, the air dried pellet was re-suspended in 15 µl Hi Di formamide (Applied Biosystems, UK). The re-suspended DNA was denatured by heating at 94 °C for 3 minutes prior to sequencing. Sequencing reactions were run on the ABI Prism 3100 Genetic analyser using a 50 vcm array or for longer sequencing reads, the 80 vcm array. Sequence data was automatically collated and analysed using the ABI sequencing analysis software and was subsequently manually verified. The data for each fragment was analysed using DNAStar software (Lasergene) and aligned using ClustalW (http://www.ebi.ac.uk/clustalw/).

The sequencing of whole genomes was performed at the Advanced Genomics Facility based at the University of Liverpool, UK, using a GS-FLX 454 genome sequencer.

2.2.9.7.1. DNA Sequence Analysis

Sequences generated by the ABI Prism 3100 genetic analyser were collated and analysed using the ABI sequencing analysis software. The quality of the sequences was assessed by eye and any anomalies were corrected. Sequences were either exported to the BioEdit software (http://www.mbio.ncsu.edu/BioEdit/BioEdit.html) to align the forward and reverse sequences of each sample, or were exported to

SeqmanII (DNAstar, Lasergene, Madison, WI) to assemble multiple sequences. Mismatches were checked by eye, and a consensus sequence was taken.

2.2.9.7.2. Database Searches

Similarities between virus sequences and published sequences were determined using BLAST (Basic Local Alignment Search Tool) within the National Centre for Biotechnology Information (NCBI) (http://www.//ncbi.nlm.nih.gov/BLAST) database. Searches were undertaken at the nucleotide level (BLASTn) to identify homologous sequences as well as identifying any potential contamination. The tBLASTx protein database translated the nucleotide sequence and was used to examine sequences at the protein level to determine any homology. Protein sequences were examined using BLASTP. Significant similar sequences and the associated BLAST 'scores' were recorded.

2.2.9.7.3. Computer Analysis of Sequences

Multiple sequences of whole genes and partial genes were aligned using ClustalW using the default settings (Thompson *et al.,* 1994). Phylogenetic analyses of the alignments were undertaken using Neighbour and Parsimony, in PHYLIP, version 3.68 (Felsenstein, 2005). The robustness of all alignments were tested with the bootstrapping option (SeqBoot). Genetic distances, applicable for distance matrix phylogenetic inference, were calculated using the Protdist program in the PHYLIP package. Phylogenetic inferences based on the distance matrix (Neighbour) algorithm was applied to the alignments. The best tree or majority rule consensus tree was selected using the consensus program (Consense). The trees were visualised and drawn using the TREEVIEW software version 3.5 (Page, 1996). Inteins were analysed using the NEB InBase database (http://www.neb.com/neb/inteins.html).

2.2.9.7.4. Genome Sequence Assembly and Finishing

Complete sequencing of the OtV-1 and OtV-2 genomes was performed at the Advanced Genomics Facility based at the University of Liverpool, UK, using a GS-FLX 454 genome sequencer. The resulting reads were then *de novo* assembled with the 454's own Newbler assembler software, version 1.1.03.24. Resulting contigs were next screened according to coverage depth to filter out the low coverage algal host contamination and bacterial contamination from the viral DNA. The remaining contigs were subsequently ordered and oriented using *Ostreococcus virus* OsV5 complete genome (NC_010191) as a reference sequence, using MUMmer 3.2. (Kurtz *et al.,* 2004). The OtV genomes were sequenced to an average of 10-fold coverage. Primers were designed to fill gaps between contigs, and the resultant Sanger sequence data was merged with the 454-generated contigs to form the completed genome sequences. Putative open reading frames (ORFs) were then identified, *de novo*, using glimmer3 using the provided g3-iterated.csh script (Delcher *et al.,* 2007). These ORFs were blasted against the reference sequence to identify those ORFs, which may have been split due to frameshift errors caused by the 454 sequencer. Where found, relevant putative coding sequence (CDS) regions were joined and annotated accordingly. Coding regions of less than 65 amino acids were excluded from subsequent analysis. All remaining coding regions were then preliminarily annotated after blast search against the reference to identify similarity. Putative tRNA genes were identified with aid of tRNAscan-SE version 1.23 (Lowe, 1997). Annotation was then checked and supplemented manually within the Artemis software tool (release 11) (Rutherford *et al.,* 2000).

2.2.9.7.5. Genome Annotation

Whole genome sequences of OtV1 and OtV2 were analysed and annotated using the software program Artemis (Rutherford *et al.,* 2000) with putative coding sequences (CDSs) being generated based on predicted ORFs, correlation scores for each potential reading frame, GC content and codon usage indices. Similarities of putative coding sequences were detected using BLAST and any homologous sequences were recorded. CDSs were assigned putative functions and were colour coded based on their function. Comparative genomic analysis was conducted between the genomic data obtained in this study and similar virus genomes in the database using the software program Artemis Comparison Tool (ACT) (Carver *et al.,* 2005).

Chapter 3

The detection and isolation of viruses infecting marine nano- and picophytoplankton

3.1. Introduction

Viruses are now widely accepted as the most abundant biological entities in natural waters and are an integral component of the marine ecosystem. The first evidence of infective algal viruses in seawater was reported in 1979, when a virus infecting the marine eukaryotic algal species *Micromonas pusilla* was successfully isolated (Mayer and Taylor, 1979). These researchers observed that *M. pusilla* cells, isolated during a study of nanoplankton in British Colombia coastal waters, declined rapidly in culture. Addition of filtrates generated from the medium of these cultures produced the lysis of healthy *M. pusilla* cells. Transmission electron microscopy (TEM) analysis of the lysate revealed polyhedral virus-like particles (VLPs) of 130 - 135 nm in diameter. This initial report on *M. pusilla* viruses was followed by the isolation and characterisation of a number of other marine eukaryotic algal viruses. Viruses infecting *M. pusilla* have been isolated and described since this initial report (Cottrell and Suttle, 1995; Sahlsten, 1998; Sahlsten and Karlson, 1998; Brussaard *et al.,* 2004; Zingone *et al.,* 2006).

Since the 1970s, the significance of viruses as agents of mortality for phytoplankton in aquatic systems has been highlighted. Viruses have been found in algal cells from all major taxonomic classes comprising some key species, such as *Synechococcus* sp. (Suttle and Chan, 1993; Wilson *et al.,* 1993), *Emiliania huxleyi* (Bratbak *et al.,* 1993), *Phaeocystis pouchetti* (Jacobsen *et al.,* 1996) and harmful species such as *Aureococcus anophagefferens* (Milligan and Cosper, 1994; Gastrich *et al.,* 1998) and *Heterosigma akashiwo* (Nagasaki and Yamaguchi, 1997). At present, a number of algal viruses have been isolated and characterised and are now widely accepted as

being diverse and dynamic within the microbial community (Cottrell and Suttle, 1991; Short and Suttle, 2002; Schroeder *et al.,* 2003; Brussaard, 2004b).

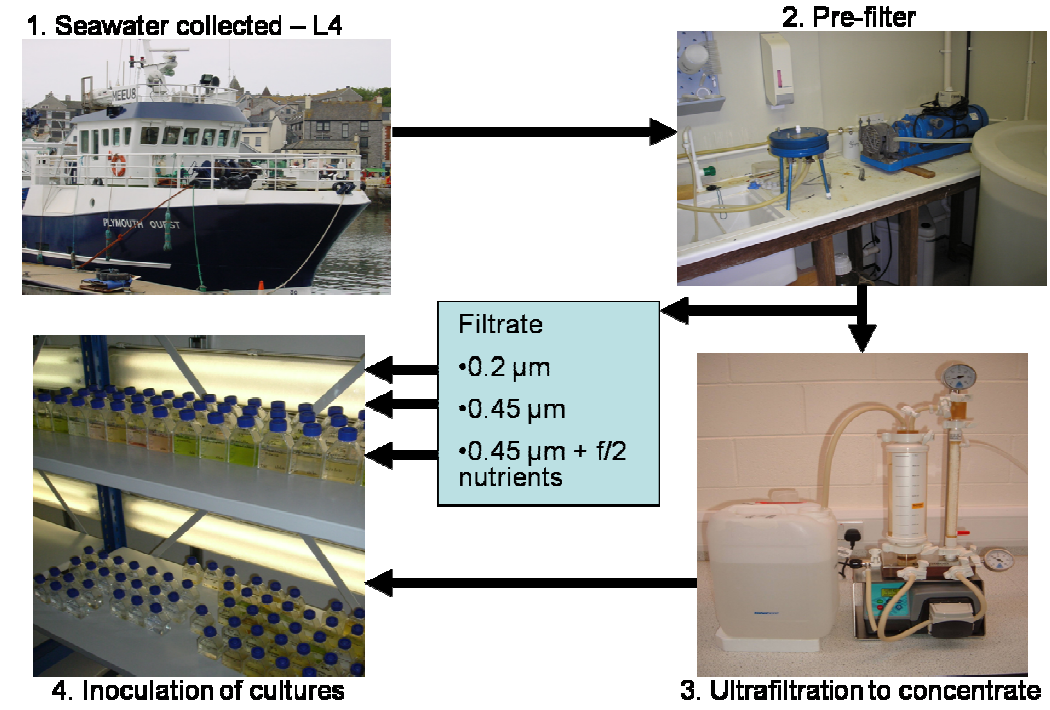
The isolation and characterisation of a virus that infects a *Chlorella*-like species led to the formation of the *Phycodnaviridae* family (Van Etten *et al.,* 1983b; Van Etten and Meints, 1999). *Paramecium bursaria Chlorella* virus (PBCV-1) is the prototype of this family of large viruses, which infect both freshwater and marine eukaryotic algae. Members and prospective members of the *Phycodnaviridae* are continually being discovered. They all have icosahedral morphology, an internal lipid membrane, and large dsDNA genomes of 157 to 560 kb (Van Etten *et al.,* 2002; Wilson *et al.,* 2005b; Schroeder *et al.,* 2009).

Availability of experimental host-virus systems is the key to understanding more about the features and ecological significance of algal viruses and to develop molecular methods that may be applied in field studies. Therefore, it has become necessary to bring more host-virus systems into the laboratory, where they can be characterised and described. Virus-like particles have now been observed in over 44 taxa of marine algae (reviewed by Van Etten *et al.,* 1991). However, only a small number have been isolated and even fewer brought into culture and characterised in detail. Those that have been adequately described have been putatively assigned to the family *Phycodnaviridae.* The taxonomic divisions within this family are supported by sequence analysis of DNA polymerase genes (Chen and Suttle 1995a; b; 1996).

The majority of algal viruses in culture have been isolated from unicellular host organisms (microalgae). All microalgal virus infections described to date are lytic.

Generally, the isolation of lytic algal viruses follows a relatively simple protocol. A prefiltered natural seawater sample is added directly or, following concentration, to one, or various strains, of the algal species of interest. A loss in chlorophyll autofluorescence is then monitored. Natural samples can be concentrated by ultrafiltration (Suttle *et al.,* 1991) and/or prefiltered, or centrifuged at low speed, to remove cellular material. A loss of autofluorescence indicates a loss of cells in comparison to a noninfected control culture (Suttle *et al.,* 1990). An overview of the various steps that can be employed in algal virus isolation, and were indeed used in this present study, is presented in Figure 3.1.

Screening with whole-water samples ensures no viruses are lost, as may occur during filtration or concentration. The use of whole-water samples has led to the isolation of a range of algal viruses (Nagasaki and Yamaguchi, 1997; Tarutani *et al.,* 2001; Castberg *et al.*, 2002). This method is the simplest and quickest approach to virus isolation. Large volumes of water are required, however, as no concentration steps are included.



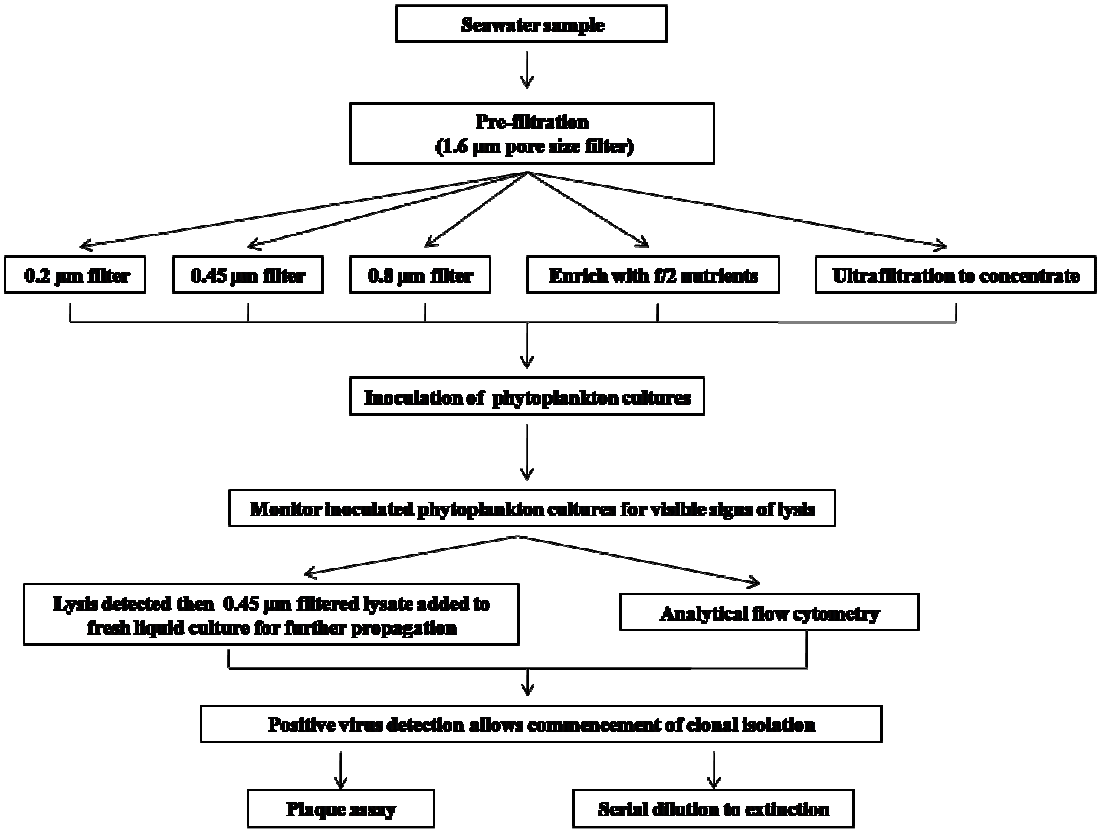


Figure 3.1. Flow diagram of procedure for detection and isolation of viruses infecting phytoplankton.

Although a number of viruses may be removed by prefiltration and concentration steps, the use of concentrated water samples greatly increases the detection limits of screening. Viruses have been concentrated using techniques such as tangential flow filtration, vortex flow filtration and ultracentrifugation (Paul *et al.,* 1991; Suttle *et al.,* 1991; Wommack *et al.,* 1995). Although large viruses are selectively removed using these techniques, the possibility of detecting less common viruses is enhanced. Serial dilutions of lysate are screened for the propagation of viruses. In theory, only one virus is required to cause the eventual lysis of an entire population, so the highest dilution that results in lysis will contain the one virus and this will give a clonal isolate. Clonal isolation is important as this ensures a single virus type is present. However, endpoint dilutions are not 100% accurate and the process must be repeated a minimum of three consecutive times to ensure the clonal status of the isolate. To obtain purified clonal cultures of infectious viruses, typically several cycles of a liquid serial dilution procedure are used. This process can be time-consuming and labour-intensive, although is helped if cultures can be grown in microtitre plates.

The most efficient method for obtaining virus clones is the use of plaque purification assays, since individual plaques represent the propagation of a single virus. Plaque assays were originally developed to count and measure infectivity of bacteriophages (Safferman and Morris, 1963). The plaque assay technique has since been adapted for the isolation of viruses infecting *Chlorella* species (Van Etten *et al.,* 1983b), cyanobacteria (Suttle and Chan, 1993; Wilson *et al.,* 1993) and *Emiliania huxleyi* (Wilson *et al.,* 2002b). The use of plaque formation in the isolation of viruses has advantages. If a single plaque is isolated it can be treated as having originated from a single virus particle. Therefore, the number of infectious particles can be calculated. The plaque assay method enables a discrete form of virus growth, which can be

distinguished from a background of uninfected cells. The host cells therefore support growth of the virus and provide a contrasting background against which virus growth can be recognised. The basis of the plaque assay technique used in the isolation and purification of viruses and to determine viral titres, is to measure the ability of a single infectious virus to form a plaque on a confluent monolayer of host cells. A plaque is formed as a result of infection of one cell by a single virus particle followed by the replication of that virus and eventually the death of the cell. The newly replicated cell then proceeds to infect and kill the surrounding cells forming a clear plaque on the lawn of host cells. The plaque can be visualised by the naked eye or light microscopy. Each plaque represents a single virus. Therefore, clonal virus populations can be purified by isolating individual plaques. Individual plaques obtained from varying dilutions of a viral stock can be counted to determine the viral titre.

Most algal viruses isolated to date infect hosts that generally do not grow, or grow poorly, on agar plates. Plaque assay purification methods, analogous to those used in the purification of bacteriophages, are therefore inapplicable in the isolation of viruses infecting such hosts. Instead viruses are isolated and purified in liquid cultures and many viruses, including those that infect dinoflagellates (Tarutani *et al.,* 2001), prymnesiophytes (Jacobsen *et al.,* 1996; Sandaa *et al.,* 2001) and raphidophytes (Nagasaki and Yamaguchi, 1997; Lawrence *et al.,* 2001), have been isolated using this technique. This end-point method gives an all-or-none response and the virus is diluted until only a few inoculants contain a single infective particle. To generate a pure viral stock by end-point dilution, the aim is to dilute the virus, such that, if multiple cultures are exposed to the diluted inoculants, any cultures that become infected will have received only a single infectious unit. To ensure this, the

virus should be diluted to the point where only 10% of the cultures become infected. The enumeration of infectious units may be achieved by adding serial dilutions of the inoculated culture to replicate cultures and monitoring for lysis at each dilution level. In this assay it must be assumed the lysed cultures contain at least one infectious virus particle. The number of viruses present in the dilution series is then estimated using statistical analysis.

Algal viruses are detected by techniques routinely used in virology, such as TEM, which can also provide information about morphology of the virus particles. Certain algal viruses can also be discriminated using analytical flow cytometry (AFC) (Brussaard *et al.,* 2000; 2004a). AFC is a high throughput method and can enable many of the fluorescent stained viruses to be distinguished from bacteriophages. The use of nucleic acid-specific stains can allow viruses to be enumerated by AFC (Brussaard, 2004a) and also epifluorescence microscopy (Hennes and Suttle, 1995; Noble and Fuhrman, 1998). AFC allows the discrimination of subpopulations, based on their fluorescence and scatter characteristics.

In this chapter, the screening of seawater from natural samples is described in order to isolate novel viruses infecting phytoplankton and to characterise them further. The host phytoplankton strains used to test for susceptibility to viruses were representative of the nano- and picophytoplankton assemblages. Picoeukaryotes (cells smaller than 3 µm in diameter) and nanophytoplankton (cells 3 µm to 20 µm in diameter) are widely distributed in aquatic environments. Their significant biomass and high productivity suggest that they play a major role in oceanic, coastal, and freshwater systems (Diez *et al.,* 2001; Massana *et al.,* 2002). A total of 12 viruses

were isolated in this study that infect the picoeukaryote species *Ostreococcus tauri* and *Micromonas pusilla.*

3.2. Materials and Methods

3.2.1. Algal cultures and growth conditions

The different unialgal phytoplankton strains and species representatives of different taxonomic groups (not axenic) used for virus isolation are listed in Table 2.1. All algal species, except *Micromonas pusilla* and *Ostreococcus tauri* strains, were cultured in f/2 medium (Guillard, 1975). The Mamiellaceae species were grown in K medium (Keller *et al.,* 1987) (Table 2.3). All cultures were incubated under the growth conditions outlined in Section 2.2.

3.2.2. Isolation and purification of viruses

Virus lysates were isolated from natural seawater originating from a sampling site, L4, in the Western English Channel according to the procedure described in Section 2.2.3. Briefly, filtered (GF/F, Whatman) seawater, was processed as described previously in Section 2.2.3. Processed filtrate (1 ml) was added to exponentially growing host phytoplankton strains in f/2 medium (Guillard *et al.,* 1975) or K medium (Keller, 1987) (50 ml). Once clearing of the host culture was observed, the lysate was passed through a 0.45 µm syringe filter (Gelman), after which the lysate was used to reinfect an exponentially growing algal host culture. After recurrent lysis and reinfection, viral isolates were made clonal by plaque purification assay or serial dilution to extinction three times, as described in Section 2.2.4.2.

3.2.3. Analytical Flow Cytometry

Analytical flow cytometry was used to monitor algal growth and detect the presence of virus populations in inoculated cultures and is described in Section 2.2.7.

3.3. Results

3.3.1. Algal cultures used for virus screening

At the outset of this study a total of 106 phytoplankton strains, representing species in the nano- and picophytoplankton size ranges (Table 2.1) were established in a culture collection. Healthy strains in the mid-exponential phase of growth were screened for the presence of viruses in natural seawater samples. The phytoplankton species used to screen for viruses were selected to represent a range of taxonomic classes. Representatives of certain genera were included, for which viruses have been described to infect in previous studies, e.g. *Heterosigma akashiwo* (Nagasaki and Yamaguchi, 1997), *Chyrsochromulina* sp. (Suttle and Chan, 1995) and *Pyramimonas* sp. (Sandaa *et al.,* 2001). Six months after the establishment of the culture collection, six strains of *Ostreooccus tauri* and one strain of *Micromonas pusilla* were added to the collection.

Seawater sampled in the Western English Channel coastal waters, during the period December 2005 to September 2007, was screened for the occurrence of viruses lytic to the representative strains of marine microalgae. The original culture collection of 106 phytoplankton strains was reduced to 38 strains after a period of 12 months, according to healthy algal growth status and the ongoing potential for successful virus detection within the constraints of this study. A comprehensive summary of the outcome for each algal strain is presented in Table 3.1. The culture collection was reduced in size due to a number of factors. These included poor growth of certain algal species, resulting in the death of a number of cultures (Table 3.1), despite subculturing efforts. Healthy growing cultures, which were used to screen for viruses but did not undergo lysis, were no longer maintained after several months of inoculation on a weekly basis. Cultures which had been seen to show visible signs of lysis were maintained for further rounds of inoculation. However, if re-infection

did not result in repeated lysis or virus isolation failed then these cultures were eventually no longer maintained (Table 3.1). Eventually, 18 months following the start of this study, this collection was scaled down further to a final maintained algal culture collection of six strains of *Ostreococcus tauri* and a single strain of *Micromonas pusilla* (Table 3.2).

Table 3.1..Summary of status of phytoplankton strains used in this study for virus screening and isolation.

Species	Culture code*	Growth medium	Status of culture
Ostreococcus tauri	RCC 745	K	Healthy. One virus strain isolated.
Ostreococcus tauri	RCC 501	K	Healthy. Two virus strains isolated.
Ostreococcus tauri	RCC 143	K	One virus strain isolated. Lost from culture due to poor growth Feb 2008
Ostreococcus tauri	RCC 393	K	Healthy. Two virus strains isolated.
Ostreococcus tauri	RCC 344	K	One virus strain isolated. Lost from culture due to poor growth May 2007.
Ostreococcus tauri	RCC 356	K	One virus strain isolated. Lost from culture due to poor growth August 2007.
Micromonas pusilla	RCC434	K	Healthy. Four virus strains isolated.
Corethron hystrix	PCC 615	f/2	Poor growth. Lost from culture after two subculturing rounds.
Odontella sinesis	PCC 624	f/2	Poor growth. Lost from culture after three subculturing rounds.
Stephanopyxis palmeriana	PCC 625	f/2	Poor growth. Lost from culture after two subculturing rounds.
Coscinodiscus wailesii	PCC 695	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture May 2006.
Asterionellopsis glacialis	PCC 607	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture July 2006.
Bacillaria paxillifera	PCC 665	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture Oct 2006.
Dunaliella primolecta	PCC 81	f/2	Healthy growth. No lysis seen following inoculation. Lost from culture March 2008.
Dunaliella minuta	PCC430	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture March 2008.
Dunaliella tertiolecta	PCC 83	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture March 2008.
Chlorella stigmatophora	PCC 85	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Chlorella salina	PCC 309	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Nannochloris atomus	PCC 583	f/2	Poor growth. Lost from culture after two subculturing rounds.
Chlamydomonas concordia	PCC 491	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Feb 2007.
Chlamydomonas reginae	PCC 399	f/2	Poor growth. Lost from culture after three subculturing rounds.
Tetraselmis striata	PCC 443	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture May 2007.
Tetraselmis impellucida	PCC 429	f/2	Poor growth. Lost from culture after two subculturing rounds.
Tetraselmis convolutae	PCC 372	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Feb 2007.

Species	Culture code	Medium	Status of culture
Tetraselmis suecica	PCC 305	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Feb 2007.
Tetraselmis marina	PCC 308	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Feb 2007.
Tetraselmis marina?	PCC 570	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Tetraselmis rubens	PCC 456	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Tetraselmis sp.	PCC 315	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Tetraselmis sp.	PCC 511a	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Tetraselmis sp.	PCC 512	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Tetraselmis sp.	PCC 513	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Pyramimonas amylifera	PCC 246	f/2	Poor growth. Lost from culture after three subculturing rounds.
Pyramimonas obovata	PCC 280	f/2	Poor growth. Lost from culture after two subculturing rounds.
Pyramimonas urceolata	PCC 299	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Nov 2006.
Pyramimonas parkeae	PCC 492	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Nov 2006.
Ochromonas sp.	PCC 2	f/2	Poor growth. Lost from culture after two subculturing rounds.
Ochromonas sp.	PCC 504	f/2	Poor growth. Lost from culture after two subculturing rounds.
Pseudopedinella sp.	PCC 361	f/2	Poor growth. Lost from culture after two subculturing rounds.
Pseudopedinella sp.	PCC 438	f/2	Poor growth. Lost from culture after three subculturing rounds.
Heterosigma akashiwo	PCC 12a	f/2	Poor growth. Lost from culture after two subculturing rounds.
Heterosigma akashiwo	PCC 239	f/2	Poor growth. Lost from culture after two subculturing rounds.
Heterosigma akashiwo	PCC 461	f/2	Poor growth. Lost from culture after two subculturing rounds.
Nannochloropsis oculata	PCC 663	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture May 2007.
Nannochloropsis salina	PCC 591	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Cryptomonas calceiformis	PCC 175	f/2	Poor growth. Lost from culture after two subculturing rounds.
Cryptomonas calceiformis	PCC 412	f/2	Poor growth. Lost from culture after three subculturing rounds.
Cryptomonas maculata	PCC 174	f/2	Poor growth. Lost from culture after two subculturing rounds.
Hemiselmis virescens	PCC 157	f/2	Poor growth. Lost from culture after two subculturing rounds.
Hemiselmis brunnescens	PCC 14	f/2	Poor growth. Lost from culture after three subculturing rounds.
Hemiselmis sp.	PCC 631	f/2	Poor growth. Lost from culture after two subculturing rounds.
Alexandrium tamrense	PCC 173	f/2	Poor growth. Lost from culture after three subculturing rounds.
Heterocapsa triquetra	PCC 169	f/2	Poor growth. Lost from culture after three subculturing rounds.

Species	Culture code	Medium	Status of culture
Amphidinium carterae	PCC 127	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Oct 2006.
Gymnodinium veneficum	PCC 103	f/2	Poor growth. Lost from culture after three subculturing rounds.
Gymnodinium simplex	PCC 368	f/2	Poor growth. Lost from culture after three subculturing rounds.
Eutreptiella gymnastica	PCC 462	f/2	Poor growth. Lost from culture after three subculturing rounds.
Isochrysis galbana	PCC 565	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture Jan 2008.
Isochrysis aff. galbana	PCC 8	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture June 2007.
Isochrysis sp.	PCC 240	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture June 2007.
Isochrysis sp.	PCC 506a	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture June 2007.
Isochrysis sp.	PCC 562	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture June 2007.
Imantonia rotunda	PCC 133	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Nov 2006.
Imantonia rotunda?	PCC 577	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture July 2007.
Ochrosphaera neapolitania	PCC 162	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Nov 2006.
Ochrosphaera?	PCC 341	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Oct 2006.
Ochrosphaera?	PCC 434(5)	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Oct 2006.
Coccolithus pelagicus	PCC 667	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Sept 2006.
Pleurochrysis elongata	PCC 535	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Oct 2006.
Pleurochrysis carterae	PCC 156	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture May 2007.
Pleurochrysis carterae	PCC 17	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture May 2007.
Pleurochrysis carterae?	PCC 378(1)	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Oct 2006.
Prymnesium patelliferum	PCC 527	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Prymnesium patelliferum	PCC 527a	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Ocy 2006.
Prymnesium patelliferum	PCC 527b	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2006.
Prymnesium sp.	PCC 569	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture July 2006.
Prymnesium sp.	PCC 593	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Sept 2006.
Prymnesium sp.	PCC 594	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Mar 2007.
Prymnesium sp.	PCC 596	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2007.
Prymnesium sp.	PCC 598	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2007.
Chrysotila stipitata	PCC 377	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful.

Species	Culture code	Medium	Status of culture
Chrysotila stipitata	PCC 432	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2007.
Chrysotila lamellosa	PCC 475	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful. Lost from culture June 2007.
Chrysotila sp.	PCC 509	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2007.
Chrysotila sp.	PCC 510a	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2007.
Chrysotila sp.?	PCC 257	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Jan 2007.
Hymenomonas globosa	PCC 536	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Phaeocystis globosa	PCC 64	f/2	Poor growth. Lost from culture after three subculturing rounds.
Phaeocystis globosa	PCC 147	f/2	Poor growth. Lost from culture after three subculturing rounds.
Phaeocystis globosa	PCC 540	f/2	Poor growth. Lost from culture after three subculturing rounds.
Geophyrocapsa oceanica	PCC 572	f/2	Poor growth. Lost from culture after three subculturing rounds.
Dicrateria inornata	PCC 564	f/2	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful.
Chrysochromulina sp.	PCC 307	f/2	Poor growth. Lost from culture after two subculturing rounds.
Chrysochromulina strobilus	PCC 43a	f/2	Poor growth. Lost from culture after two subculturing rounds.
Chrysochromulinaaff.polylepis	PCC 200	f/2	Poor growth. Lost from culture after two subculturing rounds.
Pavlova salina	PCC 154	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture April 2007.
Pavlova aff. lutheri	PCC 554	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture April 2007.
Pavlova lutheri	PCC 75	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture April 2007.
Pavlova pinguis	PCC 471	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Pavlova sp.	PCC 484	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Pavlova aff. salina	PCC 486	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Dec 2006.
Pavlova gyrans	PCC 93	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture April 2007.
Pavlova virescens	PCC 515	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture April 2007.
Apistonema sp.	PCC 508	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Feb 2007.
Cricosphaera?	PCC 351	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Nov 2006.
Unidentified prymnesiophyte	PCC 168	f/2	Poor growth. Lost from culture after two subculturing rounds.
Rhodosarus marinus	PCC 472	f/2	Healthy growth. No lysis seen following inoculations. Lost from culture Nov 2006.
Pelagomonas calceolata	RCC 100	K	Healthy growth. Lysis seen following inoculation but virus isolation unsuccessful
*PCC = Plymouth Culture	Collection	RCC =	Roscoff Culture Collection

Table 3.2. *Prasinophyceae* strains used for clonal virus isolation in this study.

Genus	Strain code[a]	Geographic origin	Depth (m)	Clade[b]
Ostreococcus	RCC 344	Atlantic Ocean off Morocco	5	A
Ostreococcus	RCC 356	English Channel	0-5	A
Ostreococcus	RCC 393	Mediterranean	90	B
Ostreococcus	RCC 143	Tropical Atlantic	120	B
Ostreococcus	RCC 745/OTH 95	Thau lagoon, France	0-5	C
Ostreococcus	RCC 501	Mediterranean	0-5	D
Micromonas	RCC 434	Mediterranean	0-5	-

[a] RCC corresponds to the strain number in the Roscoff Culture Collection
[b] Clade designations based on ITS sequence phylogeny (Rodriguez *et al.,* 2005)

3.3.2. Virus screening

Seawater for the screening of viruses in this study was collected on a weekly basis (conditions permitting) over a 20 month period from mid-December 2005 to the start of September 2007. The regular screening of seawater samples at mostly weekly intervals during a 20 month period in this study resulted in the positive detection of a number of viruses in natural waters. Seawater sampled from April to October 2006 and from March to August 2007 resulted in the positive screening for viruses infecting certain phytoplankton species (Table 3.3). No detectable signs of lysis were observed for seawater sampled over the late autumn and winter months. The clonal viruses in this study were isolated in spring to autumn months (Table 3.3) but not during winter months.

Of the 106 phytoplankton strains used to screen for viruses, clearing of inoculated liquid cultures was observed in a total of 10 host species, shown in Table 3.4. Lysis of the host cells was usually observed between 5-7 days following inoculation. The visible appearances of select examples of cleared cultures are illustrated in Figure 3.2. Lysis was only observed in liquid cultures inoculated with concentrated seawater.

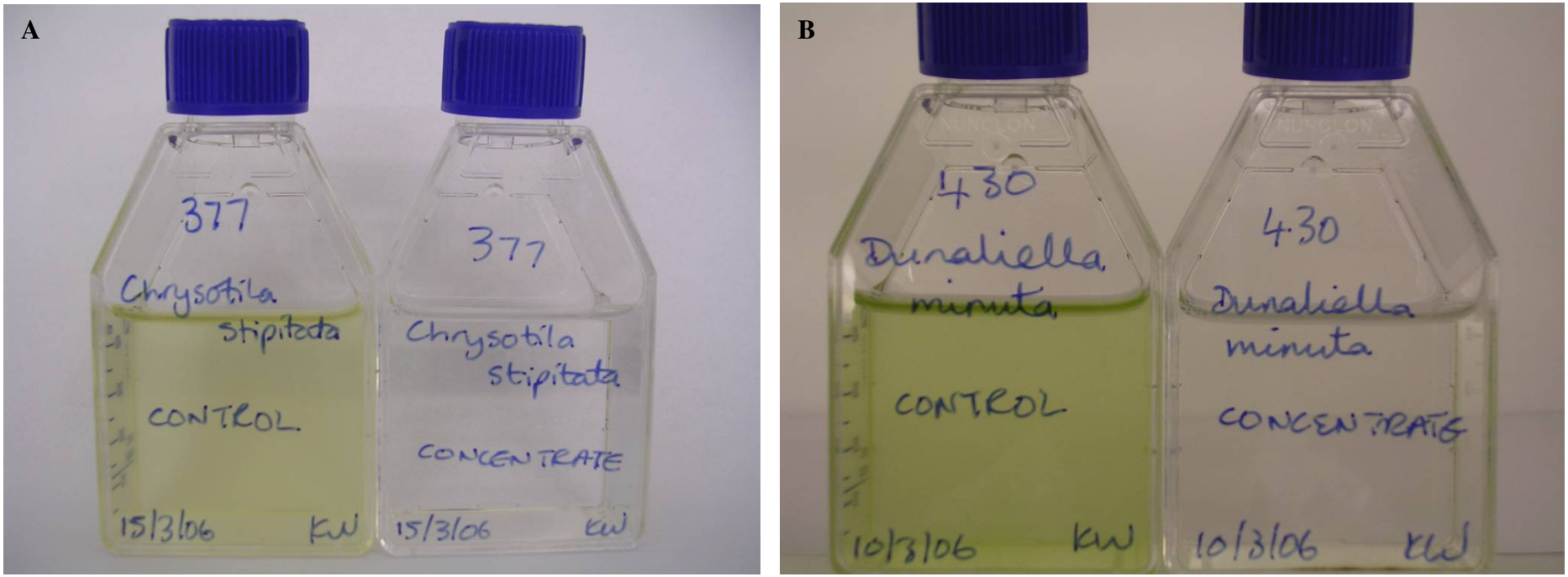


Figure 3.2. Clearing of liquid cultures inoculated with concentrated seawater, compared to non-inoculated controls.

A = *Chrysotila stipitata* B = *Dunaliella minuta*

.

Table 3.3. Phytoplankton strains where lysis was observed in this study, following inoculation with concentrated seawater samples. Positive AFC detection indicates a putative virus group was visible in inoculated cultures but absent in control noninoculated cultures.

Division	Class	Species	Positive AFC detection
Chlorophyta	Chlorophyceae	*Dunaliella minuta*	No
Chlorophyta	Chlorophyceae	*Dunaliella tertiolecta*	No
Haptophyta	*Prymnesiophyceae*	*Isochrysis galbana*	Yes
Haptophyta	*Prymnesiophyceae*	*Isochrysis sp.*	No
Haptophyta	*Prymnesiophyceae*	*Dicrateria inornata*	Yes
Haptophyta	*Prymnesiophyceae*	*Chrysotila lamellosa*	No
Haptophyta	*Prymnesiophyceae*	*Chrysotila stipitata*	Yes
Chlorophyta	*Prasinophyceae*	*Ostreococcus tauri* [a]	Yes
Chlorophyta	*Prasinophyceae*	*Micromonas pusilla*	Yes
Heterokontophyta	Pelagophyceae	*Pelagomonas calceolata*	No

[a] A detailed description of the different *Ostreococcus tauri* strains seen to lyse in this study are summarised in Table 3.2.

Table 3.4. Summary of lysis observations in ten Species strains of screened phytoplankton

L4 Sample Date	*Dunaliella minuta*	*Dunaliella tertiolecta*	*Isochrysis galbana*	*Isochrysis sp.*	*Dicrateria inornata*	*Chrysotila lamellosa*	*Chrysotila stipitata*	*Ostreococcus tauri*	*Micromonas pusilla*	*Pelagomonas calceolata*
14/12/2005		-	-		-	-	-			
19/12/2005		-					-			
03/01/2006		-				-		-	-	
09/01/2006	-	-	-	-				-	-	
16/01/2006	-	-	-	-	-	-	-	-		
23/01/2006	-	-	-	-	-	-		-	-	
06/02/2006	-	-	-	-	-		-	-	-	
14/02/2006	-	-	-	-		-	-	-	-	
22/02/2006	-	-	-	-	-	-	-	-	-	
28/02/2006	-	-	-	-	-	-		-	-	
06/03/2006	-	-	-	-	-	-	+	-	-	
16/03/2006	+	-	-	-	-	-	-	-	-	
21/03/2006	-	-	-	-	-	-	-	-	-	
31/03/2006	-	-	-	-	-	-	-	-	-	
03/04/2006	-	+	-	-	-	+	-	-	-	
11/04/2006	+-	-	-	-	-	-	-	-	-	
19/04/2006	-	-	-	-	-	-	-	-	-	-
26/04/2006	-	-	-	-	-	-	-	-	-	-
04/05/2006	-	-	-	-	-	-	-	-	-	
17/05/2006	-	-	-	-	-	-	-	-	-	
23/05/2006	-	-	-	-	-	-	-	-	-	
07/06/2006				-		-				-
13/06/2006	-	-	-	-	-	-	-	-	-	-
27/06/2006	-	-	+	-	+		-	-	-	
03/07/2006	-	-	-	+	-			-	-	+
12/07/2006	-	-	-	-	-	-	-	-	-	-

Table 3.4. Contd.

	Species screened									
L4 Sample Date	***Dunaliella minuta***	***Dunaliella tertiolecta***	***Isochrysis galbana***	***Isochrysis sp.***	***Dicrateria inornata***	***Chrysotila lamellosa***	***Chrysotila stipitata***	***Ostreococcus tauri***	***Micromonas pusilla***	***Pelagomonas calceolata***
17/07/2006	-	-	-	-	-			-	-	+
05/08/2006	-	-	-	-	-	-	-	-	+	
14/08/2006	-	-	-	-				+	-	-
22/08/2006	-	-		-	-	-		-	-	-
06/09/2006	-	-	-	-	-	-	-	+	-	-
12/09/2006	-	-	-	-	-			-	-	-
26/09/2006	-	-	-	-	-	-	-	-	-	-
04/10/2006	-	-	-		-			-	+	-
09/10/2006	-				-	-	-	-	+	-
31/10/2006	-				-			-	-	-
06/11/2006					-	-	-	-	+	
12/12/2006					-	-	-	-	-	
18/12/2006					-	-		-	-	
15/01/2007								-	-	
24/01/2007								-	-	
29/01/2007								-	-	
16/02/2007								-	-	
21/02/2007								-	-	
27/02/2007								-		
09/03/2007								-		
12/03/2007								+		
20/03/2007								-		
26/03/2007								+		
10/04/2007								+		
17/04/2007								-		
27/04/2007								-		

Table 3.4. Contd.

	Species screened									
L4 Sample Date	***Dunaliella minuta***	***Dunaliella tertiolecta***	***Isochrysis galbana***	***Isochrysis sp.***	***Dicrateria inornata***	***Chrysotila lamellosa***	***Chrysotila stipitata***	***Ostreococcus tauri***	***Micromonas pusilla***	***Pelagomonas calceolata***
30/04/2007								-		
08/05/2007								+		
30/05/2007								-		
04/06/2007								+		
19/06/2007								-		
25/06/2007								-		
03/07/2007								-		
10/07/2007								-		
16/07/2007								-		
23/07/2007								-		
20/08/2007								+		
28/08/2007								-		

3.3.2.1. Analytical flow cytometry analysis

Following primary clearing of liquid cultures and, in order to detect the presence of virus populations, lysed cultures were passed through a 0.45 µm filter and prepared for analysis in the analytical flow cytometer (AFC). AFC analysis of samples diluted in TE buffer and stained with DNA-binding dye SYBR Green I revealed the presence of bacteria in all samples and was used to aid confirmation of presence or absence of virus-like particles in inoculated cultures. These groups were distinguished based on their side scatter (SSC) and green fluorescence (GFL) characteristics. Possible virus-like particles were detected in inoculated cultures of the following species: *Ostreococcus tauri; Micromonas pusilla; Chrysotila stipitata; Dicrateria inornata* and *Isochrysis galbana*. (Table 3.3; Figs. 3.3., 3.4. and 3.5).

The virus group seen for *I. galbana* (Fig. 3.3.) had a substantially higher GFL value than those seen for some of the other virus lysates such as the OtV and MpV lysate AFC signatures (Figs. 3.4. and 3.5.). AFC analysis of the *D. inornata* lysate revealed the presence of two possible virus populations (Fig. 3.3). The AFC signature of the *C. stipitata* lysate revealed a very small population of virus-like particles (Fig. 3.3). Attempts to re-infect and propagate the *I. galbana*-specific virus, the *D. inornata*-specific virus and the *C. stipitata* virus failed and repeated lysis of the host cultures was not seen after the initial round of infection. The reasons for this are not known but one explanation may be due to resistance arising in the host cultures following exposure to the viruses.

The AFC signatures of the OtV and MpV viruses are shown in Figures 3.4. and 3.5. In the control non-inoculated cultures of both species, host cells are clearly visible but are absent from infected culture profiles. AFC analysis of the lysates showed signatures

with similar characteristics between *O. tauri* lysates and likewise between *M. pusilla* lysates (Figs. 3.4. and 3.5.).

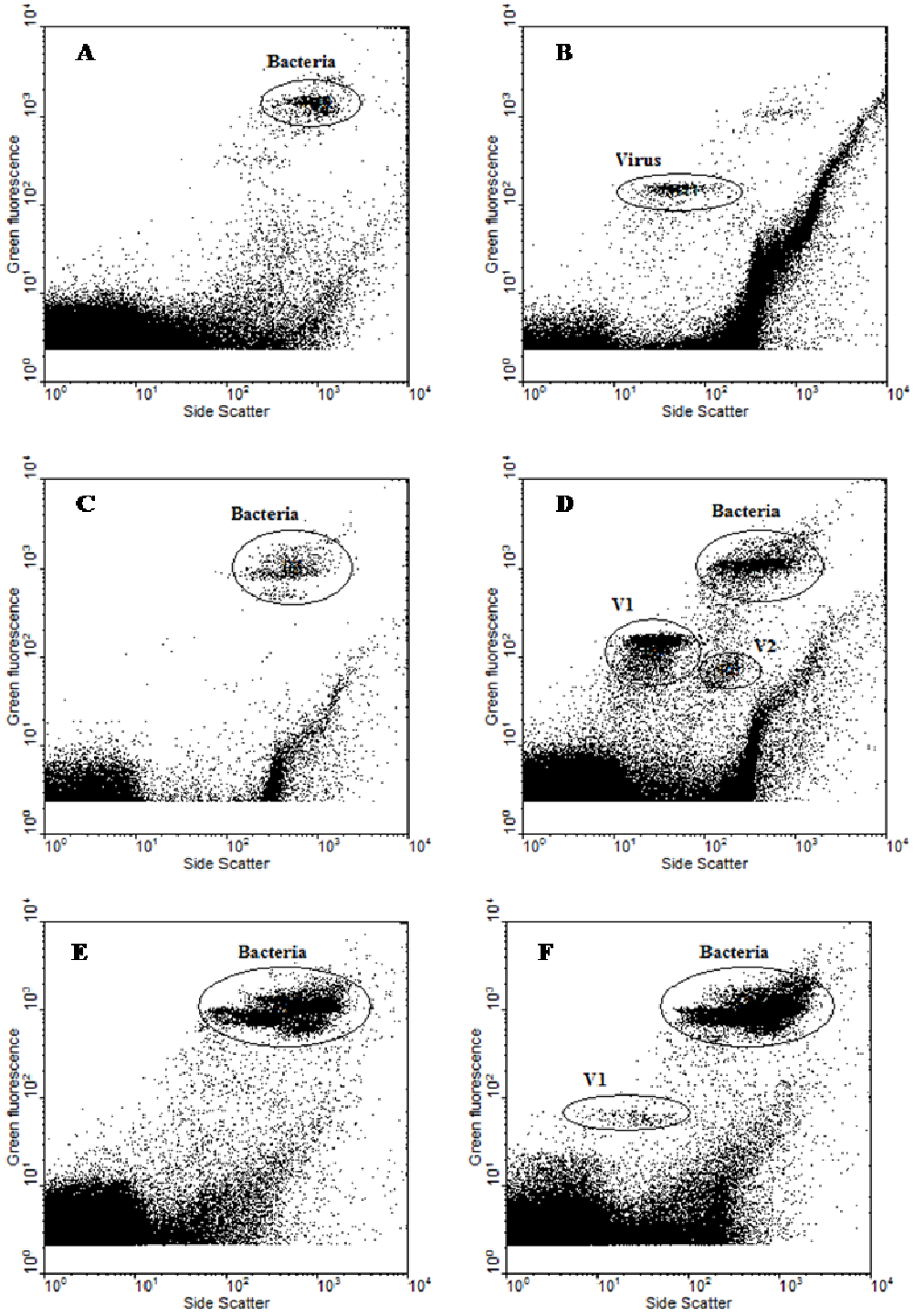


Figure 3.3. Representative biparametric flow cytometry plots showing populations of bacteria and possible virus population (labelled virus, V1, V2) in three algal species. The cultures analysed were unfiltered. A = control culture of *Isochrysis galbana*; B = culture of Isochrysis galbana inoculated with concentrated seawater; C = culture of *Dicrateria inornata* ; inoculated with concentrated seawater; D = culture of *Dicrateria inornata* inoculated with concentrated seawater; E = control culture of *Chrysotila stipitata*; F = culture of *Chrysotila stipitata* inoculated with concentrated seawater.

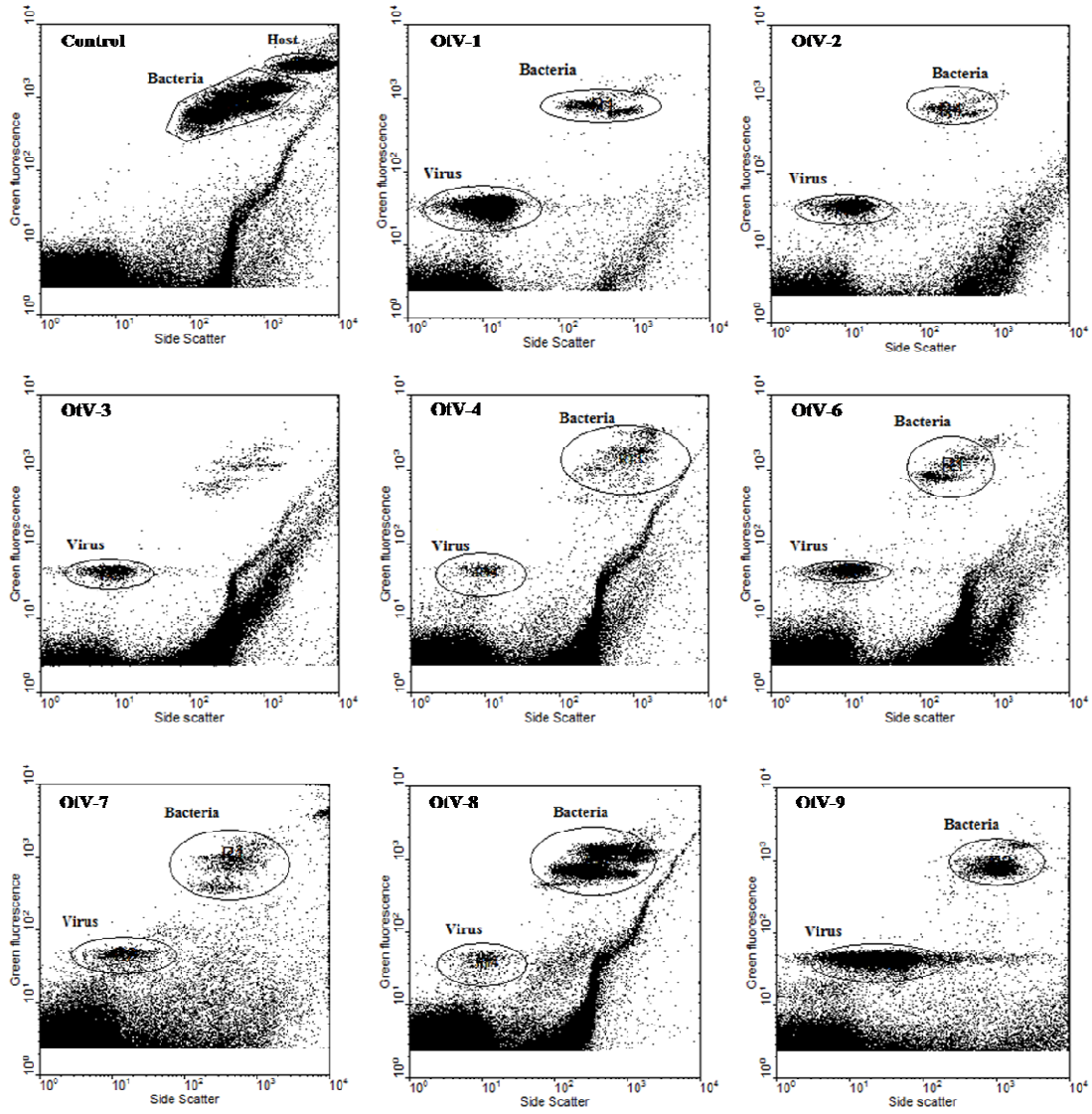


Figure 3.4. Representative biparametric flow cytometry plots showing populations of host algal cells of *Ostreococcus tauri*, bacteria and 8 strains of *Ostreococcus tauri*-specific viruses (OtVs), referred to as OtV-1, OtV-2, etc. The control plot is of a healthy non-inoculated control sample of *O. tauri.*

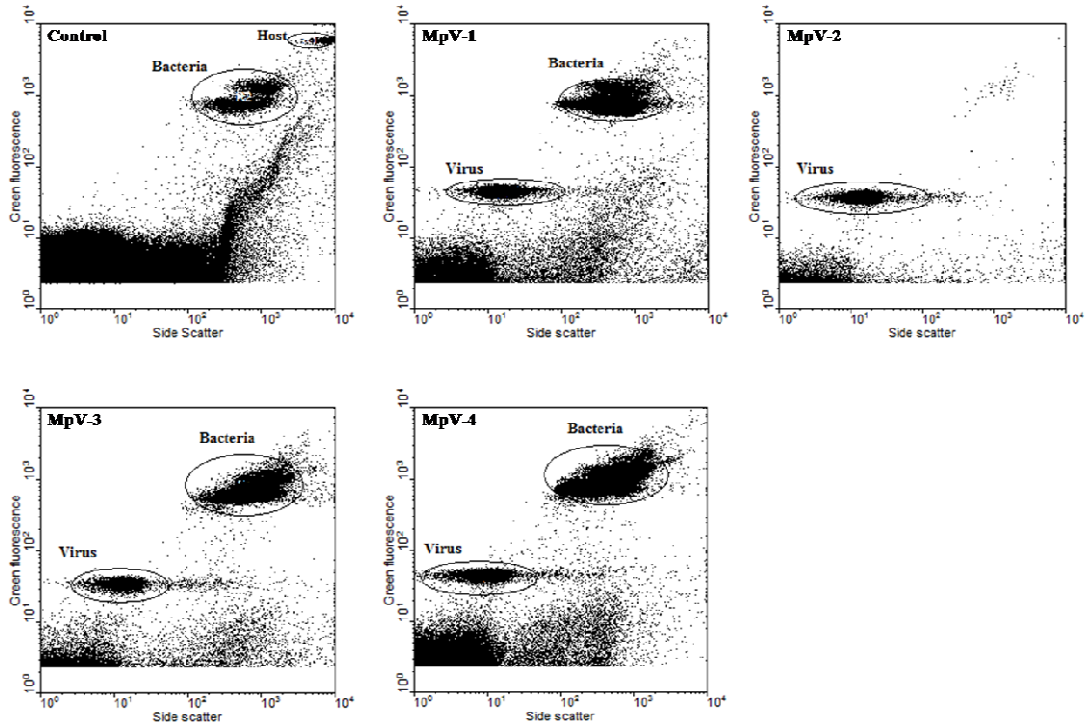


Figure 3.5. Representative biparametric flow cytometry plots showing populations of host algal cells of *Micromonas pusilla*, bacteria and 4 strains of *M.pusilla*-specific viruses (MpVs), referred to as MpV-1, MpV-2, etc. The control plot is of a healthy non-inoculated control sample of *M. pusilla.*

3.3.3. Purification of virus isolates

Following the wide screening of cultures and the subsequent detection of putative virus populations using analytical flow cytometry (AFC), steps were taken, as described in Section 2.2.4., to obtain clonal viruses. Of the ten phytoplankton strains shown in Table 3.3, clonal viruses were successfully isolated that infect two of the ten species, namely *Ostreococcus tauri* and *Micromonas pusilla*. However, repeated inoculations of liquid cultures of *C. stipitata, D. inornata* and *I. galbana* following signs of lysis in initial experiments, failed to reproduce visible signs of clearing in subsequent cultures. Following the successful isolation of viruses in the initial round of screening, a range of virus isolates that infect *O. tauri* and *M. pusilla* was successfully isolated and purified from the seawater samples. In total, four virus isolates were obtained that infect a single strain of *M. pusilla* and eight viruses were isolated that infect six strains of *O. tauri* (Table 3.5).

Table 3.5. Details of viruses isolated in this study that infect members of the *Prasinophyceae.*

Host species	Host strain	Date seawater sampled	Virus identity	Purification method[a]
Ostreococcus tauri	OTH 95	06/09/2006	OtV-1	Plaque assay
Ostreococcus tauri	RCC 393	07/05/2007	OtV-2	S.D.
Ostreococcus tauri	RCC 344	14/08/2006	OtV-3	S.D.
Ostreococcus tauri	RCC 356	12/03/2007	OtV-4	S.D.
Ostreococcus tauri	RCC501	26/03/2007	OtV-6	S.D.
Ostreococcus tauri	RCC 501	02/04/2007	OtV-7	S.D.
Ostreococcus tauri	RCC393	04/06/2007	OtV-8	S.D.
Ostreococcus tauri	RCC 143	07/08/2007	OtV-9	S.D.
Micromonas pusilla	RCC 434	08/09/2006	MpV-1	S.D.
Micromonas pusilla	RCC 434	14/09/2006	MpV-2	S.D.
Micromonas pusilla	RCC 434	22/09/2006	MpV-3	S.D.
Micromonas pusilla	RCC 434	16/10//2006	MpV-4	S.D.

N.B. All seawater was sampled from L4 sampling station in Western English Channel.
[a] S.D. = Serial dilution to extinction, described in Section 2.2.4.2.2.

A plaque purification method was successful for the virus, OtV-1 (Fig. 3.6). All other OtV isolates and all MPV isolates were purified using a serial dilution to extinction assay, as described in Section 2.2.4.2.2. Visual inspection of plates of *O. tauri* during the plaque purification procedure revealed that the host grew uninhibited on plates turning a brown-green colour after 5-7 days. Circular plaques approximately 1-2mm in diameter were observed after 4-5 days incubation for OtV-1 (Fig. 3.6). A 1 mm plaque was shown to contain approximately 10^4 plaque forming units (PFU) following further propagation.

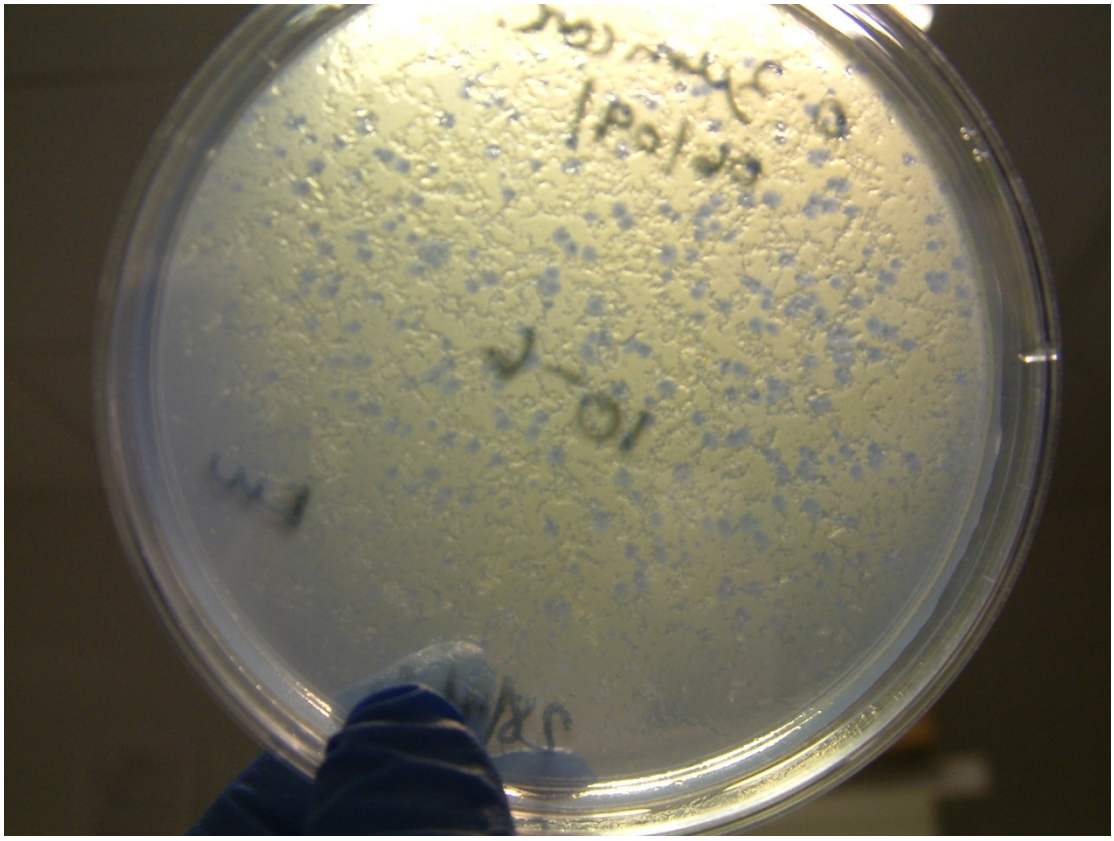


Figure 3.6. Plaques formed on a lawn of *Ostreococcus tauri* strain OTH 95 following inoculation with OtV-1.

3.4. Discussion

Although clearing of liquid cultures of a number of the phytoplankton species in this study was observed, subsequent attempts to repeat lysis of cultures were not always successful. The use of techniques such as AFC and TEM are not the definitive final proof of the presence of a virus. Confirmation that an infectious agent is present usually follows the fulfilment of Koch's postulates. To fulfil Koch's postulates, the infectious agent must be filterable, transferable and induce the same disease symptoms, following propagation in a healthy host. As lysis was not seen following transfer of inocula to a freshly growing host in the majority of the species studied here, the nature of the pathogen cannot be confirmed as viral. As cultures used in the study were not axenic, it is possible the algicidal activity observed was of a bacterial origin. However, the isolated clonal viruses that infected *O. tauri* and *M. pusilla* in this study successfully fulfilled Koch's postulates. As these viruses were filterable through a 0.2 µm filter, it was possible to reduce the presence of bacterial contamination, although complete elimination could not be guaranteed through filtering alone.

Following the initial success in the wide screen for suitable algal species, the *O. tauri*-specific viruses and *M. pusilla*-specific viruses were relatively straightforward to isolate and susceptible host strains usually lysed between 2 to 5 days after the addition of filtered concentrated seawater. Similar methods have been used to increase the natural virus concentration prior to virus isolation (Suttle *et al.,* 1991; Milligan and Cosper, 1994). The concentration of large volumes (20 litres) of seawater prior to inoculation proved to be the most successful method to increase the success of viral infection. The employment of this approach can concentrate viruses from seawater at efficiencies approaching 100% (Suttle *et al.,* 1991) and it has proven to be effective in the isolation of viruses lytic to microalgae in this study.

Viruses infecting the prasinophyte species *O. tauri* have been reported, with VLPs previously observed in transient blooms of *O. tauri* (O'Kelly *et al.,* 2003). Indeed, as this thesis was being prepared, a virus strain referred to as OtV-5 has recently been described (Derelle *et al.,* 2008). The OtV-5 virus was isolated from water sampled from Bages Lagoon in the Mediterranean Sea, in early 2006. The OtV5 virus was isolated using a modified plaque purification technique (Derelle *et al*., 2008). The genome of this virus has been sequenced and annotated (Derelle *et al.,* 2008). Following the isolation and characterisation of a range of viruses infecting *O. tauri* during this study, it will be of great interest to compare viruses from such different geographical origins. The collection of OtV strains isolated in UK coastal waters and described in this study can now undergo comparative analysis with OtV strains isolated in Mediterranean coastal waters. Likewise, the MpV isolates in this study could also be compared to viruses that infect *M. pusilla* in previous reports (Chen *et al.,* 1996; Zingone *et al.,* 2006).

The strain OtV-1 isolated in this study from seawater sampled at the L4 station was seen to infect the same strain (OTH 95) of *O. tauri* as the OtV-5 strain isolated from Mediterranean coastal waters (Derelle *et al.,* 2008). Direct comparisons between these two characterised viruses should highlight the degree of diversity between the geographically distinct isolates. A previous study comparing MpV virus diversity reported that there was no significant relationship between host susceptibility and geographical origin of host strains or with season of collection of host and viruses (Zingone *et al.,* 2006). Other studies found host strains were only susceptible to viruses from the same geographical locations as the host (Sahlsten, 1998; Lawrence *et al.,* 2001; Tai *et al.,* 2003). It is of note that in this present study, viruses were isolated from

seawater sampled from surface waters were seen to infect both high-light adapted and low-light adapted host strains in the case of the *O. tauri* host strains. It is postulated here that a degree of mixing of the water column results in the occurrence at various depths, including surface waters, of viruses that infect low-light adapted host species.

To date, most algal viruses described have very narrow host ranges with some exhibiting not only species-restricted host range but a strain-specific spectrum of infection. The fact that two isolates, OtV-2 and OtV-6, are seen to infect the same host strain, RCC 393, will require further detailed examination of their commonalities and differences. It is also possible that OtV-2 and OtV-6 are in fact the same virus strain and this would require further investigation, beyond analysis of the highly conserved DNA polymerase gene. Structural capsid protein composition, for example, will be a good indicator of how these two virus isolates can infect a common host. The proteins on the surface of free viruses serve as a means of virus attachment to receptors on the host cell surface allowing transfer of virus genomes into the cell. Any changes in the composition of these structural proteins between viruses may affect the binding to host receptors and consequently affect host range restriction (Tarutani *et al.,* 2006; Mizumoto *et al.,* 2008)

This investigation is not a quantitative study but rather a qualitative approach to virus isolation. The analysis of associated metadata e.g. sea temperature, UV irradiance etc. for the L4 sampling station are not included in this study, as this was not part of the aims of the study. The fact that the viruses in this study were isolated from seawater sampled during spring and summer months is most likely due to the increase of phytoplankton biomass and thus the increased likelihood of the presence of host species

in higher concentrations than during winter months. It is of interest that positive detection of viruses, particularly those seen to infect *O. tauri* and *M. pusilla,* led to an indication of possible temporal variation in viral abundance and infectivity at the L4 sampling station. A previous study investigated the temporal distribution of planktonic organisms at the L4 sampling station and reported a gradual increase in virus biomass from the winter to summer months (Rodriguez *et al.,* 2000). Viruses were reported to be significantly correlated with temperature and total phytoplankton biomass, reaching greatest abundances in the month of June. Previous studies conducted in Mediterranean coastal waters measured the abundances of *M. pusilla* and co-occurring viruses across seasons and sampling stations (Zingone *et al.,* 1999). *M. pusilla* was seen to be present from autumn to late spring although abundance declined to low levels and remained largely undetectable throughout the summer. Viruses infecting *M. pusilla* were seen to increase in density following a peak in host numbers (Zingone *et al.,* 1999). MpVs have been reported in varying concentrations in different geographical coastal locations, including sites along the US coast (Cottrell and Suttle, 1991, 1995) and the Skagerrak-Kattegat coastal areas, where concentrations were reported to be on average 2.5×10^5 viruses ml^{-1} (Sahlsten, 1998).

The wide range of virus titers observed in previous reports and the pattern seen in this study may be a reflection of spatial and temporal variation in phytoplankton host abundances. Seasonal variations resulting in lower temperatures in the winter and shorter daylight hours compared to the spring and summer months, result in reduced nutrient availability and lower productivity in the winter and, therefore, a subsequent drop in host abundances. The L4 sampling station is characterised by large seasonal variation in meteorological conditions and exhibits a diverse succession of

phytoplankton blooms throughout the spring, summer and autumn blooms (Holligan and Harbour, 1977). During the summer months, thermohaline stratification occurs at the L4 sampling station. A previous report recorded up to 50% of host *Emiliania huxleyi* cells were infected by viruses during a bloom confined by stratified waters (Brussaard *et al.,* 1996b). Seasonal shifts in water column stability and water temperature will cause fluctuations in phytoplankton community composition (Pingree *et al.,* 1978). During the winter months the environment in the Western English Channel is undoubtedly far more dynamic, experiencing unstable meteorological conditions, as well as a reduction in sea temperatures. Such variations will inevitably have consequences for virus abundances too. Elevated UV irradiance in coastal surface waters during spring and summer months could potentially cause damage to viruses, reducing their infectivity. However, previous studies have outlined how light-dependent repair, possibly by photoreactivation, may help to maintain high levels of viruses in surface waters (Weinbauer *et al.,* 1997; Wilhelm *et al.,* 1998).

One possible reason for the success in finding viruses lytic to *O.tauri* and *M. pusilla* may be that these prasinophyte species are ubiquitous with a widespread distribution in marine waters (Throndsen, 1976; Throndsen and Kristiansen, 1991; Guillou *et al.,* 2004; Zhu *et al.,* 2005; Countway and Caron, 2006; Fuller *et al.,* 2006b; Worden, 2006; Li *et al.,* 2008).

The observed lack of susceptibility to viral infection in a number of the original 106 phytoplankton strains in this study may be attributed to one of several possibilities. The method used in this study for isolating algal viruses is only selective for lytic viruses that infect algae kept in culture under laboratory conditions. This study is focussed on

lytic viruses and any existence of lysogenic infection was not investigated here. It is possible that such infections may have occurred following inoculation but in the absence of a lytic infection stage, such phenomenon would remain undetected. However, it cannot be discounted that the negative results in detecting lytic infections in the majority of the 106 cultures inoculated could have been due to the presence of lysogenic infection in the algae. To date, all viruses described that infect microalgae have been lytic. Lysogenic infection in microalgae is yet to be demonstrably proven. The presence of RNA viruses was not reported here but again this cannot be excluded as a possibility and simply remained undetected. Phytoplankton species infected by RNA viruses include the raphidophytes *Heterocapsa circularisquama* (Nagasaki *et al.,* 2005b; Nagasaki *et al.,* 2006; Mizumoto *et al.,* 2008) and *Heterosigma akashiwo* (Tai *et al.,* 2003) and certain diatom species (Nagasaki *et al.,* 2004; Shirai *et al.,* 2008). Indeed, a double-stranded RNA virus has been isolated and characterised that infects a strain of the prasinophyte species *Micromonas pusilla* (Brussaard *et al.,* 2004). The characteristics and features of RNA viruses that infect phytoplankton species may have a lower rate of detection and isolation success. However, the few reports of their existence indicates this is an area of algal virology that demands further investigation and no doubt this will result in several exciting discoveries and insights into viral ecology.

3.5. Conclusions

A phytoplankton culture collection of 106 species, representing a broad range of nano- and picophytoplankton, was established in order to conduct an exhaustive screening of seawater, sampled on a weekly basis over two years, for viruses. Lysis of liquid cultures inoculated with concentrated seawater was observed in ten phytoplankton species. Although flow cytometric analysis indicated a virus may have been responsible

for the lysis seen in a number of these ten species, further attempts to isolate and purify the viral agent for some cultures proved unsuccessful. From the ten species that were seen to lyse, a final total of twelve virus strains were isolated, which infect the prasinophytes *Ostreococcus tauri* and *Micromonas pusilla.* A seasonal variation was observed, with viruses most readily isolated from seawater sampled in summer and winter months. This may be due to an increase in host abundance as sea temperatures increase. This study has demonstrated that viruses infecting the prasinophytes *Ostreococcus tauri* and *Micromonas pusilla* can be readily isolated from natural seawater sampled from coastal areas.

Chapter 4

Characterisation of viruses infecting

Ostreococcus tauri and *Micromonas pusilla*

4.1. Introduction

This chapter focuses on viruses that infect two species of the order Mamiellales, namely *Ostreococcus tauri* and *Micromonas pusilla*. Particular attention is given to a virus that infects a high-light adapted strain of *O. tauri* (OTH 95), a virus that infects a low-light adapted strain of *O.tauri* (RCC 393) and four viruses that infect a strain of *M. pusilla* (RCC 474) and these are characterised and described in detail.

The development of molecular probes coupled with techniques such as fluorescent in situ hybridisation (FISH), have established that prasinophytes of the order Mamiellales, particularly the species *Micromonas pusilla* (Butcher, 1952) and *Ostreococcus tauri* (Chretiennotdinet *et al.,* 1995), dominate picoeukaryote communities in different oceanic regions (Not *et al.,* 2002; 2004; Romari and Vaulot, 2004; Worden *et al.,* 2004; Zhu *et al.,* 2005; Fuller *et al.,* 2006b). The *Prasinophyceae* are an early diverging class, branching at the base of the extant green plant lineage. Members of the *Prasinophyceae* are often minute in cell size and morphologically simple. This is best exemplified by the prasinophyte genus *Ostreococcus*.

Discovered in 1994 (Chretiennotdinet *et al.,* 1995), *Ostreococcus tauri* is the smallest (less than 1 µm in diameter), free-living eukaryote described to date. It is a naked, non-flagellated coccoid cell with a minimal cellular organisation consisting of a single chloroplast, single mitochondrion, single Golgi body, and a relatively small genome (~ 12.5 Mbp) (Derelle *et al.,* 2006). From an evolutionary perspective *O. tauri* is a primitive or ancestral chlorophyte. *O. tauri* can be considered an ideal model system especially from a physiological and ecological viewpoint. The genus *Ostreococcus* has a cosmopolitan distribution with several strains having been detected in coastal and oligotrophic North Atlantic waters and the Mediterranean, Indian and Pacific Oceans

(Worden *et al.,* 2004; Zhu *et al.,* 2005; Countway and Caron, 2006; Piganeau and Moreau, 2007).

Compared to the ecotypic differentiation and adaptive success displayed by *Prochlorococcus* (the most abundant marine prokaryotic picophytoplankter) (Moore, 1998), different *Ostreococcus* ecotypes from surface or deep waters provide evidence of niche adaptation (Rodriguez *et al.,* 2005). Low-light and high-light adapted *Ostreococcus* strains exhibit differences in their growth characteristics under various light regimes. The *Ostreococcus* genus has been shown to include distinct genotypes adapted to either high-light or low-light environments, thus providing the first evidence of ecotypic differentiation of light adaptation in eukaryotic picophytoplankton (Rodriguez *et al.,* 2005). Based on 18S rRNA phylogeny, the *Ostreococcus* strains currently maintained in culture are dispersed in four distinct phylogenetic clades (Guillou *et al.,* 2004; Rodriguez *et al.,* 2005). This distinction is in accordance with their photophysiological and ecological origins. For example, strains belonging to clade B have been isolated at the bottom of the euphotic zone and are unable to grow under high-light conditions. They also contain the unusual chlorophyll CS-170 (Rodriguez *et al.,* 2005).

Another important and ubiquitous member of the *Prasinophyceae* is *Micromonas pusilla,* identified as a major component of the picoplankton community in several oceanic and coastal regions including the Mediterranean Sea (Zingone *et al* 1999), the Norwegian Sea, central Californian waters and the English Channel (Throndsen and Kristiansen, 1991; Not *et al.,* 2004). One of the earliest discoveries and characterisation of a virus infecting a marine microalga was in the picoplanktonic prasinophyte

Micromonas pusilla (Mayer and Taylor, 1979). Since this initial report, *Micromonas pusilla* viruses (MpVs) have subsequently been detected at various geographical sites (Cottrell and Suttle, 1991; Sahlsten and Karlson, 1998; Zingone *et al.,* 1999). MpVs have been characterised under both laboratory conditions (Waters & Chan 1982) and the natural environment (Cottrell and Suttle, 1991, 1995). However, still relatively little is known about various aspects of these MpVs, including the role of viral infection on the abundance and distribution of the host species. High viral infection rates (estimated to be between 10% to 34%, using a modified dilution method) were reported in a *Micromonas* population (Evans *et al.,* 2003) and mortality due to viral lysis, coupled with grazing, exceeded estimates of the potential production of *Micromonas* sp. population and led to a decline in the host population.

The viruses isolated in this study seen to infect *O. tauri* and *M. pusilla* are characterised in terms of their morphology, phylogeny, growth kinetics and genome size.

4.2. Materials and Methods

4.2.1. Infection Kinetics

Infection kinetics of each virus isolate was determined by adding dilutions of the virus to exponentially growing host strains and following their growth compared to non-infected cultures, inoculated with autoclaved filtered seawater which served as virus-free negative controls. After infection, a 1 ml aliquot of cell suspension was collected from each culture at the same time point, once a day for 5 - 7 days post-infection, to enumerate host cells, and a 1 ml aliquot was fixed in 0.5% (v/v) glutaraldehyde for enumeration of viruses (Section 2.2.7). Phytoplankton and virus abundances were enumerated using analytical flow cytometry (Section 2.2.7).

Infection kinetics of virus isolates were determined by adding the virus at a MOI of one to exponentially growing (~2 x 10^6 cells ml^{-1}) cultures of host phytoplankton strains and following their growth compared to non-infected controls. Host cells in the cultures were enumerated in a FACScan analytical flow cytometer (Beckton Dickinson) (Section 2.2.7). Growth of the host cultures was monitored over a 7 - 10 day period.

4.2.2. Transmission Electron Microscopy (TEM)

The OtV and MpV isolates from this study were examined by TEM (Section 2.2.8). Briefly, viruses were either viewed by mounting whole samples onto electron microscope nickel grids or prepared by ultrathin sectioning techniques and mounted onto grids. All grids were stained with 2% (w/v) uranyl acetate and viewed in a JEOL 100S transmission electron microscope.

4.2.3. Pulsed Field Gel electrophoresis (PFGE)

PFGE was carried out on clonal viruses isolated during this study as described in Section 2.2.9.6.

4.2.4. Host range analysis

Host range determination was conducted on six different strains of *O. tauri*, one strain of *M pusilla*, three *Synechococcus* strains, two *Prochlorococcus* strains and one freshwater *Chlamydomonas reinhardtii* strain, using 8 OtV strains and 4 MpV strains (Section 2.2.6).

4.2.5. Molecular analysis and phylogeny

Nucleic acids were extracted as described in Section 2.2.9.2. Amplification of the DNA *pol* gene was then conducted by PCR using the AVS1 and AVS2 primer pair (Table 2.9) as described in Section 2.2.9.4. Amplified *pol* fragments were then cloned and sequenced as described in sections 2.2.9.5 and 2.2.9.7. Phylogenetic trees were generated as described in Section 2.2.9.7.3.

The DNA polymerase sequences used for phylogenetic analysis are described in Table 4.1 (scientific name, with abbreviation in parentheses, followed by the database accession number [referring to the National Centre for Biotechnology Information database unless otherwise stated]). The DNA polymerase sequences of the viruses OtV-2, OtV-3, OtV-4, OtV-6, MpV-1, MpV-2, MpV-3 and MpV-4 were obtained in this study.

Table 4.1. Details of DNA polymerase sequences used for phylogenetic analysis

Scientific name	**Accession number**
Alcephaline herpesvirus	NC 002531
Autographa californica nuclear polyhedrosis baculovirus (AcNPV)	P18131
Bombyx mori nucleopolyhedrovirus (BmNPV)	NP047469
Chilo iridescent virus (CIV)	CAC19195
Chrysochromulina brevifilum virus PW1 (CbV-PW1)	AAB49739
Chrysochromulina ericina (CeV)	ABU23716
Ectocarpus siliculosus virus (EsV-1	NP_077578
Emiliania huxleyi virus-86 (EhV-86)	AJ890364
Equid herpesvirus	YP_053075.1
Feldmania species virus (FsV)	AAB67116
Feldmannia irregularis virus (FirrV-1)	AAR26842
Fowlpox virus (FOWPV)	NP_039057
Heterosigma akashiwo virus (HaV)	AB194136.1
Human herpesvirus	CAA28464
Lymantria dispar nuclear polyhedrosis baculovirus (LdNPV)	BAA02036
Micromonas pusilla virus SP1 (MpV-SP1)	AAB66713
Micromonas pusilla virus-PL1	AAB49747
Micromonas pusilla virus-SG1	AAB49746
Mimivirus	YP_142676
Molluscum contagiosum virus (MOCV)	NP043990
Ostreococcus tauri virus 5 (OtV5)	ABY28020
Paramecium bursaria Chlorella virus (PBCV-1)	AAC00532
Paramecium bursaria Chlorella virus NY2A (PBCV-NY2A)	AAA88827
Phaeocystis globosa virus 01 (PgV-01)	ABD62757
Pyramimonas orientalis virus (PoV-01)	ABU23717
Vaccinia virus (VACV)	A24878
Sulfurisphaera ohwakuensis	BAA93703

4.3. Results

4.3.1. Viral lysis of host strains

All of the virus isolates (both OtV and MpV; listed in Table 3.5) in this study caused almost complete lysis of the host cultures with no detectable regrowth after 12 weeks. However, there was one exception, OtV-3, which infects the *O. tauri* strain RCC 501. Following lysis of host cultures, regrowth was observed after a period of 7-10 days following clearing of the culture (Fig. 4.1). Analysis of lysed cultures, using analytical flow cytometry (AFC), showed the host population was reduced but not eliminated. When regrowth of the infected RCC 501 strain was observed, an experiment was conducted to test if the viruses present in the culture were infective. A 10 ml aliquot of the recovering cultures was centrifuged (5 minutes, 2,500 × *g*) and 100 µl of the 0.2 µm filtered supernatant was transferred to an exponentially growing uninfected culture of strain RCC 501. Lysis was observed each time and regrowth occurred repeatedly, confirming that the viruses in the cultures were infective to strain RCC 501.

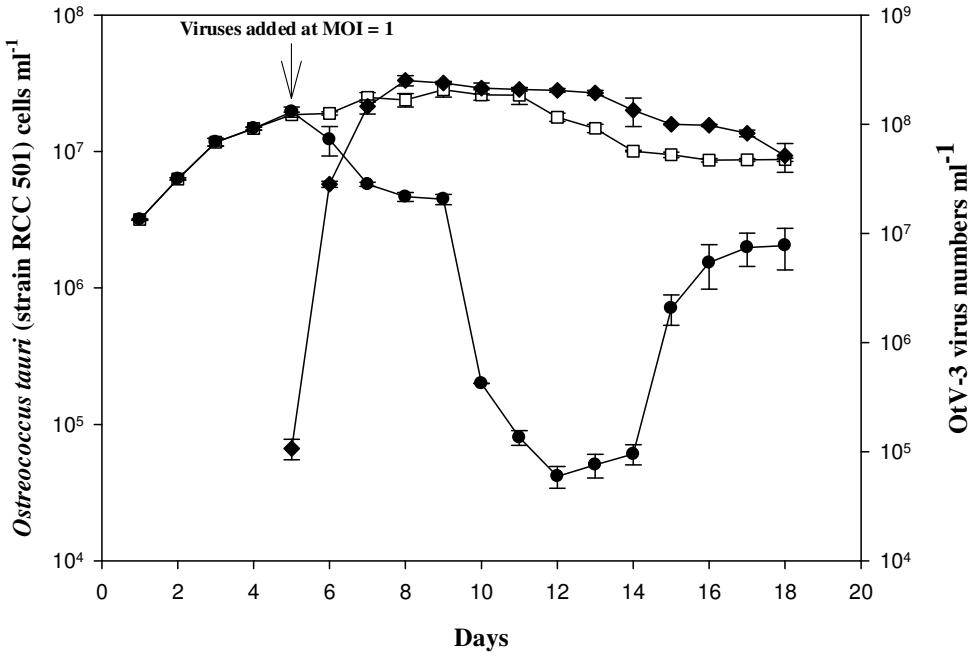


Figure 4.1. Infection kinetics of OtV-3. Growth curves of *O. tauri* strain RCC 501 in both a control non-infected culture (□) and a culture infected with OtV-3 (●), as measured by AFC. Abundance of OtV-3 is also shown (♦). Error bars represent the standard deviation of the mean (n=3). The arrow indicates the time point of viral infection. Regrowth was observed in infected cultures. Lysed cultures returned to a light green appearance following clearing of liquid cultures due to viral lysis.

4.3.2. Host range analysis

A range of 12 virus isolates were successfully isolated from seawater samples collected during this study (Table 3.5). Eight of these viruses infected six strains of *Ostreococcus tauri* and four infected a single strain of *Micromonas pusilla.* The host range of all 12 virus isolates was tested using 13 phytoplankton strains, including six *O. tauri* strains, a single *M. pusilla* strain, three *Synechococcus* strains, two *Prochlorococcus* strains and *Chlamydomonas reinhardtii*. All of these isolated viruses were subsequently shown to have a very narrow host range, with each virus isolate successfully infecting only a single host strain (Table 4.2).

The virus, OtV-1, was only lytic to the high-light adapted *O. tauri* strain OTH 95 (Table 4.3) demonstrating the high specificity of this virus. The low-light-adapted *O. tauri* strain RCC 393 was susceptible to infection by two viruses, OtV-2 and OtV-6, which had been isolated from seawater screened using this particular algal strain. The RCC 393 strain was not lysed by any of the other virus strains used in this host range analysis (Table 4.2). A second low-light-adapted *O. tauri* strain, RCC 143, was seen to lyse following the addition of the virus OtV-7, but was not susceptible to infection by any other virus tested in the host range study.

Table 4.2. Host range study of the *Ostreococcus tauri* virus (OtV) and the *Micromonas pusilla* virus (MpV) isolates with the *O. tauri* and *M. pusilla* host strains and cyanobacteria and freshwater algal strains kept in culture. Culture lysis (+); no evidence of culture lysis (–).

Host culture strain	Virus isolate												
Species	Culture code	MpV-1	MpV-2	MpV-3	MpV-4	OtV-1	OtV-2	OtV-3	OtV-4	OtV-6	OtV-7	OtV-8	OtV-9
Ostreococcus tauri	OTH95	–	–	–	–	+	–	–	–	–	–	–	–
Ostreococcus tauri	RCC 501	–	–	–	–	–	–	+	+	–	–	–	–
Ostreococcus tauri	RCC 393	–	–	–	–	–	+	–	–	+	–	–	–
Ostreococcus tauri	RCC 143	–	–	–	–	–	–	–	–	–	+	–	–
Ostreococcus tauri	RCC 356	–	–	–	–	–	–	–	–	–	–	–	+
Ostreococcus tauri	RCC 344	–	–	–	–	–	–	–	–	–	–	+	–
Micromonas pusilla	RCC 434	+	+	+	+	–	–	–	–	–	–	–	–
Synechococcus	WH 7803	–	–	–	–	–	–	–	–	–	–	–	–
Synechococcus	WH 8102	–	–	–	–	–	–	–	–	–	–	–	–
Synechococcus	RS 9917	–	–	–	–	–	–	–	–	–	–	–	–
Prochlorococcus	SS 120	–	–	–	–	–	–	–	–	–	–	–	–
Prochlorococcus	MIT 9313	–	–	–	–	–	–	–	–	–	–	–	–
C. reinhardtii	CC-124	–	–	–	–	–	–	–	–	–	–	–	–

The *O. tauri* strain RCC 501 was susceptible to infection by two viruses, OtV-3 and OtV-4. These two viruses were detected and isolated in this study from seawater screened using the *O. tauri* strain RCC 501. The *O. tauri* strain RCC 344 was seen to lyse following infection with OtV-8 only. The *O. tauri* strain RCC 356 was only susceptible to infection by OtV-9.

The *M. pusilla* strain RCC 434, was susceptible to infection by the four viruses, MpV-1 - 4, isolated from seawater screened using this particular algal strain as host. The four viruses MpV 1 - 4 were unable to cause lysis of any of the *O. tauri* strains.

None of the OtV isolates or the MpV isolates were seen to infect and lyse any of the cyanobacterial species, both *Synechococcus* and *Prochlorococcus,* included in this host range analysis or the freshwater species, *Chlamydomonas reinhardtii* (Table 4.2).

4.3.3. *Ostreococcus* viruses

4.3.3.1. OtV-1

The clonal virus, OtV-1, causing lysis of *Ostreococcus tauri* strain OTH 95, was isolated from seawater collected at sampling station L4 on 06/09/2006. The virus was filterable through a 0.2 µm polycarbonate membrane filter. Lysis was evident when the culture turned from a dense green colour to a translucent appearance with evidence of cellular debris at the bottom of the culture flask. Serial dilutions of these lysed cultures down to 10^{-6} ml^{-1} lysed exponentially growing *O. tauri* OTH 95 strain host cultures, which indicated that lysed cultures contained at least 10^6 infectious virus particles ml^{-1}. Host cultures lysed by the virus became pale in colour, presumably due to the loss or degradation of photosynthetic pigments.

A necessary and important step, following the concentration of virus lysate, is to purify the virus fraction and this can be achieved by caesium chloride density gradient centrifugation. Following this procedure, a tight band was obtained (Fig. 4.2) of purified OtV-1 virus with a density of ~ 1.2 g cm^2 .

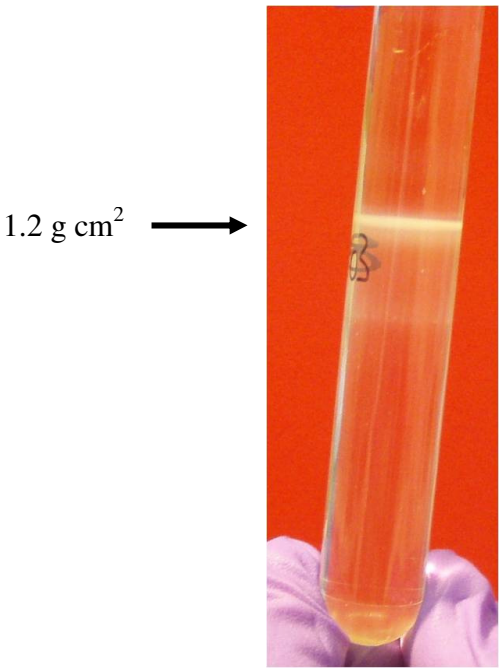


Figure 4.2. Caesium chloride gradient of OtV-1. The visible band is of purified OtV-1 viruses shown to have a density of 1.2 g cm^2. This band was extracted for use in further characterisation studies.

4.3.3.1.1. Infection Kinetics of OtV-1

The addition of viruses to exponentially growing *O. tauri* cells usually resulted in the decrease in cell number 2-3 days after virus addition (Fig. 4.3). The infected culture collapsed three days post-infection, while the non-infected cultures grew as normal. The appearance of an OtV-1 group, during analysis of samples using a flow cytometer (AFC), correlated with the decline in *O. tauri* abundance.

At the point of infection, using a multiplicity of infection (moi) of one, host cell density was on average 4×10^6 cells ml^{-1} (Fig. 4.3). However, with a concurrent rise in virus numbers, the infected cultures were seen to start to decrease in abundance two days after infection, eventually reaching lowest abundances 3-4 days post-infection. Based on the decline in the algal host population and the increase in extracellular virus particles, a conservative estimate of the burst size can be estimated. Virus concentrations calculated by AFC, reached highest average concentrations of 1.2×10^8 ml^{-1} (Fig. 4.3). This would give a burst size estimate of approximately 30 viruses per cell.

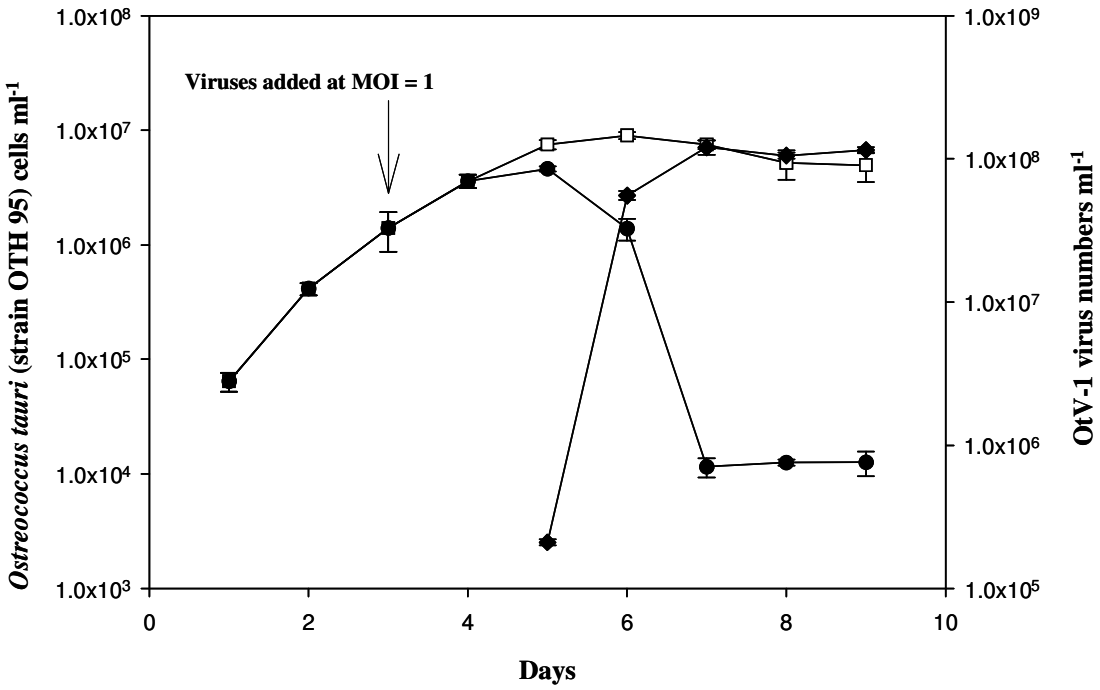


Figure 4.3. Growth curves of *O. tauri* strain OTH 95 in both a control non-infected culture (□) and a culture infected with OtV-1 (●), as measured by AFC. Abundance of OtV-1 is also shown (♦). Error bars represent the standard deviation of the mean (n=3). The arrow indicates the time point of viral infection.

4.3.3.1.2. Morphology of OtV-1

Transmission electron micrographs of negatively stained whole mounts of lysed *O. tauri* strain OTH 95 cells inoculated with OtV-1 showed free viral particles (Fig. 4.4 a-c). The OtV-1 virus particles showed the hexagonal outlines typical for the distinctive icosahedral symmetry of large dsDNA algal viruses (Fig. 4.4 a-c.). The OtV-1 virion sizes ranged between 100 to 120 nm with no visible tail but with an electron dense core surrounded by a distinct capsid (Fig. 4.4 a-c.). Although no tail-like structure is visible, a small protrusion is discernible in the transmission electron micrographs on the external surface of the capsid shells of the OtV-1 viruses (Fig. 4.4 a, b). This appendage may play a role in attachment of the virus to the host cell during adsorption in early infection.

An experiment was conducted to investigate the infection process of *O. tauri* strain OTH 95 by OtV-1, at regular intervals using transmission electron microscopy over a 30 hour period, starting at the point of infection. Examples of healthy uninfected host *O. tauri* OTH 95 cells are shown in the thin section electron micrographs in Figure 4.5. The single organelles can be viewed.

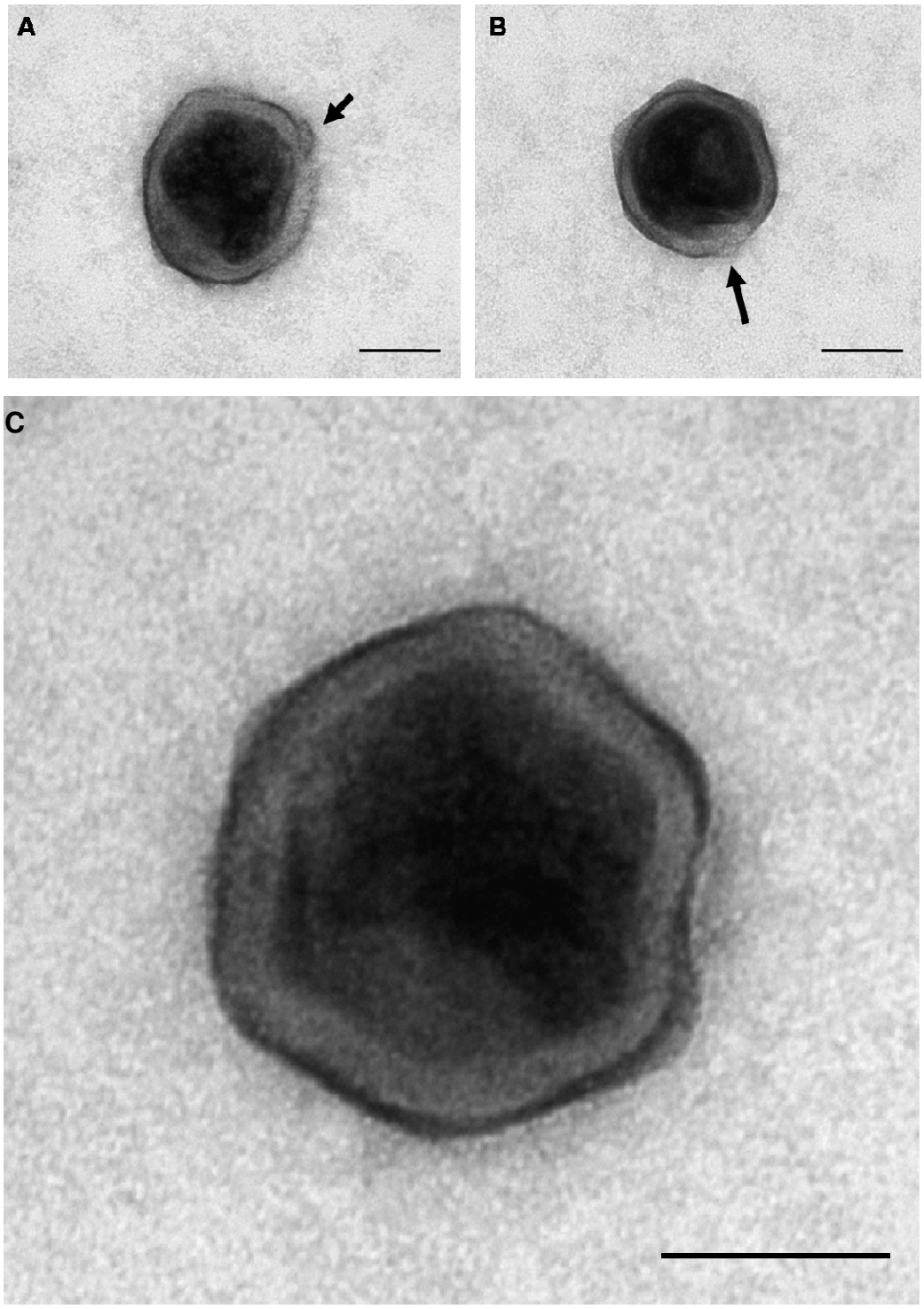


Figure 4.4. a-c. Transmission electron micrographs of negatively stained whole mounts of *Ostreococcus tauri* virus–1 (OtV-1). Detail of virus-like particles show a distinct capsid surrounding an electron dense core with a hexagonal outline suggesting icosahedral symmetry. A tail-like stub (arrows in A and B) indicates the possible attachment structure. Scale bars represent 50 nm.

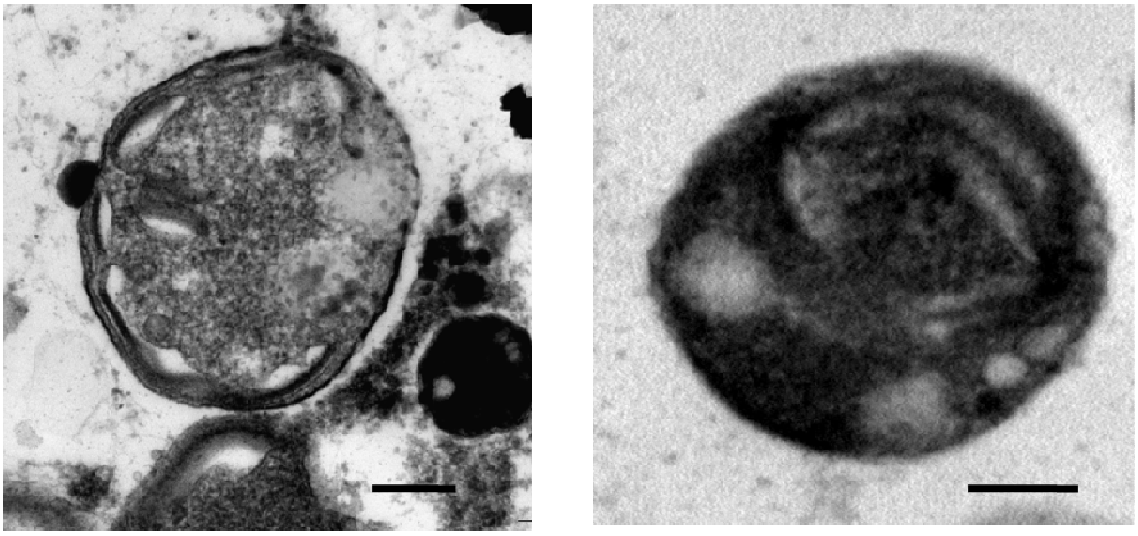


Figure 4.5. Transmission electron micrographs of *O. tauri* strain OTH 95 cultures. Thin sections of healthy cells. Scale bars represent 200 nm.

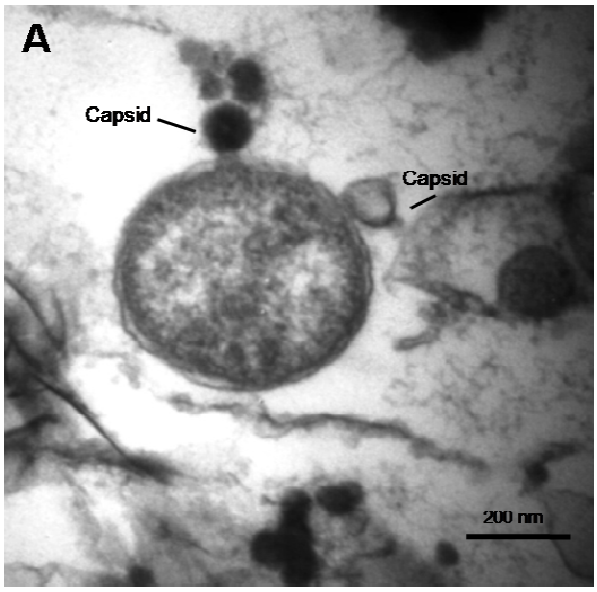
In the micrographs of material fixed at 2 hours post-infection (p.i.) (Fig. 4.6 a), two virus particles are attached to a single host cell. One of the virus capsid structures appears translucent indicating some, or all, of the genetic material has been emptied from the capsid. The second virus particle, seen attached to the top of the host cell, has a heavily stained internal region, indicating the capsid still contains the genetic material within.

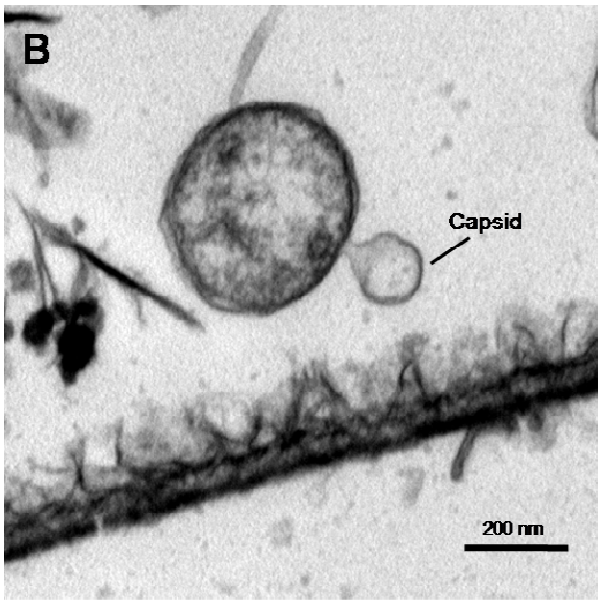
The micrograph shown in Figure 4.6 b is of a thin section of host culture fixed four hours p.i. with OtV-1. A capsid shell can be distinguished on the exterior of the cell and is fused, via a vertex, with the outer membrane of the host cell. The interior of the capsid appears translucent (Fig. 4.6 b), indicating it no longer contains genetic material, as this has been injected into the interior of the cell. The capsid structure has lost some of its integrity and no longer displays the hexagonal shape and symmetry of free OtV-1 virus particles, as seen in Figure 4.4 a-c, which still contain genetic material. The capsid structure appears larger (diameter of the particle is approximately 200 nm) in comparison to free unattached viruses that are yet to infect, which are approximately 100 – 120 nm in diameter. The attached empty capsid also appears to 'bulge', having lost the integrity of its shape.

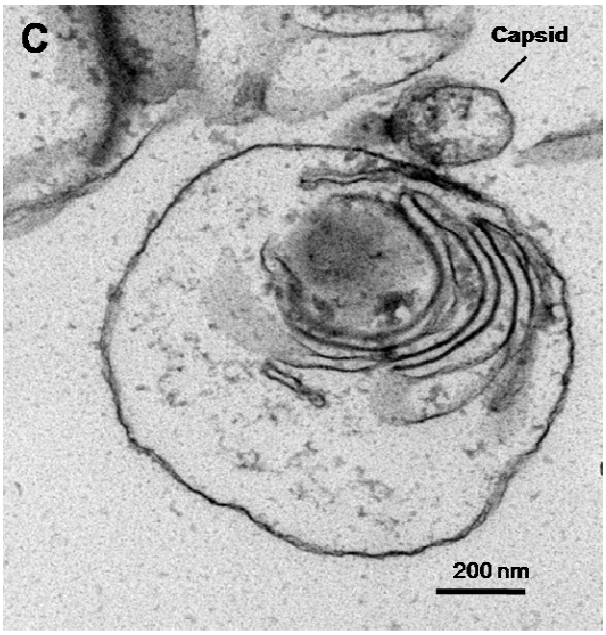
At 8 hrs p.i., empty capsid shells are seen attached to host cell membranes (Fig. 4.6 c and d). Virion assembly is discernible as internally stained hexagonal shapes are recognisable within the cell cytoplasm (Fig. 4.6 d). Empty capsid shells again appear enlarged and continue to remain attached to the cell membrane (Figs. 4.6 c and d). The regular shape of the capsids at 8 hours p.i. is lost. The integrity of host cell organelles appears to be in the process of degradation, leaving almost no visible structures intact

(Fig. 4.6 c). The structures that are visible are probably the remaining thylakoids of the chloroplast. This would be in keeping with the virus requiring the photosynthetic centres of the host cell to remain functioning for as long as possible during infection to provide energy for virus production.

The TEM image captured at 12 p.i. (Fig. 4.6 e) shows a capsid emptied of its contents and still attached to a host cell, which is now devoid of any internal organelles. Although the cell interior appears to have degraded, the cell membrane remains mostly intact.


A
Capsid
Capsid
200 nm


B
Capsid
200 nm


C
Capsid
200 nm

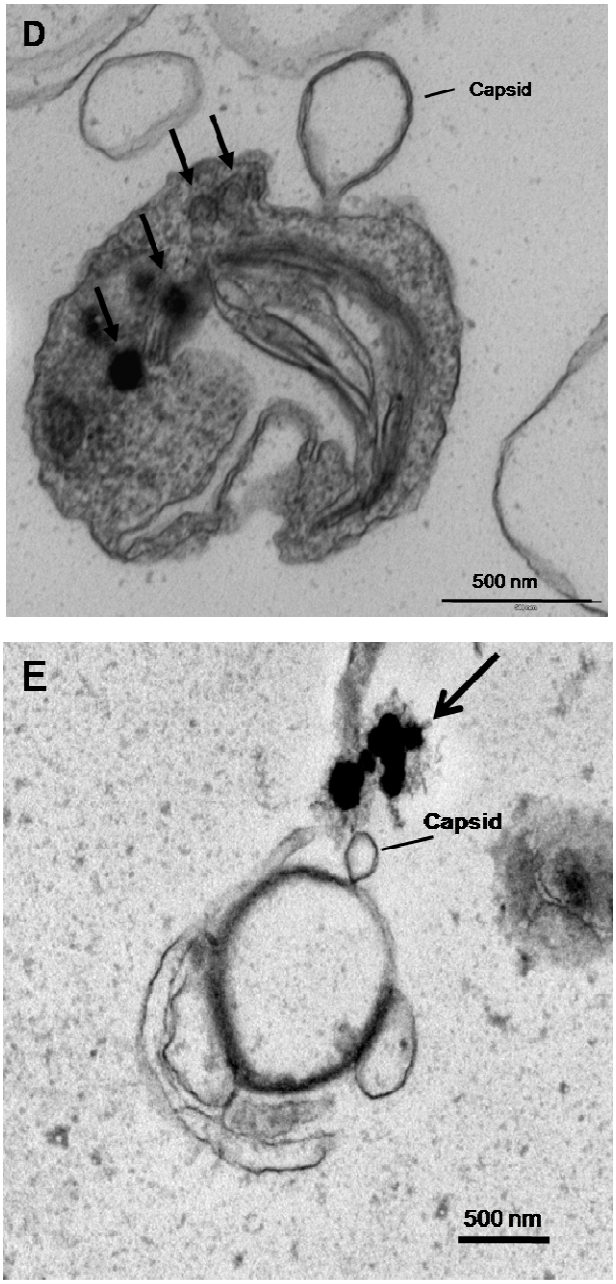


Figure 4.6. a-e. TEM micrographs of thin sections of *Ostreococcus tauri* (strain OTH 95) and OtV-1 prepared between 0 hours to 12 hours post-infection (p.i.).
(a) 2 hours p.i. Two viruses attached to the same *O. tauri* cell. Capsids are labelled: one is translucent, indicating loss of contents; the other contains genetic material. (b) 4 hours p.i. The OtV-1 virus is attached via a vertex to the host cell membrane and has lost the genetic material from the core of the capsid shell. (c) and (d) 8 hours p.i. The host cell in (c) can be seen to have lost almost all of the organelles, with only possible thylakoid structures of the chloroplast remaining. In (d) virions at different stages of assembly can be seen (arrows). The empty capsid shell is visible. The host interior contains heavily stained shapes identified as virions with packaged DNA. Host cells in (c) and (d) are in various stages of degradation. (e) 12 hours p.i. Empty viral capsid remains attached to the host cell membrane. The interior of the host cell appears almost completely empty. Arrow indicates viral particles.

A range of images captured at 24 – 30 p.i. revealed more cells in the latter stages of infection (Fig. 4.7 a-c). At 24 hours p.i., a deformed capsid shell is seen to remain attached to the exterior of the host cell (Fig. 4.7. a,b). Host cells contain virus-like particles at various stages of assembly (Fig. 4.7 a-c). Some of the visible cores of the virions are not heavily stained and therefore, do not appear to have been DNA packaged. The attached empty capsid shell in Figure 4.7 b is very large at approximately 250 nm in diameter and is clearly fused with the cell plasmalemma.

Organelles are no longer visible within the host cell interior at 30 hours p.i. and virus-like particles can be seen within the cell (Fig. 4.7 c). Unlike the earlier images, no discernible empty capsid shell is attached to the outer cell membrane (Fig. 4.7 c). A free unattached virus is visible on the outside of the cell. The host cell membrane is intact and this micrograph image does not indicate clearly how new viruses exit the cell.

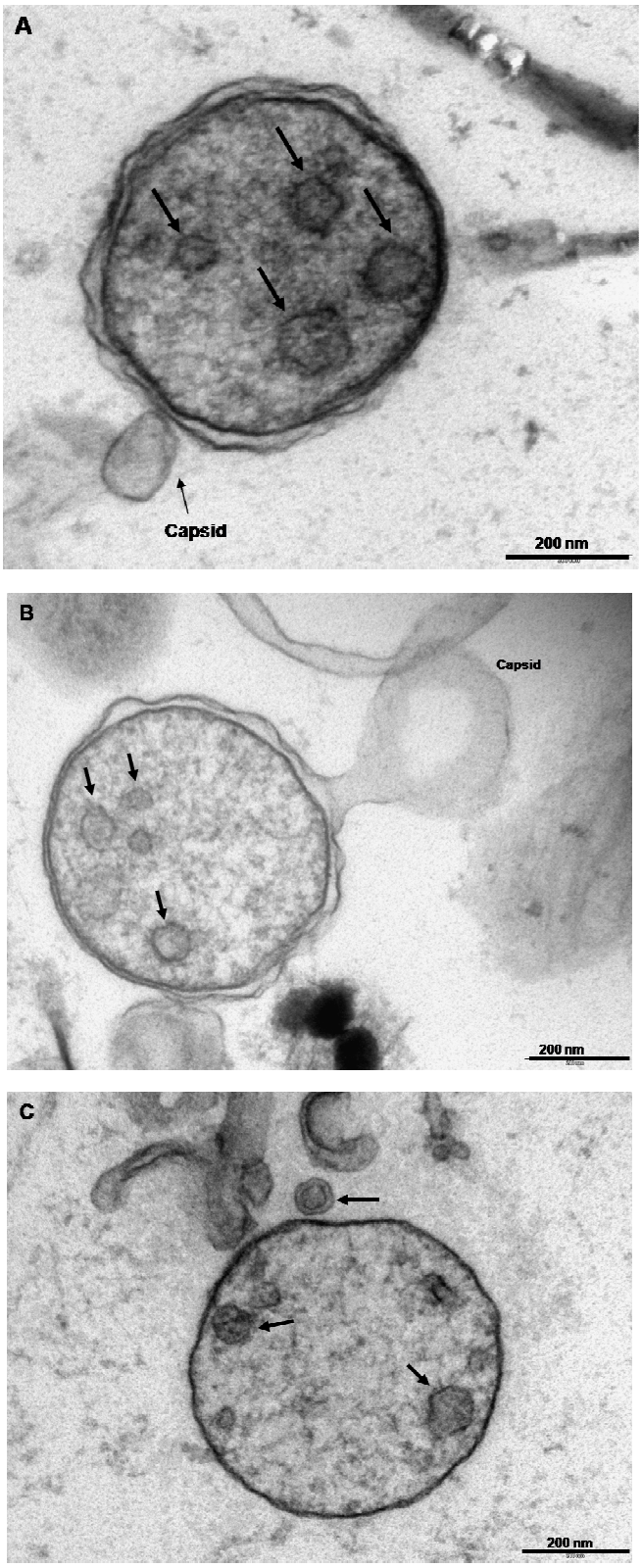


Figure 4.7.a-c. TEM micrographs, thin sections of *Ostreococcus tauri* (strain OTH 95) and OtV-1 prepared between 24 hours (a, b) and 30 hours (c) post-infection (p.i.).
(a, b) Empty viral capsids seen attached to the host cell membrane. Virions assembling in interior of host cells (arrows).

4.3.3.2. OtV-2

4.3.3.2.1. Infection kinetics of OtV-2

The addition of OtV-2 viruses to exponentially growing *O. tauri* strain RCC 393 cells usually resulted in a decrease in cell number 2-3 days after infection (Fig. 4.8). The infected culture collapsed three days post-infection, while the non-infected cultures grew as normal. The appearance of an OtV-2 group correlated with the decline in *O. tauri* strain RCC 393 abundance.

At the point of infection, using a moi of one, host cell numbers were on average 3×10^6 cells ml^{-1} (Fig. 4.8). However, with a concurrent rise in virus numbers, the infected cultures were seen to start to decrease in abundance, eventually reaching lowest abundances 3-4 days post-infection. Based on the decline in the algal host population and the increase in extracellular virus particles, a conservative estimate of the burst size can be estimated. Virus numbers reached highest average concentrations of 8.2×10^7 ml^{-1} (Fig. 4.8). This would give a conservative burst size estimate of approximately 27 viruses per cell.

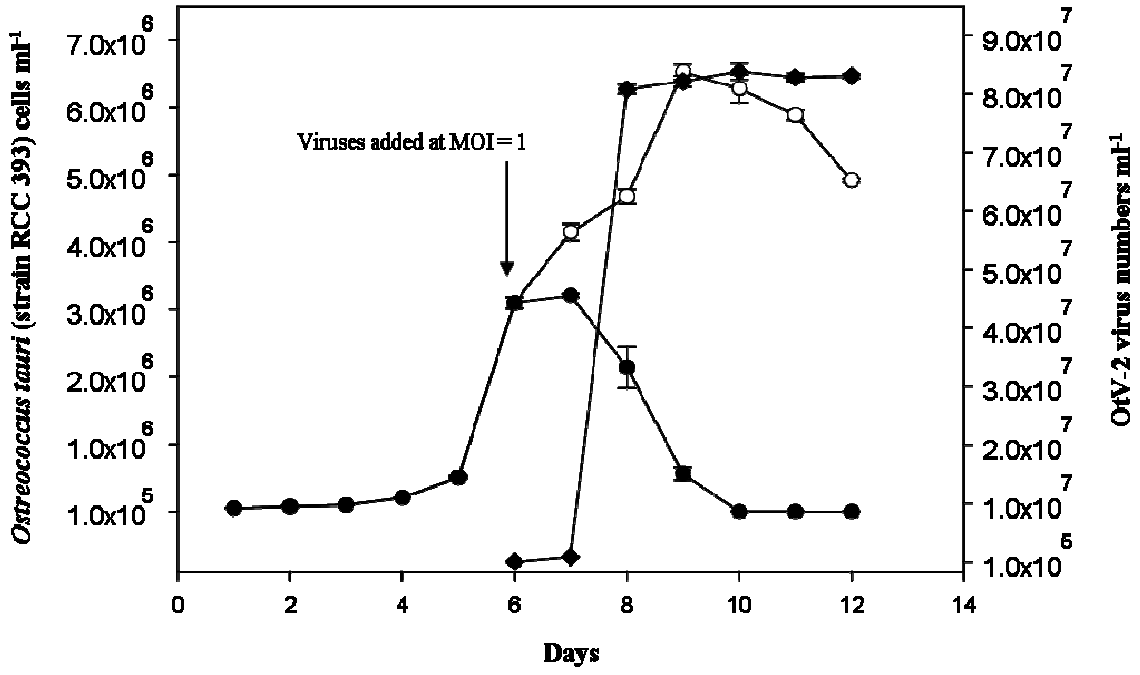


Figure 4.8. Growth curves of *O. tauri* strain RCC 393 in both a control non-infected culture (□) and a culture infected with OtV-2 (●), as measured by AFC. Abundance of OtV-2 is also shown (♦). Error bars represent the standard deviation of the mean (n=3). The arrow indicates the time point of viral infection.

4.3.3.2.2. Morphology of OtV-2

TEM images of the negatively stained whole mounts of OtV-2 virus particles showed the hexagonal outlines typical for the distinctive icosahedral symmetry of most large dsDNA algal viruses (Figs. 4.9 and 4.10 a-h). The OtV-2 virion sizes ranged between 100 to 120 nm (Fig. 4.9) with no visible tail but with an electron dense core surrounded by a capsid. The size of the OtV-2 viruses is comparable with the sizes of the OtV-1 viruses.

4.3.3.2.3. Infection kinetics of remaining OtV isolates

Complete lysis of the remaining *O. tauri* strains RCC 501, RCC 143, RCC 344 and RCC 356 was observed 2 - 4 days after virus addition (Figs. 4.11 a-d.). The growth kinetics of each of the four virus isolates OtV-6, OtV-7, OtV-8 and OtV-9 are shown.

The addition of each of these four viruses to a susceptible exponentially growing *O. tauri* strain usually resulted in a decrease in cell number 2-3 days p.i. (Fig. 4.11 a-d). The infected cultures collapsed 3-4 days post-infection, while the non-infected cultures grew as normal. The appearance of an OtV group, assessed by AFC, correlated with the decline in *O. tauri* host abundance. Conservative burst size estimates are summarised in Table 4.3 for each of the four virus isolates.

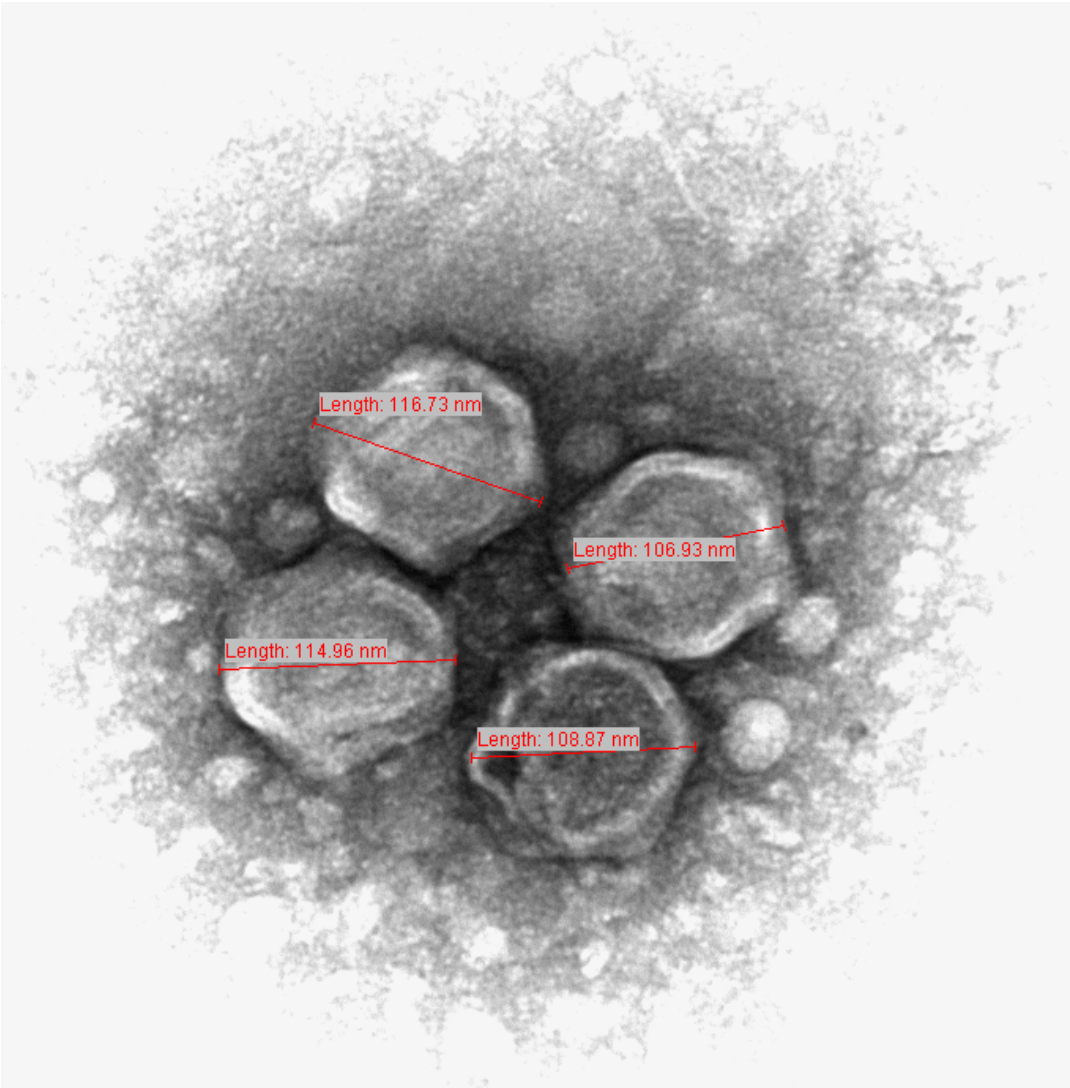


Figure 4.9. Transmission electron micrographs of negatively stained whole mount of *Ostreococcus tauri* virus–2 (OtV-2). Red lines indicate the diameter of individual virus particles in nanometres.

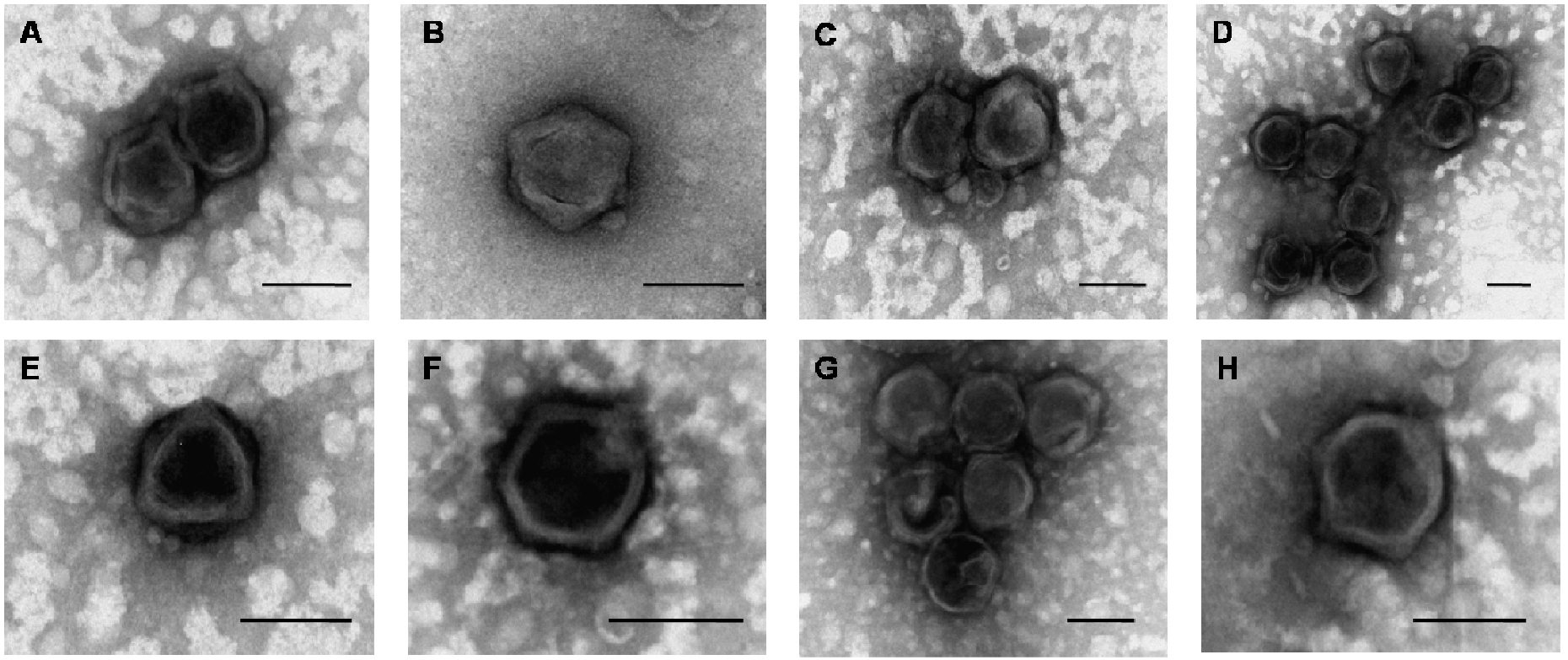


Figure 4.10. a-h. Transmission electron micrographs of negatively stained whole mounts of *Ostreococcus tauri* virus–2 (OtV-2). Detail of virus-like particles show a distinct capsid surrounding an electron dense core with a hexagonal outline suggesting icosahedral symmetry. Scale bars represent 100 nm.

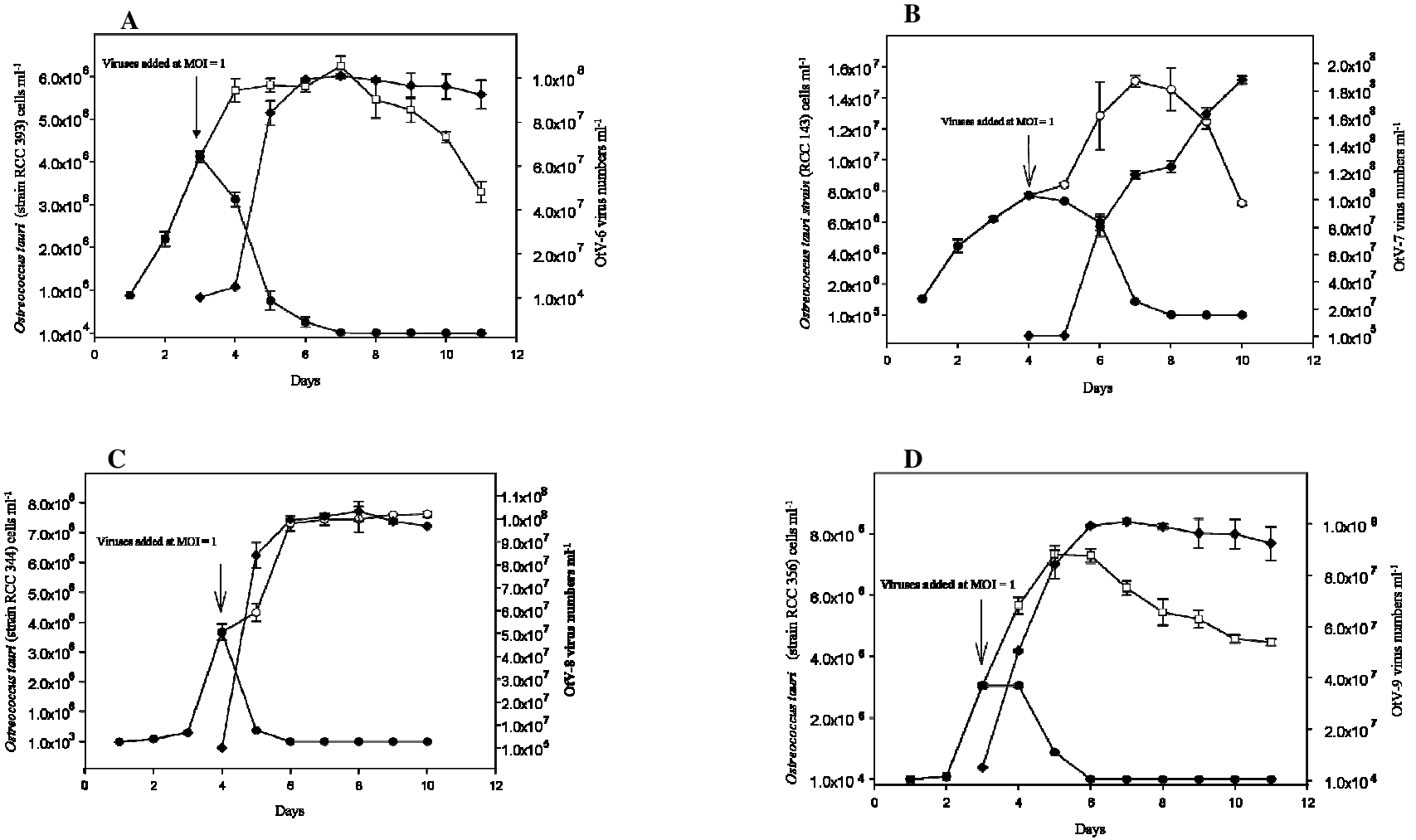


Figure 4.11. a-d Growth curves of a) *O. tauri* strain RCC 393 in a control non-infected culture (□) and a culture infected with OtV-6(●); b) *O. tauri* strain RCC 143 in a control non-infected culture (□) and a culture infected with OtV-7 (●); c) *O. tauri* strain RCC 344 in a control non-infected culture (□) and a culture infected with OtV-8 (●); d) *O. tauri* strain RCC 356 in a control non-infected culture (□) and a culture infected with OtV-9 (●) Abundance of viruses is also shown (♦). Error bars represent the standard deviation of the mean (n=3). The arrow indicates the time point of viral infection.

4.3.4. MpV viruses

4.3.4.1. Infection kinetics

Complete lysis of *M. pusilla* was observed 2-3 days after virus addition. The growth kinetics of each of the four viruses MpV-1, MpV-2, MpV-3 and Mpv-4, isolated in this study are shown (Fig. 4.12 a-d). The addition of MpV viruses to exponentially growing *M. pusilla* strain RCC 434 cells usually resulted in a decrease in cell number 2-3 days after MpV addition. The infected cultures collapsed three days post-infection, while the non-infected cultures grew as normal. The appearance of a MpV group, assessed by AFC, correlated with the decline in *M. pusilla* host abundance.

MpV-1 virus numbers reached highest average concentrations of 4.0×10^8 ml^{-1} (Fig. 4.12 a) to give a conservative average burst size estimate of approximately 67 viruses per cell. MpV-2 virus numbers reached highest average concentrations of 2.2×10^8 ml^{-1} (Fig. 4.12 b) to give a conservative average burst size estimate of approximately 73 viruses per cell. MpV-3 virus numbers reached highest average concentrations of 2.0×10^8 ml^{-1} (Fig. 4.12 c) to give a conservative average burst size estimate of approximately 80 viruses per cell. MpV-4 virus numbers reached highest average concentrations of 3.0×10^8 ml^{-1} (Fig. 4.12 d) to give a conservative average burst size estimate of approximately 60 viruses per cell.

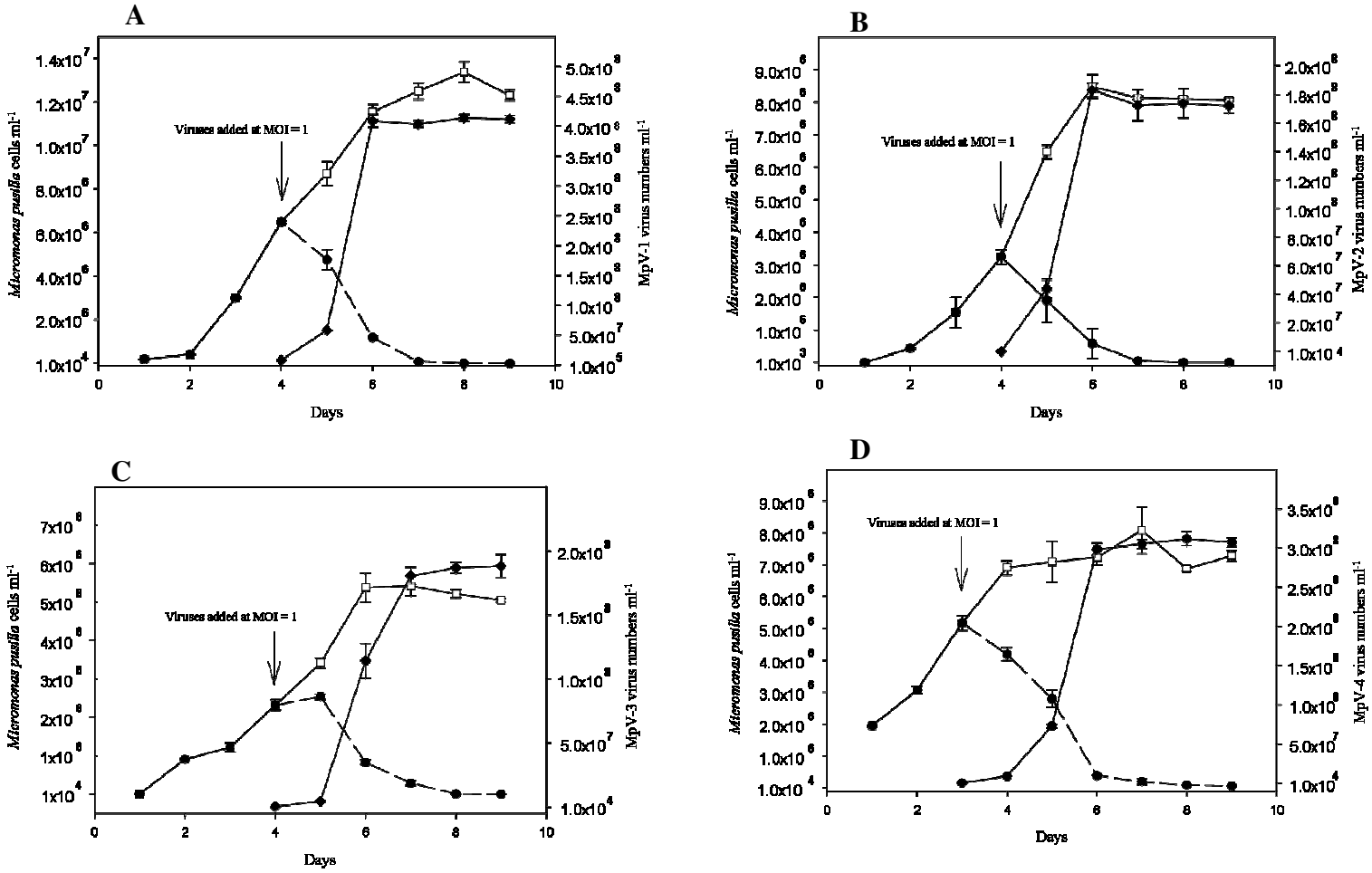


Figure 4.12. a-d. Growth curves of *M. pusilla* strain RCC 434 in both a control non-infected culture (□) and a culture infected with (a) MpV-1; (b) MpV-2; (c) MpV-3; (d) MpV-4 (●), as measured by analytical flow cytometry. Abundance of MpV particles is also shown (♦). Error bars represent the standard deviation of the mean (n=3). The arrow indicates the time point of virus addition.

4.3.4.2. Morphology of MpV isolates

TEM images of the MpV isolates show viruses in the size range of 115 nm to 135 nm. These viruses are slightly larger than the OtV viruses characterised in this study. TEM images of the negatively stained MpV virus particles show the hexagonal outlines typical for the distinctive icosahedral symmetry of most large dsDNA algal viruses (Fig. 4.13 a-f). No tail is visible but an electron dense core is seen surrounded by the hexagonal capsid shell (Fig. 4.13 a-f). In Figure 4.13 e, a possible 'star'-like shape in the capsid is discernible.

4.3.5. Genome size of OtV and MpV isolates

Pulsed field gel electrophoresis (PFGE) was conducted to estimate the genome size of the MpV and OtV virus isolates in this study. Clear bands were seen for all four MpV isolates and five of the OtV isolates (OtV-1, OtV-2, OtV-3, OtV-4 and OtV-6) indicating genome sizes of approximately 198 kb for the MpV strains and approximately 193 kb for the OtV strains (Figure 4.14). The isolates OtV-7, OtV-8 and OtV-9 failed to produce bands on the PFGE gel. This may have been due to insufficient genetic material in the prepared samples.

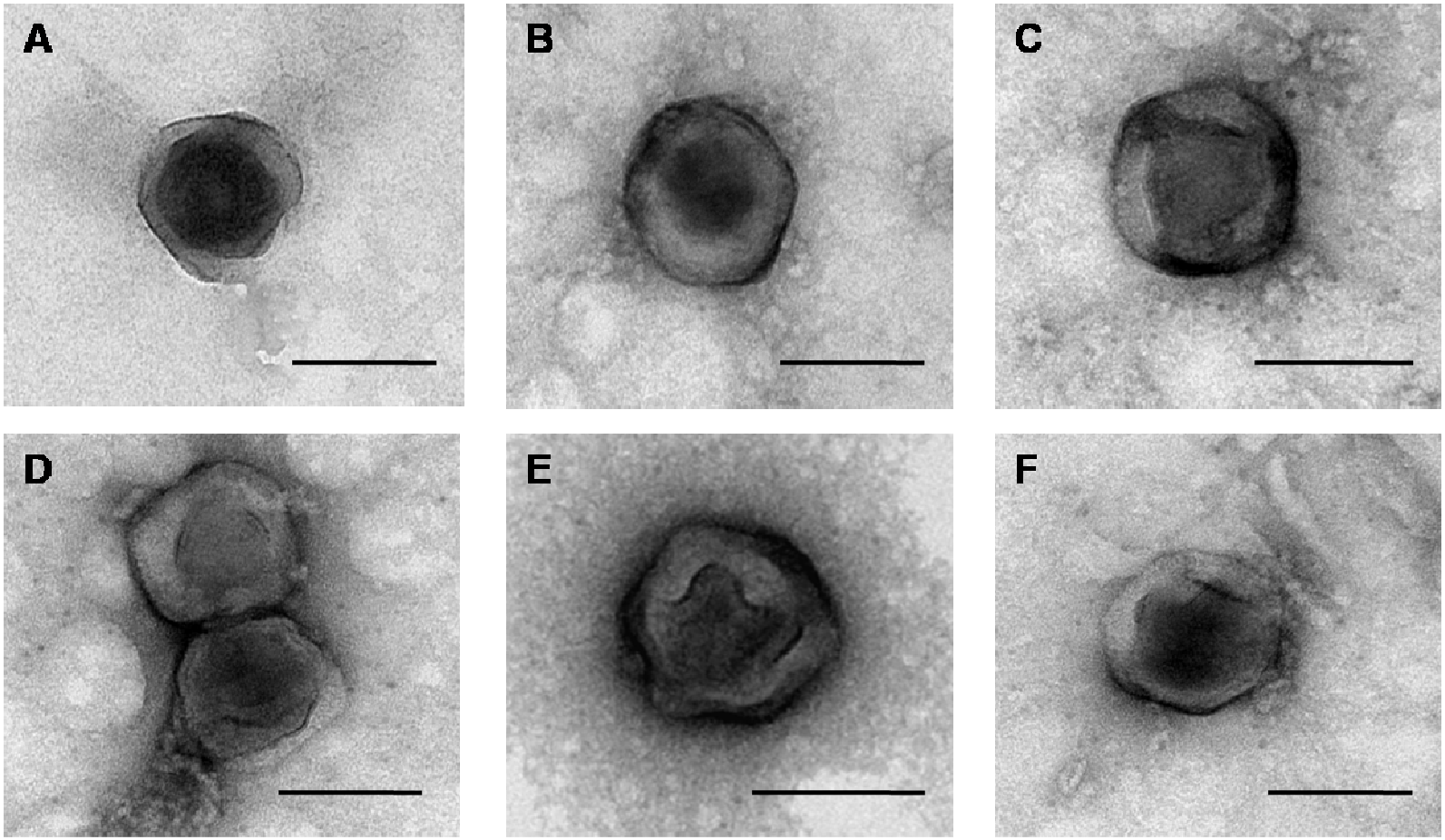


Figure 4.13.a-f Negatively stained TEM micrographs of MpV viruses, whole mounts. (a) MpV-1; (b), (c), (d) MpV-2; (e) MpV-3; (f) MpV-4. Scale bars represent 100 nm.

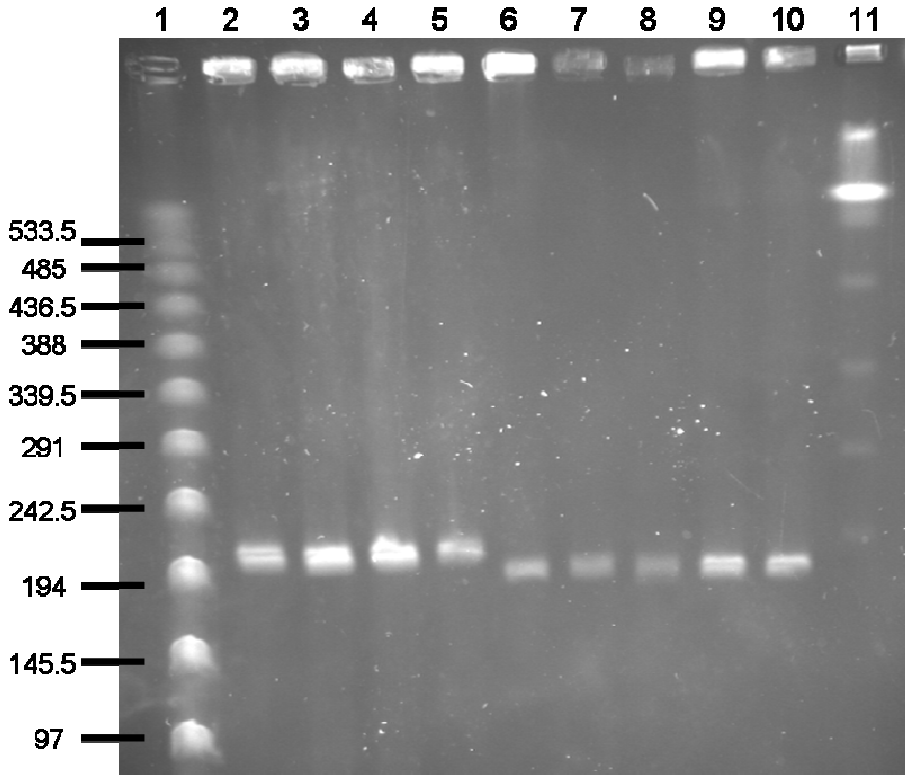


Figure 4.14. Pulsed field gel electrophoresis profiles of samples from concentrated lysates of *O. tauri* and *M. pusilla* infected with MpV strains and OtV strains isolated in this study. Lane 1 contains DNA size marker phage lambda concatamers from 97 kbp to approximately 600 kbp. Lane 11 contains a *Saccharomyces cerevisiae* ladder marker. Lane description: **2**, MpV-1; **3**, MpV-2; **4**, MpV-3; **5**, MpV-4; **6,** OtV-1; **7,** OtV-2; **8,** OtV-3; **9,** OtV-4; **10,** OtV-6, respectively. The virus genomes migrated to a position corresponding approximately to the 194 kb marker band.

4.3.6. Phylogeny based on DNA polymerase genes

Of the 12 virus isolates obtained, all 4 of the MpV isolates were used for phylogenetic analysis. Of the 8 OtV viruses isolated here, PCR products were obtained for OtV-2, OtV-3, OtV-4 and OtV-6 (Fig. 4.15). Fragments amplified by the AVS1 and AVS2 primers from the MpV and OtV viruses corresponded to approximately 680 bp (Figs. 4.15 and 4.16 b). The degenerate primer pair AVS1 and AVS2 was unable to amplify the DNA polymerase gene of OtV-1 and OtV-9. Insufficient quantities of viral lysate proved inadequate for DNA extraction prior to use in PCR targeting the DNA *pol* gene of the OtV-9 virus. In addition, the OtV isolates, OtV-7 and OtV-8 were not included in the phylogenetic analysis. Due to poor growth of the host strains RCC 344, RCC 356 and RCC 143, it was not possible to obtain sufficient concentrations of their co-occurring viruses, OtV-7, OtV-8 and OtV-9. Efforts to restore the *Ostreococcus* host strain cultures, RC 344, RCC 356 and RCC 143, to a healthy state proved unsuccessful. All three host strains were lost from the culture collection in this study before their co-occurring viruses could be included in a phylogenetic analysis. In total, four OtV isolates and four MpV isolates were used for phylogenetic analysis.

PCR primers developed by Chen *et al.,* (1996) were used to amplify the DNA *pol* gene fragments from MpV isolates. PCR products of the expected size (approximately 680 bp) were amplified from the MpV virus lysates (Fig. 4.16 a). The PCR products were subjected to a second round of amplification using the nested primers AVS1 and POL (Chen and Suttle, 1995a). All amplifications using the AVS1 and POL primer set resulted in a single band of ca. 500 bp for MpV-1, 2, 3 and 4 (Fig. 4.16 b).

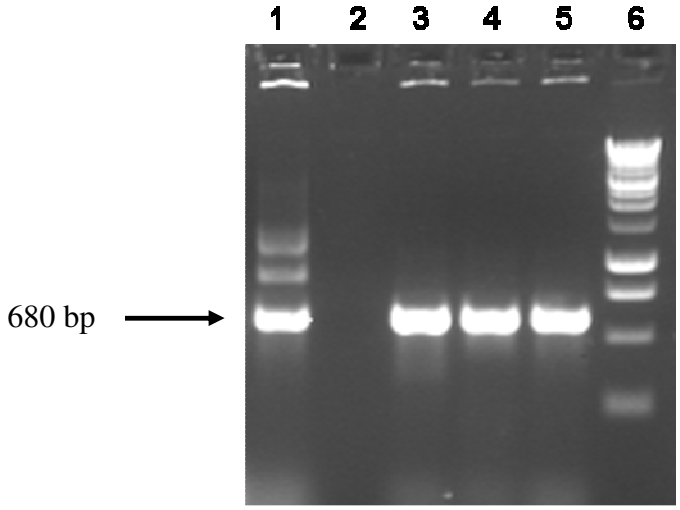
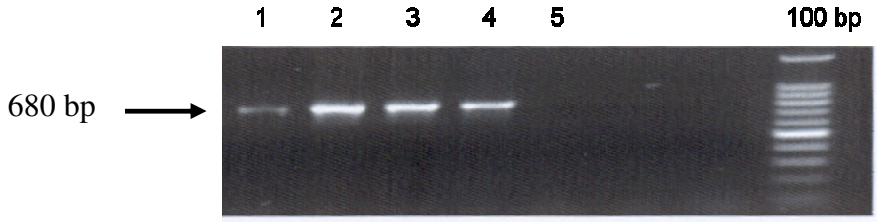


Figure 4.15. Analysis of PCR products by agarose gel electrophoresis.
Fragments from four virus isolates which infect *O. tauri* strains were amplified by the primer set AVS1 and AVS2. Lanes 1 to 5, OtV-2, -1, -3, -4, and -6 respectively. Lane 6, molecular weight marker (1 kb ladder, with bands corresponding to 10,000, 8,000, 6,000, 5,000, 4,000, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500 and 250 bp, respectively).

A)



B)

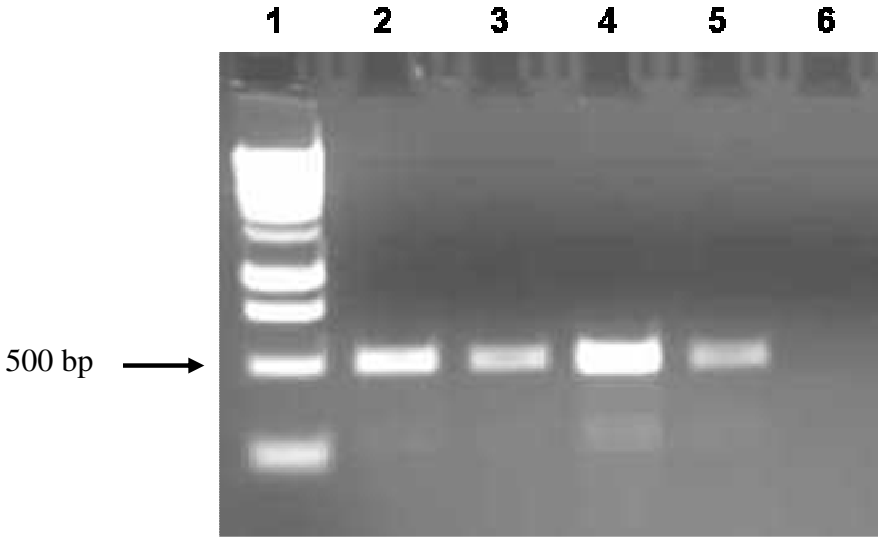


Figure 4.16. a-b. Analysis of PCR products by agarose gel electrophoresis.

A) Fragments from the four virus isolates which infect *M. pusilla* were amplified by the primer set AVS1 and AVS2. Lane **1**, MpV-1; **2**, MpV-2; **3,** MpV-3; **4,** MpV-4; **5,** negative (no DNA template) control. Outer most lane (100 bp) contains DNA molecular weight marker 100 kb ladder, with bands corresponding to 1,500, 1,000, 900, 800, 700, 600, 500, 400, 300, 200, 100 bp, respectively. B) Amplicons from the PCR in (A) were used as templates in a nested PCR using the primer set AVS1 and POL. Lane 1 contains a 1 kb DNA molecular weight marker with bands corresponding to 10,000, 8,000, 6,000, 5,000, 4,000, 3,000, 2,500, 2,000, 1,500, 1,000, 750, 500 and 250 bp, respectively. Lanes 2 to 5, MpV-1, -2, -3, -4, respectively; lane 6, negative control.

Sequence analysis and subsequent database searching of cloned products revealed a high sequence identity to DNA polymerases isolated from members of the family of algal viruses, the *Phycodnaviridae* and DNA polymerases of eukaryotic origin. A phylogenetic reconstruction of the DNA polymerase sequences obtained (Fig. 4.17) shows that the OtV viruses isolated in this study, and OtV5 from the public database, cluster with the phycodnaviruses and form a distinct cluster with both the MpV isolates in this study and the MpV sequences in the public databases. As the MpV strains described to date have been assigned to the Prasinoviruses within the *Phycodnaviridae*, the OtV isolates can also be putatively assigned as members of the Prasinoviruses.

Multiple alignments of the DNA *pol* fragments revealed that three of the MpV isolates have identical sequence with only one isolate, MpV-2, having differences in this sequence (Fig. 4.18). The genome of one of the MpV isolates, MpV-3, was prepared for sequencing. However, this attempt proved unsuccessful, due to extensive bacterial contamination in the sequence data retrieved.

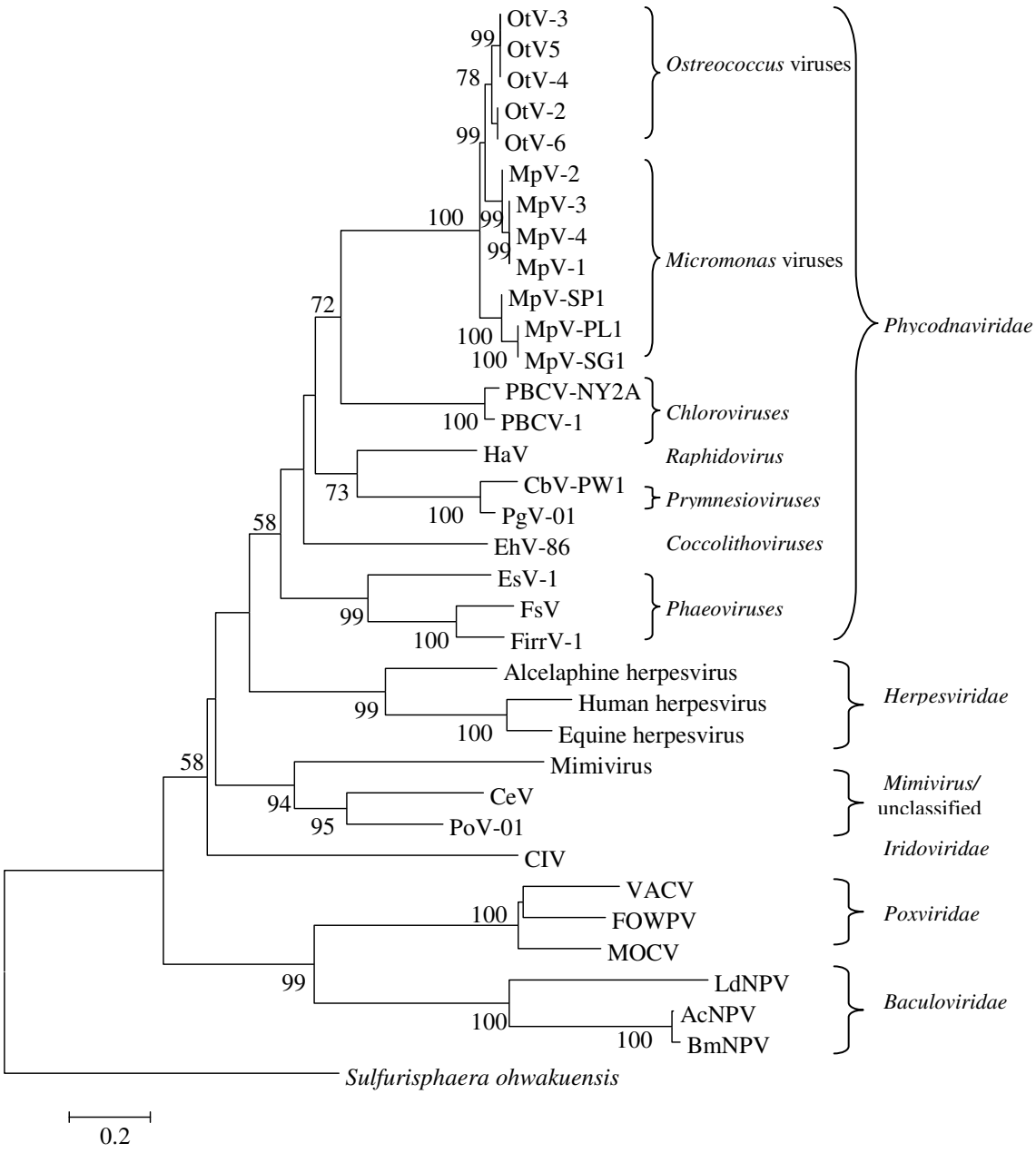


Figure 4.17. Phylogenetic relationships amongst DNA *pol* gene fragments of the family *Phycodnaviridae* and other large dsDNA viruses based on a distance matrix algorithm [Neighbour, in PHYLIP version 3.68 (Felsenstein, 2005)]. The alignment was performed (ClustalW) on the region spanning the highly conserved regions I and IV (Villarreal and DeFilippis, 2000) of the DNA *pol* genes. OtV-2, OtV-3, OtV-4, OtV-6, MpV-1, MpV-2, MpV-3 and MpV-4 were sequences from isolates obtained during this study. Numbers at nodes indicate bootstrap values retrieved from 1000 replicates for the neighbour-joining analyses. The DNA *pol* II(a) gene fragment from the Archean, *Sulfurisphaera ohwakuensis,* was used to root the tree. The bar indicates a distance of 0.2.

```
MpV-3  GAAGGGGCGACGGTGCTGGACGCCCAAAAGGGTGCGTATTACACTCCGATTACTGCCCTA
MpV-2  ............................................................
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  GATTTTGAGGCTCTGTACCCGAGTATCATGATGGCTCACACCCTCTGTTACTCCTCGTAT
MpV-2  ........................A.............A......C.................
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  GTCATGAATGAGAAGGACTATGGCAACATACCTGGTATTGAATATGAAACGTTCAAGATT
MpV-2  ............................................................
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  GGTGCAAAGACTTACAAGTTTGCACAAGATGTTCCTAGCCTCCTACCGGCTATCCTTCTA
MpV-2  ............................................................
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  GAGCTTAAGCAGTTCCGTAAAAAGGCTAAGAAGGATATGGCAGCTGCGACAGGTTATATG
MpV-2  ..................................................G.........
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  AAGGAGGTCTACAATGGTAAACAGTTGGCCTATAAAATCAGTATGAACTCAGTCTATGGA
MpV-2  ..A.........................A..C.....T.................C...
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  TTTACTGGCGCTGGTAAGGGTATTCTTCCATGTGTACCTATTGCGTCTACTACAACATTT
MpV-2  ....................................................C...
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  AGGGGTCGCGCAATGATTGAAGAGACTAAGAATTACGTTGAGAAAAACTTCCCAGGTTCA
MpV-2  ..A............................C..T.....A..G........G......
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  AAAGTGAGATATGGCGACACCGATTCAGTCATGGTTGAATTTGATGTGGGTGACCGTAAG
MpV-2  ..G..T..G....................................................
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  GGTGAAGAAGCTGTTAAGTACAGCTGGGAGATTGGTGAGAGACCTGCCGAAGAGTGCTCA
MpV-2  ..........................................G.................
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  GCTCTTTTCAAGAAGCCTAACAATCATGAATTAGAAAAGGTGTATTGGCCTTATTTGCTG
MpV-2  ........................T...............................C...
MpV-1  ............................................................
MpV-4  ............................................................

MpV-3  TACACCAAAAAACGCTACGCCGC
MpV-2  .......................
MpV-1  .......................
MpV-4  .......................
```

Figure 4.18. Multiple sequence alignment of the amplified segments of the MpV DNA *pol* genes of MpV-1, MpV-2, MpV-3 and MpV-4. Conserved bases are identified as a dot underneath the corresponding base from MpV-3.

A summary of the characteristics of the viruses isolated in this study are outlined in Table 4.3.

Table 4.3. Summary of characteristics of the viruses isolated in this study.

Virus isolate	Virus size (nm)	Burst size	Genome size	Nucleic acid extracted (Yes/No)	DNA *pol* amplicon with AVS primers (Yes/No)
OtV-1	100-120	30	~ 190 kb	Y	N
OtV-2	100-120	27	~ 190 kb	Y	Y
OtV-3	100-120	30	~ 190 kb	Y	Y
OtV-4	ND	29	~ 190 kb	Y	Y
OtV-6	ND	25	~ 190 kb	Y	Y
OtV-7	ND	25	ND	N	N
OtV-8	ND	33	ND	N	N
OtV-9	ND	33	ND	N	N
MpV-1	115-135	67	~ 195 kb	Y	Y
MpV-2	115-135	73	~ 195 kb	Y	Y
MpV-3	115-135	80	~ 195 kb	Y	Y
MpV-4	115-135	60	~ 195 kb	Y	Y

4.4. Discussion

The characterisation of algal viruses is a necessary step towards understanding their ecological role in the environment. Here, the characterisation of a new set of algal viruses that infect the smallest free-living eukaryote known to date, *Ostreococcus tauri* and the prasinophyte *Micromonas pusilla* is reported.

The algal viruses isolated in this study have a number of characteristics that suggest they belong to the family *Phycodnaviridae.* This family currently consists of five genera: *Chlorovirus, Phaeovirus, Coccolithovirus, Prymnesiovirus* and *Prasinovirus* (Wilson *et al.,* 2005a). Members of the family *Phycodnaviridae* all contain large dsDNA genomes, ranging from 154 to 560 kb (Sandaa *et al.,* 2001; Schroeder *et al.,* 2009) and infect both freshwater and marine algae. The viruses isolated in this study were shown to contain large dsDNA genomes (ca. 190-195 kb), have an icosahedral symmetry and have sizes ranging between 100 to 135 nm.

The host range analysis of all the OtV and MpV isolates revealed very high host specificity. This is a common feature of viruses that infect large dsDNA eukaryotic microalgae. Although algal viruses are generally host species-specific, there can be considerable variation in the extent of strain specificity. Resistant host strains have been reported for most eukaryotic viruses in culture, but the strains tested generally include several isolated from distinctly different geographical locations. However, certain studies have shown that host strains from the same geographical location may resist infection when challenged with the same virus clone (Waters and Chan, 1982; Cottrell and Suttle, 1991; Waterbury and Valois, 1993; Nagasaki and Yamaguchi, 1998; Tarutani *et al.,* 2001; Schroeder *et al.,* 2002). To date, most algal viruses described have very narrow host ranges with some exhibiting not only species-restricted host

range but a strain-specific spectrum of infection. The proteins on the surface of free viruses serve as a means of virus attachment to receptors on the host cell surface allowing transfer of virus genomes into the cell. Any changes in the composition of these structural proteins between viruses may affect the binding to host receptors and consequently affect host range (Baranowski *et al.,* 2001). The susceptibility of the raphidophyte *Heterosigma akashiwo* to infection by HaV was seen to differ among clonal strains in the laboratory (Tarutani *et al.,* 2006). Similarly, the dinoflagellate *Heterocapsa circularisquama* was also seen to differ greatly in its response to infection by HcRNAV (Mizumoto *et al.,* 2008).

Phylogenetic analysis of the DNA *pol* gene fragments, in conjunction with the size of the viral genomes and host range of the viruses isolated in this study, suggests they can be classified into the genus *Prasinovirus*. The fact that the DNA *pol* gene of all the MpV isolates and four of the OtV isolates could be amplified using the algal virus-specific primers, AVS1 and AVS2 (Chen & Suttle 1996), allows assignment of these viruses to the family *Phycodnaviridae* (Van Etten 1995). Although the use of the highly conserved DNA polymerase gene was a good genetic marker for classification of all of the MpV isolates, and a number of the OtV strains, it did not amplify the DNA *pol* gene fragment in the OtV-1 isolate. This may have been due to primer mismatch with the target sequence, as the AVS primer set is degenerate. Another possibility is the presence of an intein in this gene. Indeed, following the sequencing of the complete OtV-1 genome (Chapter 5, this study), the presence of an intein was revealed in the DNA *pol* gene of this virus, thus explaining why it was not possible to generate sequence data using the AVS1 and AVS2 primer pair sets.

The DNA *pol* gene fragment sequences of the MpV isolates in this study differentiated one of the isolates, MpV-2, from the other three isolates described. This is in contrast to a previous study reporting enough heterogenecity in DNA *pol* gene fragments to resolve genetic variation among eight related viruses that all infect *M. pusilla* (Chen & Suttle 1996). However, Schroeder *et al.,* (2002) also reported no differentiation between isolates in the amplified DNA *pol* target sequence of *Emiliania huxleyi* viruses (EhVs) that were isolated from the English Channel. However, the amplified fragments of the major capsid gene were not so well conserved and distinctions between isolates were possible based on sequence data of this gene. It is possible three of the MpV isolates were in fact the same virus and only MpV-2 and this virus are therefore true individual isolates. Targeting of other genes, which are less evolutionarily conserved, should resolve this.

The burst size of viruses infecting *M pusilla* has previously been reported as, on average, 72 particles per cell (Waters and Chan, 1982). The estimated burst sizes of the four MpV isolates in this study are comparable with this figure. The estimated burst size of approximately 30 particles per cell for the OtV virus isolates in this study is lower than that for the MpV isolates. The smaller host cell size of < 1µm may be a factor in the lower number of viruses released during infection. The ultrastructure of the *O. tauri* cell has been examined in detail and the cytoplasmic compartment was reported as representing a maximum of 30% of the entire cell volume (Henderson *et al.,* 2007). As this volume is also occupied by numerous ribosomes and vesicles, the number of virus particles produced will most likely be restricted due to volume constraints. In addition, if the propagation strategy of OtV-1 is to release virus progeny over a period of time without a sudden burst event, then this would explain the TEM

images having only 4 or 5 virus particles present in infected cells. The total release of progeny over an extended burst event would give the final total number of VLPs estimated.

TEM analysis of OtV and MpV isolates during this study revealed they are morphologically similar, although the MpV virus particles were larger with diameters 125 nm (±10 nm) than the OtV particles, which had diameters 112 nm (±8nm). The structure of the virus particles corresponds to those reported for other members of the *Phycodnaviridae,* with a hexagonal symmetry indicating an icosahedral capsid surrounding an electron dense core.

The TEM micrograph of MpV-3 indicated a possible 'star' shape appearance in the capsid structure. Such a structural phenomenon has been reported and characterised in the Mimivirus where it is believed to play a central role as an entry mechanism for the genetic material of the virus to pass into the host cell interior (Zauberman *et al.,* 2008). The Mimivirus capsid undergoes a large conformational change and five icosahedral faces are seen to open at a unique vertex, which has been termed the 'stargate'. It is through this membrane conduit that the viral genome is delivered. That an icosahedral dsDNA virus utilises this mechanism to enter a host cell may mean it is possible that other inner-membrane containing viruses may also infect a cell via this process. The TEM image of the MpV-3 virus particle in this study hints at a second 'stargate'-type structure.

The TEM images of whole mounts of OtV-1 show the presence of a tail stub, which is a possible attachment mechanism during adsorption of the virus to the host. Analysis of

thin sections of infected algal samples in the TEM was performed for *O. tauri* infected with OtV-1. Viruses are seen to attach to the cell membrane and, in some micrograph images, appear to fuse with the membrane itself. Following emptying of the genetic material from the capsid into the host cell, the empty capsid shell remained attached to the host membrane, adding support to a possible fusion mechanism of virus with cell membrane. This initiation of infection by specific and irreversible attachment to the host membrane is similar to the start of infection by the chlorovirus, PBCV-1 (Meints *et al.,* 1984; Meints *et al.,* 1988). Attachment of PBCV-1 to its *Chlorella* host occurs at a unique vertex and is immediately followed by degradation of the host wall at the point of contact by a virus-packaged enzyme(s) (Yamada *et al.,* 2006). The mechanism by which OtV-1 infects its host may therefore be similar to the propagation strategy of chloroviruses, such as PBCV-1.

Characterisation of chloroviruses and phaeoviruses has revealed these consist of non-enveloped viruses that 'inject' their genome into the host via viral inner membrane host plasma membrane fusion mechanism leaving an extracellular viral capsid. This is in contrast to the strategy employed by another phycodnavirus, namely the coccolithovirus EhV-86. Early infection of *Emiliania huxleyi* by EhV-86 particles is marked by envelopment of the virus into the host cell (Mackinder *et al.*, in press) either by an endocytic or envelope fusion mechanism. The differences in routes of entry by these viruses may reflect the nature of the host cell that is infected.

The attachment of the chlorovirus PBCV-1 has been well-characterised. It is understood that the internal membrane of PBCV-1 fuses with the host membrane to deliver virus DNA to the host interior and an empty capsid remains on the cell surface

(Van Etten 2003). The infection mechanism of the chloroviruses shares similarities with that of icosahedral dsDNA lipid-containing bacteriophages, such as the tectivirus, PRD1 (Bamford *et al.,* 1995). In this respect, chloroviruses are not like other DNA viruses that infect eukaryotes. Infection by the OtV-1 virus isolated in this study may therefore resemble that of the PBCV-1 entry, with uncoating occurring at the cell surface and the empty capsid remaining outside the cell. It has been proposed that the formation of a unique portal near the site of receptor interaction, through which the genome can be delivered, is believed to be the strategy used by large dsDNA icosahedral viruses (Nandhagopal *et al.,* 2002; Xiao *et al.,* 2005).

An observation of note is the increase in size of the empty OtV-1 capsid when compared to an unattached capsid still containing genetic material. The empty capsid not only appears to have enlarged to some degree, but also has a less distinct hexagonal outline, appearing more rounded in shape and lacking regular symmetry. This loss in the integrity of the capsid shell, following the injection of viral nucleic acid contents into the host cell, is an intriguing phenomenon. Does it make the host less palatable for grazers? Previous studies have focussed mainly on the impact of viruses and grazers on bacterial populations in aquatic environments (Weinbauer and Peduzzi, 1995; Fischer *et al.,* 2006; Jacquet *et al.,* 2007; Steward *et al.,* 2007). The interactions observed are complex and many are still poorly understood (Miki and Jacquet, 2008). Does the continued attachment of the capsid to the cell membrane serve as an anti-viral mechanism? One hypothesis is the attached capsid prevents the subsequent attachment of other viruses to the host cell. This could be achieved by blocking a key receptor site in the cell membrane for virus attachment. Alternatively, by remaining attached, the

capsid structure induces a change in the cell membrane, which blocks all potential adsorption receptor sites for other viruses to enter the host.

Although many viruses are icosahedral when they initially bind to one or more receptor molecules on the cell surface, such an interaction is asymmetric (Bubeck *et al.,* 2005). A breakdown in symmetry and conformation of the original infecting virion is most likely initiated to prepare for membrane penetration and release of the viral genome. Structural studies on the human rhinoviruses show an enlargement in the empty capsid following attachment (Hewat *et al.,* 2002). This report showed that substantial conformational changes in copies of all the major capsid proteins of the capsid shell occur. Transformation of the capsid is necessary for release of the virus genetic material, a RNA genome, in this case. The capsid is seen to expand and rotate allowing a passage to open, through which the virus genome was allowed to pass into the host interior, thus exiting the capsid shell (Hewat *et al.,* 2002).

The entry of viral genetic material into the host cell is a key step in the infection process. The delivery of the genomes of tailed bacteriophages has been well-documented (Bamford *et al.,* 1995). The genomes of these phage are delivered into the host cytoplasm while the capsid and tail remain bound to the cell surface (Black, 1988). Recent studies show that the DNA packaged into a phage capsid is highly pressurised (tens of atmospheres) and this provides a force for the first step of passive injection of phage DNA into the host cell (Kindt *et al.,* 2001; Tzlil *et al.,* 2003). The dsDNA of the phage is packaged under high relative pressure due to the confinement and bending of the DNA chain inside the small capsid volume of approx 50-80 nm diameter (Evilevitch *et al.,* 2004). This internal pressure is believed to be responsible for the release of the

DNA after interaction with the receptor protein in the absence of an external source of energy (Evilevitch *et al.,* 2003; 2005). This method of DNA entry into the host cell is passive and does not require the action of enzymes. The capsid is seen to remain outside the cell and osmolarity is lower than that in the cytoplasm and therefore the stabilising force of DNA inside the capsid is much smaller than it would be if the capsid was directed into the cytoplasm (Castelnovo and Evilevitch, 2007). Such a scenario describes the full passive ejection of the viral genome into the host aided by the osmotic pressure gradient between the cytoplasm and the extracellular medium and may be the underlying mechanism of entry for the OtV-1 genome from capsid to cytoplasm. Higher resolution structural analysis and biochemical characterisation are clearly required to investigate this process further.

The question now arises as to the method of release of newly assembled viruses. Unfortunately, none of the TEM images generated captured a cell releasing a burst of viruses, as occurs in many lytic infections, or viruses budding off the host cell membrane, as occurs for example in the coccolithovirus, EhV-86 (Mackinder *et al.,* 2009, submitted). It is feasible the OtV viruses are released in a similar fashion to the release of chloroviruses from PBCV-1. This occurs by localised lysis of the host cell to release progeny virions by 6-8 hours p.i. (Van Etten *et al.,* 1983a)

A notable finding in this study was the observed regrowth of the *O. tauri* strain RCC 501, following initial lysis of the culture by the OtV-3 isolate. Several other microalgae have been reported to exhibit incomplete lysis and recovery from viral infection, including *Chlorella* sp. (Van Etten *et al.,* 1991), *Emiliania huxleyi* (Bidle *et al.,* 2007), and *Phaeocystis pouchetti* (Thyrhaug *et al.,* 2003; Haaber and Middelboe, 2009). For

example, an infected culture of *Emiliania huxleyi*, seen to contain replicating viruses, did not undergo complete lysis of the host population but instead was observed to recover and re-grow. No replicating viruses were observed in recovering cells (Bidle *et al* 2007).

Viral infection of certain microalgal cultures allows some host cells to survive and thus regrowth of the culture to occur within a few days (Waters and Chan, 1982; Nagasaki and Yamaguchi, 1998; Tarutani *et al.,* 2001; Tomaru *et al.,* 2004). Possible explanations include the presence of a resistant sub-population of host cells. In the case of the *Chlorella* sp. the survival of infected cultures was attributed to a pseudolysogenic condition, whereby the virus is present in a carrier state, with only a small fraction of the host population being infected continuously (Van Etten *et al.,* 1991). It is postulated this chronic infection would stimulate the production of specific or non-specific inhibitors that protect the host cells from further infection by competing with viral particles for receptor sites (Thyrhaug *et al.,* 2003). Cultures of *M. pusilla* were seen to be resistant to infection, with resistance persisting for several years in the absence of viral production (Zingone *et al.,* 2006). Observed resistance following infection by a particular viral strain did not afford protection to the host from infection by other viral strains. This would suggest a very specific protection mechanism in this particular host-virus system.

Observed regrowth of infected cultures may be a result of the production of reactive oxygen species (ROS). Oxidative bursts have previously been reported to eliminate pathogens in macroalgae, whereby production of ROS served a number of protective functions and ultimately led to programmed cell death (PCD) of infected cells (Jones

and Dangl, 1996). Photosynthesis has been seen to decrease following viral addition to microalgal cultures. The coccolithophorid *Emiliania huxleyi* was found to exhibit oxidative stress and reduced cell photosynthesis, due to disruption of photosynthesis by the generation of ROS, following virus addition (Evans *et al* 2006). ROS are widely implicated in apoptosis or PCD and elevated levels of oxidative stress have been reported in autocatalysed decline of a number of phytoplankton species, including *Peridinium* sp. (Vardi *et al.,* 1999) and *Emiliania huxleyi* (Evans *et al.,* 2006).

Viruses may induce apoptosis as part of the replication process. PCD may be employed in unicellular organisms as a defence mechanism to prevent massive viral infection and demise of clonal populations (Georgiou *et al.,* 1998). Viral gene products can employ diverse regulatory roles in apoptosis in animals (Teodoro and Branton, 1997) with a number of viruses inducing apoptosis and even specifically caspases. The viral replication process is believed to require activation of host metacaspases and a study on viral infection in *E. huxleyi* reported selective caspase inhibition led to severe reduction of viral replication (Bidle *et al.,* 2007). This has also been reported in several animal viruses (Best and Bloom, 2004). Caspase activity is required for replication and point mutation of caspase cleavage sequences on viral proteins abolishes viral replication. Virally encoded proteins containing caspase cleavage sequences have been identified in *Phycodnaviridae* genome sequences of EhV-86, PBCV-1 and EsV-1 (Bidle *et al.,* 2007). Notably, a number of these caspase cleavage-containing proteins represent a group of core genes shared among the NCLDVs, which may indicate this developed at an early stage of development of this viral lineage. Bidle *et al.,* (2007) showed there can be differential metacaspase activity within a clonal virus population, which may affect virus susceptibility. The virus requires activation of host metacaspases to

facilitate replication and production during infection. However, the activation and level of metacaspase expression during infection may differ within a host population infected by a clonal virus. Resistance may not be a fixed trait and may be reversed under physiological stress when metacaspases are expressed. Highly specific co-evolutionary processes in the activation and maintenance of metacaspases during viral infection have been proposed, contributing to the diversification and specificity of both hosts and viruses.

Another study reported a reduced competitiveness in virus-resistant algae, possibly related to their ability to acquire nutrients (Haaber and Middelboe, 2009). Regrowth was seen to occur at a much earlier stage in more nutrient-rich cultures than in others following lysis. In this report, *Phaeocystis pouchetti* was grown in the more nutrient-rich f/2 medium and the *O. tauri* strain RCC 501 grown in the present study was also cultured using f/2 medium. Further investigations are also required to confirm whether the RCC 501 strain of *O. tauri* in this study is unialgal and clonal. This would then discount the possibility that the regrowth seen was due to the presence of a second algal species or strain population. At present, it is only possible to speculate as the underlying mechanisms of algal resistance have not yet been determined.

4.5. Conclusion

A range of twelve virus isolates that infect the prasinophyte species' *O. tauri* and *M. pusilla* were isolated and characterised in this study. Characterisation of their morphology using TEM techniques revealed viruses with a icosahedral morphology and particle size ranges between 100 nm to 135 nm. The genome sizes for the isolated viruses were estimated to be between 190 to 195 kb. Phylogenetic analysis of the DNA polymerase genes of these viruses, coupled with their characterised morphological

features, indicated they can be putatively assigned to the *Prasinovirus* genus of the *Phycodnaviridae* family of viruses. Sequencing of the genomes of two viruses that infect two ecotype strains of *O. tauri* was undertaken and should yield further insight to the nature of these viruses. Resistance in one host strain of *O. tauri* was observed with regrowth of host cultures occurring after approximately 10 days following observed lysis. The underlying cause of the resistance is as yet undetermined and will require further investigation.

Chapter 5

The complete genome sequence and annotation of *Ostreococcus tauri* Virus-1, OtV-1

5.1. Introduction

The marine unicellular green alga *Ostreococcus tauri* belongs to the *Prasinophyceae*, a clade of the Chlorophyta group in the Planta kingdom (Courties *et al.,* 1998). The *Prasinophyceae* have diverged early at the base of the Chlorophyta and consequently of the green lineage (Bhattacharya and Medlin, 1998) and as such *O. tauri* holds a key phylogenetic position in the eukaryotic branch of the tree of life. Marine prasinophytes belong to a group of organisms generically known as the photosynthetic picoeukaryotes (PPEs). *O. tauri* is the smallest free-living eukaryote described to date (Chretiennotdinet *et al.,* 1995; Courties *et al.,* 1998). The *O. tauri* cell has a diameter size of less than 1µm, a naked plasma membrane with no cell wall and lacks scales or flagella. *O. tauri* has been isolated from different geographic regions and a range of depths (Diez *et al.,* 2001; Worden *et al.,* 2004).

Similar to certain cyanobacterial species, e.g. *Prochlorococcus*, *O. tauri* is reported to exhibit ecotype diversity among strains of the species. The importance of the ecological roles played by the marine cyanobacteria *Prochlorococcus* and *Synechococcus* in marine environments has been recognised, in terms of abundance and primary productivity. Over the past two decades, the diversity and ecophysiology of the cyanobacterium *Prochlorococcus* have been widely researched and detailed (Partensky, 1999). Part of the global success of *Prochlorococcus* has been attributed to the existence of distinct high- and low-light ecotypes, which occupy different niches and exploit different resources (Rocap *et al.,* 2003). *O. tauri* has been isolated from samples of various geographical origins and taken from a range of depths (Diez *et al.,* 2001; Worden *et al.,* 2004). Small-subunit rRNA sequences of *Ostreococcus* sp. from cultures and environmental samples cluster them into four different clades that are likely distinct enough to represent different species (Guillou *et al.,* 2004; Rodriguez *et al.,*

2005). Genetic distances between isolates do not appear to reflect geographic variation; instead, their genetic divergence is thought to be driven by light and nutrient gradients that give rise to physiological differences between surface and deep isolates (Rodriguez *et al.,* 2005; Cardol *et al.,* 2008; Six *et al.,* 2008)

This chapter describes the complete genome sequence of a novel algal virus, OtV-1, isolated during this study, which infects a high light strain, OTH95, of the picophytoplankton species *Ostreococcus tauri*. The host strain, OTH 95, belongs to the clade C, based on analysis of the ITS region of the SSU ribosomal RNA operon (Guillou *et al.,* 2004; Rodriguez *et al.,* 2005). OTH 95 is currently the sole member of Clade C strains of *Ostreococcus* and was sampled at a depth of between 0-5 m from the Thau lagoon (Table 3.2).

Other than a report of a transient *Ostreococcus*-like bloom off the eastern coast of the US, whose rapid decline was thought to be caused partially by viruses (O'Kelly *et al.,* 2003), there are no reports on the ecological role of *Ostreococcus*-specific viruses. Indeed, of the six *Phycodnaviridae* genera, the prasinoviruses and viruses of photosynthetic picoeukaryotes are the least represented in the literature, despite their global and ecological importance.

Data on the isolation, characterization and genome sequencing of a virus that infects *O. tauri* (OtV5) was recently reported (Derelle *et al.,* 2008). Characterisation of OtV5 revealed an 186,234 base pair dsDNA virus isolated by plaque assay, which is morphologically similar to the *Micromonas pusilla* viruses (Mayer and Taylor, 1979;

Waters and Chan, 1982; Cottrell and Suttle, 1995). Taxonomic analysis classified OtV5 as a prasinovirus in the algal virus family *Phycodnaviridae*.

OtV-1 is therefore only the second virus infecting this species and strain that has been completely sequenced. A third genome of a virus, OtV-2, that infects a low-light strain of *O. tauri* was also obtained in this study and is presented in chapter 6. These novel viral genome sequences provide further information to compare genome structure, replication strategy and conservation of core genes within the NCLDVs and will aid investigations into the diversity of viruses within the *Phycodnaviridae.*

5.2. Materials and Methods

5.2.1. Maintenance of algal host strain

The OtV-1 host, *Ostreococcus tauri* strain OTH 95 was grown in Keller (K) medium (Keller *et al.,* 1987) under the conditions described in Section 2.2.1.

5.2.2. Virus isolation

The OtV-1 virus was isolated from surface seawater collected on 06/09/2006 at the L4 sampling station in the Western English Channel (co-ordinates are 50º15'N, 04º13'W) following the method described in Section 2.2.3.

5.2.3. Clonal virus isolation of OtV-1

OtV-1 was purified by plaque assays, as described in Section 2.2.4.2.1. in order to obtain clonal virus stocks.

5.2.4. DNA preparation and sequencing

For preparation of large quantities of viruses for genome sequencing, ten litre volumes of *O. tauri* exponentially growing culture were inoculated with 10 ml of OtV-1 at a multiplicity of infection of one. Lysed cultures were passed sequentially through 0.8 µm and 0.2 µm filters to remove large cellular debris. Virus filtrates were concentrated by ultrafiltration to ~ 50 ml using a Quixstand benchtop system and hollow fibre cartridges with a 30, 000 pore size (NMWC) (GE Healthcare Amersham Biosciences). The 50 ml concentrate was then further concentrated to ~10 ml using the Mid-Gee benchtop system and hollow fibre cartridge with a 30, 000 pore size (NMWC) (GE Healthcare Amersham Biosciences). Aliquots (3 ml) of the concentrated OtV-1 lysate were purified by caesium chloride gradient purification as described in Section 2.2.5.2. Complete sequencing of the OtV-1 genome was performed at the Advanced Genomics Facility based at the University of Liverpool, UK.

5.2.5. Sequence assembly and finishing

The OtV-1 sample was sequenced as described in Section 2.2.10.3.

5.2.6. Genome annotation

The OtV-1 genome was annotated as described in Section 2.2.10.4.

5.2.7. Analysis of repeat regions

To identify repetitive sequences within the OtV-1 genome, a dot-plot analysis was performed using the LBDotView Version 1.0 (http://www.lynnon.com/dotplot.html) (Huang and Zhang, 2004). This analytical tool can compare one genome on the x-axis with a second genome on the y-axis. Here, the OtV-1 genome is compared against itself, enabling the precise location and orientation of homologous sequences to be located within the plot.

5.2.8. Phylogenetic analyses

Phylogenetic analyses were performed with the amino acid sequences of the DNA polymerase gene and intein. Phylogenetic trees were calculated from confidently aligned regions of homologous proteins by use of the PHYLIP v3.68 (Felsenstein, 2005) package. DNA polymerase sequences from various representative viruses and several environmental samples that were very likely to be related viruses were multiply aligned with the ClustalW program (Thompson, 1994). The designations and GenBank accession numbers of DNA polymerase sequences used for phylogenetic analysis Aare shown in Table 5.1 (scientific name, with abbreviation in parentheses, followed by the database accession number [referring to the National Center for Biotechnology Information database]).

Table 5.1. Details of DNA polymerase sequences used for phylogenetic analysis

Scientific name	**Accession number**
Alcephaline herpesvirus	NC 002531
Autographa californica nuclear polyhedrosis baculovirus (AcNPV)	P18131
Bombyx mori nucleopolyhedrovirus (BmNPV)	NP047469
Chilo iridescent virus (CIV)	CAC19195
Chrysochromulina brevifilum virus PW1 (CbV-PW1)	AAB49739
Chrysochromulina ericina (CeV)	ABU23716
Ectocarpus siliculosus virus (EsV-1	NP_077578
Emiliania huxleyi virus-86 (EhV-86)	AJ890364
Equid herpesvirus	YP_053075.1
Feldmania species virus (FsV)	AAB67116
Feldmannia irregularis virus (FirrV-1)	AAR26842
Fowlpox virus (FOWPV)	NP_039057
Heterosigma akashiwo virus (HaV)	AB194136.1
Human herpesvirus	CAA28464
Lymantria dispar nuclear polyhedrosis baculovirus (LdNPV)	BAA02036
Micromonas pusilla virus SP1 (MpV-SP1)	AAB66713
Micromonas pusilla virus-PL1	AAB49747
Micromonas pusilla virus-SG1	AAB49746
Mimivirus	YP_142676
Molluscum contagiosum virus (MOCV)	NP043990
Ostreococcus tauri virus 5 (OtV5)	ABY28020
Ostreococcus tauri virus-1 (OtV-1)	FN386611
Ostreococcus tauri virus-2 (OtV-2)	This study
Paramecium bursaria Chlorella virus (PBCV-1)	AAC00532
Paramecium bursaria Chlorella virus NY2A (PBCV-NY2A)	AAA88827
Phaeocystis globosa virus 01 (PgV-01)	ABD62757
Pyramimonas orientalis virus (PoV-01)	ABU23717
Vaccinia virus (VACV)	A24878

Intein sequence motifs of the protein splicing domain were identified according to the method of Pietrokovski (1998b). These six motifs contained a total of 75 conserved positions. All were present in all inteins, except for motif N4, which was missing from or divergent in some inteins. The *Drosophila melanogaster* hedgehog protein HINT domain was used as an outgroup for the inteins. The intein sequences used for phylogenetic analysis were all inserted into type B DNA polymerases. The organisms' scientific names, with intein names in parentheses, and database accession numbers (referring to the National Center for Biotechnology Information database) are shown in Table 5.2.

Table 5.2. Type B DNA polymerase intein allele sequences used for phylogenetic analysis

Scientific name	**Accession number**
D. melanogaster hedgehog HINT domain (HH DROME)	Q02936
Haloarcula marismortui ATCC 43049 (Hma)	YP_136425.1
Acanthamoeba polyphaga Mimivirus (APMV)	AAV50591
Chrysochromulina ericina virus (CeV)	A7U6F1
Heterosigma akashiwo virus (HaV01)	BAE06251
Ostreococcus tauri virus-1	This study
Methanococcus jannaschii (Mja Pol-1)	Q58295
Methanococcus jannaschii (Mja Pol-2)	Q58295
M. jannaschii (Mja TFIIB)	F64397
Pyrococcus sp. GBD (Psp-GBD Pol)	AAA67132.1
Thermococcus aggregans (Tag Pol-1)	CAA73475.1
T. aggregans (Tag Pol-2)	CAA73475.1
T. aggregans (Tag Pol-3)	CAA73475.1
Thermococcus fumicolans (Tfu Pol-1)	P74918
Thermococcus fumicolans (Tfu Pol-2)	P74918
T. hydrothermalis (Thy Pol-2)	CAC18555.1
Thermococcus kodakaraensis KOD1 (Tko Pol-1)	S71551
Thermococcus kodakaraensis KOD1 (Tko Pol-2)	S71551
Thermococcus litoralis (Tli Pol-1)	S42459
Thermococcus litoralis (Tli Pol-2)	S42459
Thermococcus peptonophilus strain SM2 (Tpe Pol)	E13953
Thermococcus sp. strain GE8 (Tsp-GE8 Pol-1)	CAC12850.1
Thermococcus sp. strain GE8 (Tsp-GE8 Pol-2)	Q9HH84
Thermococcus sp. strain GT (Tsp-GT Pol-2)	ABD14869

Details of the DNA topoisomerase II sequences used for phylogenetic analysis are shown in Table 5.3.

Table 5.3. DNA topoisomerase II sequences used for phylogenetic analysis

Scientific name	**Accession number**
Paramecium bursaria Chlorella virus AR158	YP_001498788.1
Paramecium bursaria Chlorella virus NY2A	YP_001497977.1
Paramecium bursaria Chlorella virus-1	NP_048939.1
Chlorella virus Marburg 1	AA495770.1
Paramecium bursaria Chlorella virus MT325	ABT14100.1
Paramecium bursaria Chlorella virus FR483	YP_001426181.1
Acanthamoeba polyphaga Mimivirus	YP_142834.1
Chilo iridescent virus	NP_149508.1
Emiliania huxleyi virus-86	YP_294202.1
African swine fever virus	P34203.1
Ostreococcus tauri virus-1	This study
Ostreococcus tauri	CAL56339
Ostreococcus lucimarinus	ABO99175

5.3. Results and Discussion

5.3.1. Description of the viral genome

The OtV-1 sequence has been deposited in the GenBank database (accession number FN386611). The linear genome OtV-1 was assembled into a contiguous sequence of 191,761 bp, which is in close agreement with the predicted size determined by pulsed-field gel electrophoresis (Section 4.3.5). A putative protein-coding region or CDS was defined as a continuous stretch of DNA that translates into a polypeptide that is initiated by an ATG translation start codon and extends for 65 or more additional codons. Using these criteria, 232 CDSs were identified in the ~190 kb genome of OtV-1. The CDS names were based on the virus name ***O****streococcus* ***t****auri* **V**irus-1, abbreviated to OtV-1. The CDSs were then numbered consecutively in the order in which they appeared in the

genome, following assembly and alignment with the reference genome, OtV5 (accession no. EU304328).

The OtV-1 genome has an overall G + C content of 45.36%, a coding density of 1.199 genes per kb and an average CDS length of 750 bases per gene (Table 5.4). CDSs are equally distributed on both strands (53% on the positive strand and 47% on the negative strand). Although no major repeat regions were identified, Dot Plot analysis (Lowe, 1997) which compared the genome with itself, identified the location and orientation of three small but distinct families of repeats (designated A, B and C) in the genome (Fig. 5.1). All three repeat families occur in coding regions of the genome. Family A repeats fall within a 1260 bp segment, (62,930 to 64,290 bp) and correspond to a GC-rich region (average 63.5%) within a predicted putative haemagluttinin protein product (OtV1_075). The remaining two repetitive regions occur between 93800 and 96600 bp (Family B repeats) and 99,090 and 100,850 bp (Family C repeats), respectively. Both these repeat regions fall within the same putative CDS, which encodes a putative viral A-type inclusion body (OtV1_117). All three repetitive regions consist of direct repeats. The orientation of the OtV-1 genome was compared to the published OtV5 genome and a high level of co-linearity was observed (Fig. 5.2). In addition, 87.5% i.e. 203 of 232, CDSs in OtV-1 have orthologs in OtV5, with only 29 CDSs being unique to OtV-1. A general comparison of OtV5 genome features with the OtV-1 genome also showed similar attributes (Table 5.4), with OtV5 having a comparable size (186,234 bp) and G+C composition (45.29%), though with a slightly smaller mean gene length (702 bp) giving rise to more predicted genes (268).

Table 5.4. Comparison of the general characteristics of the genomes of OtV-1 with OtV5 (Derelle *et al.,* 2008), *Ostreococcus tauri* strain OTH 95 (Derelle *et al.,* 2006), PBCV-1 (Li *et al.,* 1997), and EhV-86 (Wilson *et al.,* 2005b).

	OtV-1	OtV5	*O. tauri* OTH95	PBCV-1	EhV-86
Genome size (bp)	191,761	186,234	12,560,000	330,743	407,339
% G + C	45.36	45.29	58	40	40
Predicted CDSs (65 aa)	232	268	8,166	367	472
Mean gene size (bp/gene)	750	702	1300	899	863
% known genes	31.0	22	77	50	21
% hypothetical genes	13.0	20			
% genes with no matches to genes of known function	56.0	57	23	50	79

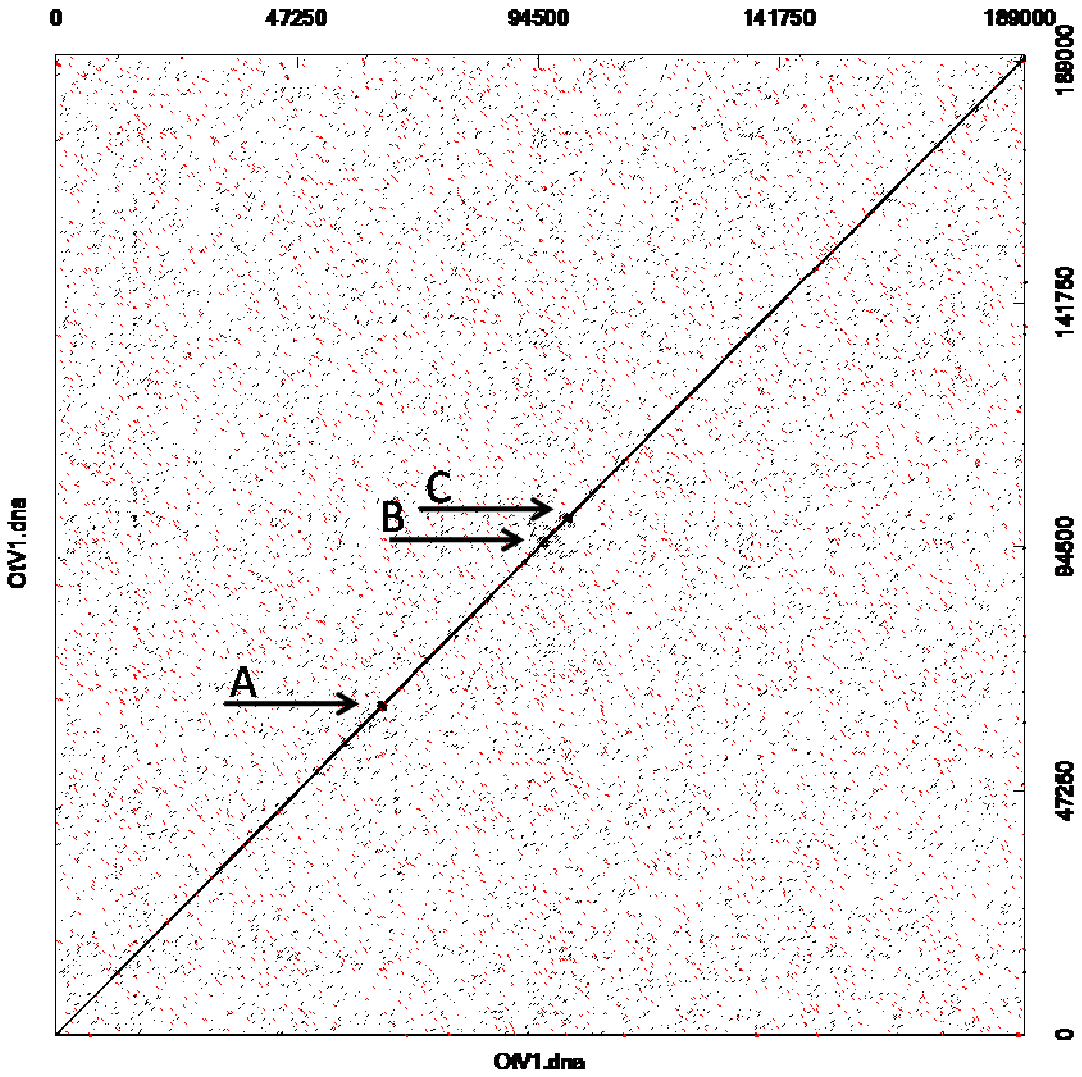


Figure 5.1. Dot Plot analysis (LBDot) of the full-length OtV-1 genome and grouping of repeat families. Family A are found between 63 kb and 64 kb, Family B and Family C are found between 94 kb and 97 kb.

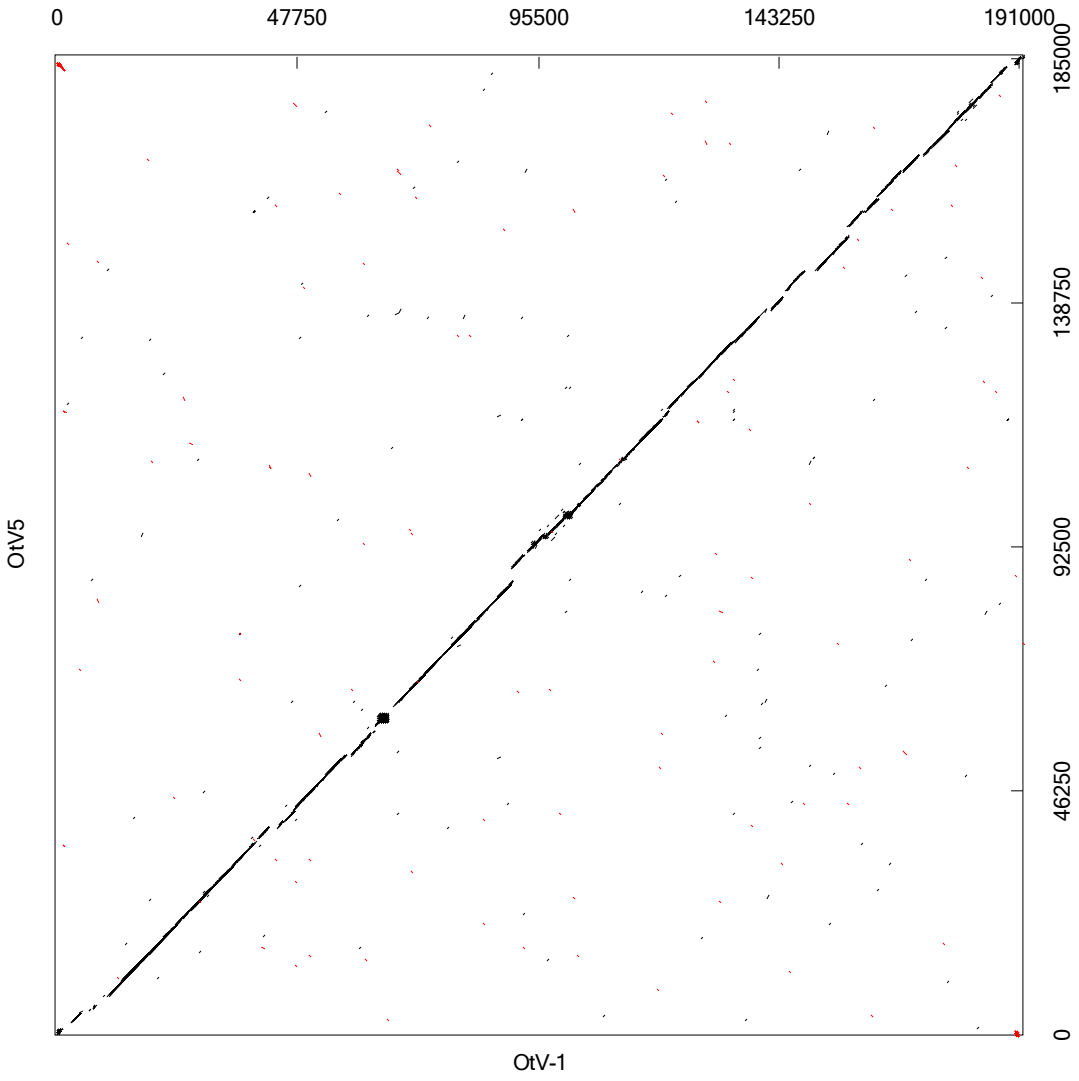


Figure 5.2. Dot Plot analysis (LBDot) of the full-length OtV-1 genome and the full-length OtV5 genome.

5.3.2. Identity of putative CDSs

Every CDS was compared with the non-redundant database at NCBI (http://www.ncbi.nlm.nih.gov/). The Pfam and COG databases were used to identify conserved domains and proteins within the OtV-1 CDSs. Of the 232 CDSs identified in the OtV-1 genome, 162 (69%) encode putative proteins where no function has been attributed (Table 5.4). Of these unknown CDSs, 32 share close similarity with other hypothetical proteins in databases. Therefore, 130 (56%) of the CDSs do not correspond to known proteins, highlighting a potentially high degree of functional novelty in this virus. Predicted metabolic functions of the remaining 73 CDSs (31%), which have homology to known proteins, were split into 9 functional groups (Table 5.6). Ten CDSs could be assigned putative functions involved in DNA replication, recombination and repair; six in nucleotide metabolism and transport; 7 in transcription; 19 in protein and lipid synthesis, modification and degradation; one in signalling; four in DNA methylation; 11 in sugar metabolism; 8 encoding capsid proteins; and a further 7 with miscellaneous functions. The OtV-1 genome is represented in a circular map in Figure 5.3.

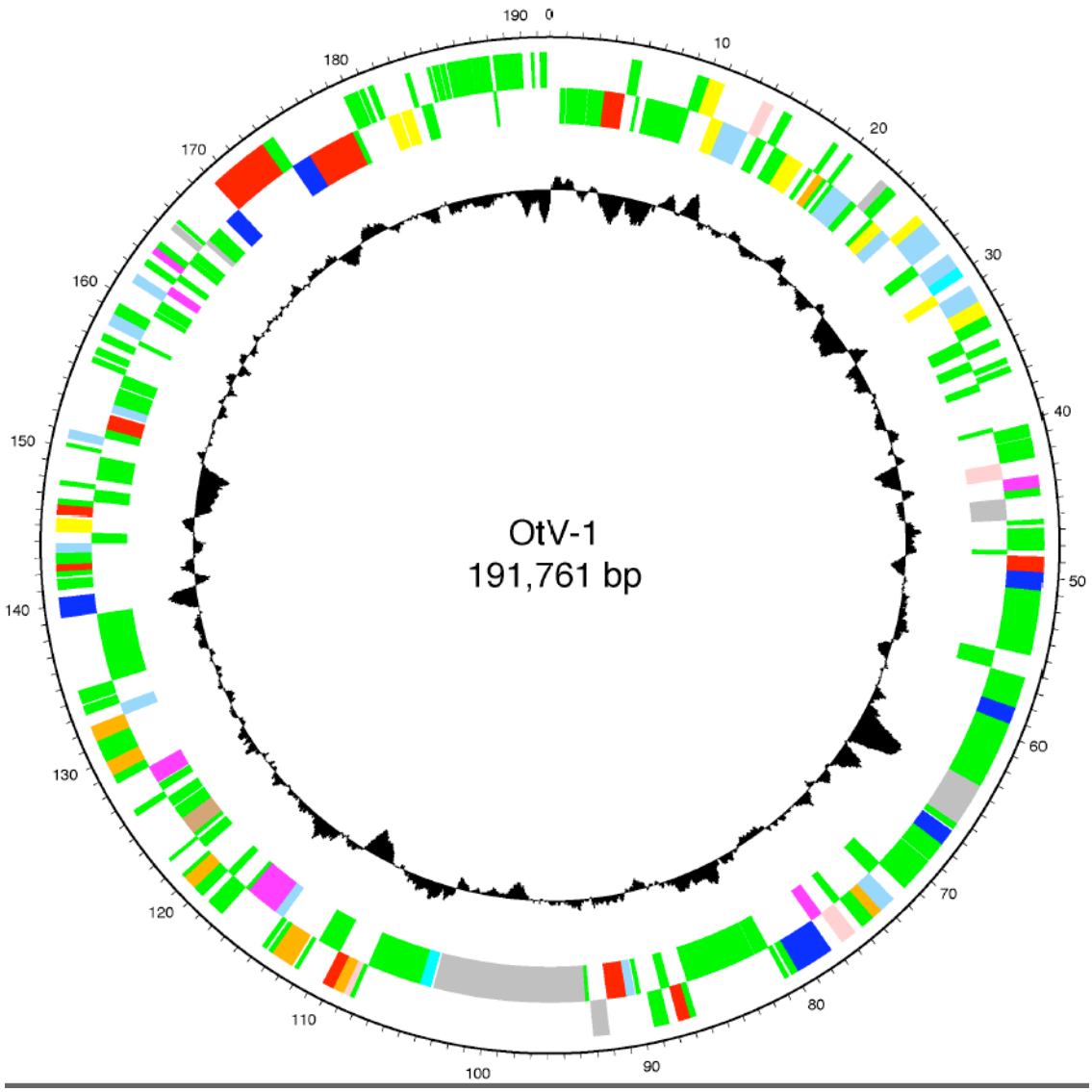


Figure 5.3. Circular representation of the 191,761 bp OtV-1 genome. The outside scale is numbered clockwise in kbp. Circles 1 and 2 (from the outside in) are CDSs (forward and reverse strands, respectively) starting with CDS OtV1_001 at position 689 bp (colour-coded green and just below the 1-kb marker on the reverse strand). The black inner line represents G+C content.

Key to colour code:

- No known function
- DNA replication, recombination and repair
- Protein synthesis and modification
- Nucleotide transport/metabolism
- Lipid manipulating enzymes
- Transcription
- Signalling
- Sugar manipulation
- DNA restriction/methylases
- Capsid proteins
- Miscellaneous

Table 5.5. Annotated coding sequences on the OtV-1 genome.

CDS	Start	End	Strand	OtV5 ortholog	Putative function/features	Top significant hit (*after OtV5)	E-value
OtV1_001	689	1054	-	264f	Unknown		
OtV1_002	1123	2523	-	162f	Unknown	*Rhodopirellula baltica**	2E-07
OtV1_003	2580	2813	-		Unknown		
OtV1_004	2840	3643	-	004r	Unknown		
OtV1_005	3673	5013	-		Exonuclease	*Pelodictyon phaeoclathratiforme*	5E-17
OtV1_006	5130	5771	+	147r	Unknown		
OtV1_007	5818	6081	-		Unknown		
OtV1_008	6515	7006	-		Unknown		
OtV1_009	7025	7435	-		Unknown		
OtV1_010	7453	8700	-		Unknown		
OtV1_011	8624	9373	+		Unknown		
OtV1_012	9406	10383	+	011f	GDP-mannose 4,6-dehydratase	*Synechococcus* sp.*	1E-92
OtV1_013	10350	11177	-	012r	Glycosyl transferase family 2	*Aquifex aeolicus**	2E-07
OtV1_014	11201	12940	-	013r	Cell division protein FtsH2	*Prochlorococcus marinus**	e-147
OtV1_015	12987	13691	+	014f	Methyltransferase FkbM	*Desulfovibrio desulfuricans**	1E-13
OtV1_016	13652	14110	-	015r	Unknown	*Dictyostelium discoideum**	1E-17
OtV1_017	14100	14351	-	016r	Unknown		
OtV1_018	14426	15001	+	017r	Unknown		
OtV1_019	14998	15720	-	018r	Unknown		
OtV1_020	15692	15967	-	019r	Unknown		
OtV1_021	15981	17189	-	020r	Glycosyl transferase, group 1	ATCV- 1*	1E-54
OtV1_022	17685	17996	-	022r	Unknown		
OtV1_023	18035	18403	+	023f	Unknown		
OtV1_024	18366	18869	-	024r	Transcription elongation factor IIS	ATCV-1 1*	2E-17
OtV1_025	18905	19141	-	025f	Unknown		
OtV1_026	19181	19504	+	026f	Unknown		
OtV1_027	19433	19723	-	027r	Unknown		
OtV1_028	19731	21251	-	028r	Asparagine synthetase B	*Gramella forsetii**	7E-91
OtV1_029	21270	21719	-	029r	Unknown		
OtV1_030	21761	22369	+	030f	Phosphate starvation-inducible protein, PhoH family protein	*Thermoanaerobacter pseudethanolicus**	8E-41
OtV1_031	22378	23052	+	031f	Unknown	Prochlorococcus phage P-SSM2*	2E-32
OtV1_032	22979	23323	-	032r	Unknown		
OtV1_033	23316	24080	-	033r	Alpha-1,2-galactosyltransferase	*Schizosaccharomyces pombe**	1E-11
OtV1_034	24104	24868	-	034r	TPR domain-containing protein	*Acaryochloris marina**	3E-23
OtV1_035	24867	25493	+	035f	Glycosyl transferase family 2	*Clostridium botulinum**	6E-08
OtV1_036	25527	27188	+	037f	Acetolactate synthase	*Methanosphaera stadtmanae**	5E-84
OtV1_037	27185	27931	-	038r	Unknown	*Chlorobaculum parvum**	2E-31
OtV1_038	27986	29002	+	039f	3-dehydroquinate synthetase	*Magnetospirillum gryphiswaldense**	3E-38
OtV1_039	28999	29643	+	040f	Oxo-acyl acyl carrier protein dehydrogenase	*Rhodopseudomonas palustris**	3E-30
OtV1_040	29634	30293	-	041r	NAD-dependent epimerase/dehydratase	*Geobacter lovleyi**	8E-21
OtV1_041	30326	31465	+	042f	DegT/DnrJ/EryC1/StrS aminotransferase family	*Trichomonas vaginalis* G3*	e-103
OtV1_042	31478	32413	+	043f	dTDP-D-glucose 4,6-dehydratase	*Nematostella vectensis**	2E-64
OtV1_043	32439	33179	+	044f	Unknown		
OtV1_044	33153	34007	-	045r	Unknown		
OtV1_045	34045	34515	+	047f	Unknown		
OtV1_046	34638	34934	-	048r	Unknown		
OtV1_047	35012	35287	-	049r	Unknown		
OtV1_048	35314	35676	+	050f	Unknown		

Table 5.5. Contd

CDS	Start	End	Strand	OtV5 ortholog	Putative function/features	Top significant hit (*after OtV5)	E-value
OtV1_049	35875	36261	+	051f	Unknown		
OtV1_050	36226	36732	-	052r	Unknown	PBCV_AR158*	6E-15
OtV1_051	36744	39256	-	053r	Unknown		
OtV1_052	39311	39592	-	056r	Unknown		
OtV1_053	39641	40438	+	057f	Unknown		
OtV1_054	40506	41177	+	058f	Unknown	PBCV-NY2A*	4E-18
OtV1_055	41190	41762	+		Unknown		
OtV1_056	41752	42807	-		DNA cytosine-methyltransferase	PBCV_AR158*	1E-58
OtV1_057	42870	43652	+	059f	FAD-dependent thymidylate synthase	*Dictyostelium discoideum**	3E-61
OtV1_058	43660	43896	+		Unknown		
OtV1_059	43925	44140	+		Unknown		
OtV1_060	44143	45627	-	062r	ABC-1 domain protein	*Synechococcus elongatus**	1E-65
OtV1_061	45675	45980	+	063f	Unknown		
OtV1_062	46210	47295	+	065f	Unknown		
OtV1_063	47292	47612	+	066f	Unknown		
OtV1_064	47609	47926	-		Unknown		
OtV1_065	48005	48967	+	067f	VV A18 helicase	PBCV-NY2A*	2E-58
OtV1_066	48984	50117	+	068f	Major capsid protein	*Pyramimonas orientalis* virus*	5E-39
OtV1_067	50125	54171	+	072f	Unknown		
OtV1_068	54408	55514	-	073r	Unknown	PBCV-1*	1E-26
OtV1_069	55586	55894	-	074f	Unknown		
OtV1_070	55906	56979	+	075f	Unknown		
OtV1_071	57007	57537	+	075f	Unknown		
OtV1_072	57563	58597	+	076f	Major capsid protein	*Pyramimonas orientalis* virus*	1E-42
OtV1_073	58817	62773	+	077f	Unknown	Nematostella vectensis*	1E-09
OtV1_074	62799	63365	+	080f	Hep_Hag haemagluttinin family protein	ATCV-1*	1E-17
OtV1_075	63314	64980	+	080f	Hep_Hag haemagluttinin family protein	ATCV-1*	1E-28
OtV1_076	64987	65388	+		Unknown	Vibrio campbellii	2E-07
OtV1_077	65517	66566	+	081f	Major capsid protein	*Pyramimonas orientalis* virus*	3E-47
OtV1_078	66580	66783	+	082f	Unknown		
OtV1_079	66809	67039	+	083f	Unknown		
OtV1_080	67073	67603	+	084f	Unknown		
OtV1_081	67605	67805	+	085f	Unknown		
OtV1_082	67884	69287	+	086f	Unknown	PBCV_AR158*	2E-21
OtV1_083	69292	69552	+	087f	Unknown		
OtV1_084	69553	69798	+	088f	Unknown		
OtV1_085	69823	70107	+	089f	Unknown		
OtV1_086	70051	70350	-	090r	Unknown		
OtV1_087	70378	71016	-	091r	Unknown	PBCV_AR158*	8E-28
OtV1_088	71065	72102	+	092f	N-myristoyltransferase	*Toxoplasma gondii**	7E-62
OtV1_089	72125	72787	+	093f	mRNA capping enzyme	*Encephalitozoon cuniculi**	2E-12
OtV1_090	72784	73212	+	094f	Metal dependent hydrolase	APMV*	1E-03
OtV1_091	73231	73650	+	095f	Unknown		
OtV1_092	73632	74132	-	096r	Unknown		
OtV1_093	74187	75212	+	097f	D12 class N6 adenine-specific DNA methyltransferase	*Campylobacter curvus**	1E-59
OtV1_094	75196	76029	-	098r	VV A32 virion packaging ATPase	*Heterosigma akashiwo* virus*	5E-64
OtV1_095	76041	77192	+	099f	Major capsid protein	*Pyramimonas orientalis* virus*	9E-45
OtV1_096	77208	78491	+	100f	Major capsid protein	*Heterosigma akashiwo* virus 01*	3E-37

Table 5.5. Contd

CDS	Start	End	Strand	OtV5 ortholog	Putative function/features	Top significant hit (*after OtV5)	E-value
OtV1_097	78495	78944	+	101f	Unknown		
OtV1_098	79129	79425	+	103f	Unknown		
OtV1_099	79415	80395	-	104r	Unknown		
OtV1_100	80447	81583	-	105r	Unknown	ATCV-1*	2E-10
OtV1_101	81586	81987	-	107r	Unknown		
OtV1_102	82015	82692	-	108r	Unknown	ATCV-1*	2E-18
OtV1_103	82710	83150	-	109r	Unknown	PBCV_FR483*	7E-07
OtV1_104	83226	83447	-	110r	Unknown		
OtV1_105	83448	84302	-	111r	Unknown	PBCV_AR158*	2E-24
OtV1_106	84336	84713	-	112r	Unknown	PBCV-NY2A*	1E-05
OtV1_107	84713	85279	-	113r	Unknown		
OtV1_108	85370	85753	+	114f	Unknown		
OtV1_109	85779	86525	+	115f	PCNA	PBCV-NY2A*	1E-36
OtV1_110	86500	87099	-	116r	Unknown		
OtV1_111	87111	88055	+	117f	Unknown	ATCV-1*	1E-50
OtV1_112	88475	88756	-	119r	Unknown	Bacterium Ellin 514*	9E-09
OtV1_113	88864	89421	-	126r	Prolyl 4-hydroxylase	APMV*	2E-27
OtV1_114	80465	90745	-	130r	NTPase/helicase	*Trichomonas vaginalis* G3*	6E-13
OtV1_115	90864	91877	+	131f	Echinonectin	*Nematostella vectensis**	2E-36
OtV1_116	91946	92185	-	132r	Unknown		
OtV1_117	92218	102414	-	135r	Virus A-type inclusion body	*Trichomonas vaginalis* G3*	1E-89
OtV1_118	102502	103263	-	136r	Patatin-like phospholipase	APMV*	3E-14
OtV1_119	103272	103664	-		Unknown		
OtV1_120	103667	104089	-	137r	Unknown		
OtV1_121	104095	104472	-	138r	Unknown		
OtV1_122	104510	104962	-		Unknown	PBCV_AR158	4E-05
OtV1_123	104987	105502	-	140r	Unknown	PBCV-MT325*	4E-05
OtV1_124	105552	106913	-	141r	Unknown		
OtV1_125	106936	107199	+	143f	Unknown		
OtV1_126	107232	107630	+	144f	DNA adenine methyltransferase	*Feldmannia* species*	2E-11
OtV1_127	107653	108375	+	145f	Rnase III	PBCV-1*	4E-50
OtV1_128	108351	109067	+	146f	Exonuclease (PBCV-1)	ATCV-1*	2E-37
OtV1_129	109028	110581	-	147r	Unknown	*Emiliania huxleyi* virus 86*	6E-38
OtV1_130	110811	111062	+	148f	Unknown		
OtV1_131	111245	112696	+	149f	SWF/SNF family helicase	PBCV-1*	9E-69
OtV1_132	112696	112977	+		Unknown		
OtV1_133	113150	113461	+	151f	Unknown		
OtV1_134	113458	114069	-	152r	33 kDa in vitro translation peptide	PBCV-1*	2E-09
OtV1_135	114098	116341	-	153r	Ribonucleotide reductase, ls	*Microscilla marina**	e=0
OtV1_136	116371	116601	-	154r	Unknown		
OtV1_137	116640	117047	+	155f	Unknown	APMV*	1E-07
OtV1_138	117073	117384	+	156f	Unknown	*Microscilla marina**	2E-05
OtV1_139	117405	117773	+	157f	Unknown		
OtV1_140	117739	118227	-	158r	Unknown		
OtV1_141	118271	118966	+	162f	Unknown		
OtV1_142	119007	119864	+	161f	TATA-box binding protein (TBP)	APMV*	1E-10
OtV1_143	119881	120111	+	163f	Unknown		
OtV1_144	120442	121128	-	165r	Unknown	PBCV_AR158*	8E-22
OtV1_145	121168	121380	+		Unknown		
OtV1_146	121377	121649	-	167r	Unknown		
OtV1_147	121666	122892	-	168r	Serine/Threonine protein kinase	*Herpetosiphon aurantiacus**	4E-08
OtV1_148	122905	123318	-	169r	Unknown		
OtV1_149	123353	123757	-	170r	Unknown		
OtV1_150	123885	124577	-	171r	33 kDa in vitro translation peptide	ATCV-1*	8E-23
OtV1_151	124657	125073	+	172f	Unknown		

Table 5.5. Contd

CDS	Start	End	Strand	OtV5 ortholog	Putative function/features	Top significant hit (*after OtV5)	E-value
OtV1_152	125162	125686	-		Unnamed protein product	*Ostreococcus tauri*	4E-08
OtV1_153	125787	126956	-	174r	Ribonucleotide reductase, ss	*Ostreococcus lucimarinus**	e-110
OtV1_154	127001	127546	+	176f	Unknown		
OtV1_155	127586	128524	+	177f	Transcription factor TFIIB	APMV*	2E-18
OtV1_156	128531	129259	+	178f	Unknown	PBCV-NY2A*	3E-57
OtV1_157	129278	129907	+	179f	Unknown	PBCV-FR483*	1E-04
OtV1_158	129930	130886	+		mRNA capping enzyme	PBCV-NY2A	9E-28
OtV1_159	130810	131607	-	181r	Ubiquitin-specific protease	*Arabidopsis thaliana**	9E-13
OtV1_160	131664	132260	+	183f	Unknown		
OtV1_161	132377	132637	+	184f	Unknown		
OtV1_162	132648	133238	+		Unknown		
OtV1_163	133235	133579	-	185r	Unknown		
OtV1_164	133586	135487	-	187r	Unknown		
OtV1_165	135512	137455	-	188r	Unknown	PBCV-1*	e-100
OtV1_166	137491	137769	-	189r	Unknown		
OtV1_167	137868	139136	+	190f	Major capsid protein	PBCV-NY2A*	e-101
OtV1_168	139581	140282	+		Unknown	APMV	5E-12
OtV1_169	140440	140781	+	191f	Unknown		
OtV1_170	140782	141222	+	192f	Ribonuclease H	*Candidatus Nitrospira defluvii**	4E-27
OtV1_171	141222	141908	+	193f	Unknown		
OtV1_172	141932	142528	+	194f	2OG-Fe(II) oxygenase	Uncultured bacterium*	2E-17
OtV1_173	142531	143211	-		Unknown		
OtV1_174	143221	144069	+	197f	6-phosphofructokinase	*Bacteroides thetaiotaomicron**	3E-18
OtV1_175	144265	145011	+	199f	DNA topoisomerase I	*Marine actinobacterium**	1E-14
OtV1_176	145011	145367	+	200f	Unknown	*Providencia stuartii**	1E-06
OtV1_177	145364	146170	-	201r	Glycosyl transferase, family 25	Alpha proteobacterium*	1E-03
OtV1_178	146207	146500	+	202f	Unknown		
OtV1_179	147022	147540	-		Unknown (putative endonuclease?)	PBCV-FR483	1E-16
OtV1_180	147510	148040	-	147r	Unknown		
OtV1_181	148280	148492	+	148f	Unknown		
OtV1_182	148749	149336	+	203f	Aspartyl/asparaginyl beta-hydroxylase	*Pseudomonas putida**	5E-14
OtV1_183	149333	149902	-	204r	Unknown	*Oryza sativa**	8E-17
OtV1_184	149919	150899	-	205r	DNA ligase	*Trypanosoma brucei**	8E-29
OtV1_185	151070	151684	-		ATP-dependent protease proteolytic subunit	*Emiliania huxleyi* virus-86	3E-21
OtV1_186	151662	152753	-	209r	Unknown		
OtV1_187	152885	153697	-	212r	Unknown	*Burkholderia phytofirmans**	1E-07
OtV1_188	153708	154073	+	213f	Unknown	PBCV-1*	8E-10
OtV1_189	154255	154719	+	215f	Unknown		
OtV1_190	154664	155140	-	216r	Unknown		
OtV1_191	155177	155656	+	219f	Unknown		
OtV1_192	155924	156202	-		Unknown		
OtV1_193	156232	156993	+	221f	Proline dehydrogenase	*Ostreococcus lucimarinus**	9E-45
OtV1_194	157017	157361	+	222f	Unknown	*Ostreococcus tauri**	6E-06
OtV1_195	157383	157742	+	223f	Unknown		
OtV1_196	157779	158087	-		Unknown		
OtV1_197	158169	158534	-	224r	Unknown	PBCV AR158*	2E-13
OtV1_198	158558	158869	-	225r	Unknown		
OtV1_199	158893	159501	+	226f	Aspartyl/asparaginyl beta-hydroxylase	*Acaryochloris marina**	7E-07
OtV1_200	159475	160026	-	227r	Thymidine kinase	*Vitis vinifera**	4E-34
OtV1_201	160202	160630	+	228f	Unknown		
OtV1_202	160614	161027	-	230r	Unknown	*Ostreococcus lucimarinus**	4E-16
OtV1_203	161257	161505	+	231f	dUTP pyrophosphatase	*Saccharomyces cerevisiae**	3E-25

Table 5.5. Contd

CDS	Start	End	Strand	OtV5 ortholog	Putative function/features	Top significant hit (*after OtV5)	E-value
OtV1_204	161567	161968	+	232f	Unknown		
OtV1_205	161965	162306	-	233r	Unknown		
OtV1_206	162368	162922	-	234r	Unknown		
OtV1_207	162959	163396	+	235f	Fibronectin-binding A-like protein	*Ostreococcus tauri**	8E-19
OtV1_208	163497	163730	+	236f	Unknown		
OtV1_209	163709	164059	-	237r	Rhodanese-like domain protein	*Clostridium ramosum**	5E-14
OtV1_210	164071	165072	-	238r	Unknown		
OtV1_211	165731	166831	-	239r	Major capsid protein	*Pyramimonas orientalis* virus*	5E-47
OtV1_212	166886	179728	+	240f	DNA polymerase	*Chrysochromulina ericina* virus*	e-131
OtV1_213	170715	171080	+	241f	Unknown		
OtV1_214	171191	171478	+	242f	Unknown	PBCV-1*	8E-18
OtV1_215	171727	172803	-	243r	Major capsid protein	*Pyramimonas orientalis* virus*	7E-07
OtV1_216	172816	176034	-	244r	DNA topoisomerase II	*Ostreococcus tauri**	e=0
OtV1_217	176064	176402	-	245r	Unknown		
OtV1_218	176483	177412	+	246f	Unknown		
OtV1_219	177512	177835	+	247f	Unknown		
OtV1_220	178048	178452	+	248f	Unknown		
OtV1_221	178614	179393	-	252r	Glycosyltransferase, group 1	*Synechococcus* sp. WH 780*	2E-10
OtV1_222	179526	180455	-	252r	Glycosyltransferase, group 1	*Jannaschia* sp.*	1E-07
OtV1_223	180493	180804	+		Unknown		
OtV1_224	180946	181686	-	254r	Unknown		
OtV1_225	181899	182123	+	255f	Unknown		
OtV1_226	182237	182656	+	256f	Unknown		
OtV1_227	182734	183054	+	257f	Unknown		
OtV1_228	183165	183419	+		Unknown		
OtV1_229	183446	184072	+	259f	Unknown		
OtV1_230	184082	184657	+	260f	Unknown		
OtV1_231	184698	185195	+	261f	Unknown	*Ostreococcus tauri**	3E-32
OtV1_232	185210	185827	+	262f	Unknown		
OtV1_233	185828	186031	-	263r	Unknown		
OtV1_234	186069	187856	+		Unknown		
OtV1_235	188429	188626	+	266f	Unknown		
OtV1_236	188980	189402	+		Unknown		

ATCV-1=*Acanthocystis turfacea Chlorella virus*-1; APMV=*Acanthamoeba polyphaga* Mimivirus

Table 5.6. Predicted proteins encoded by the OtV-1 genome grouped by function.

DNA, RNA, replication, recombination and repair	CDS#	**Protein, amino acid and lipid synthesis/modification**	CDS#
DNA polymerase	OtV1_212	Prolyl 4-hydroxylase alpha-subunit	OtV1_113
ATP dependent DNA ligase	OtV1_184	2OG-Fe(II) oxygenase	OtV1_172
DNA topoisomerase I	OtV1_175	Ubiquitin C-terminal hydrolase	OtV1_159
DNA topoisomerase II	OtV1_216	ATP-dependent protease proteolytic subunit	OtV1_185
PCNA	OtV1_109	FtsH2 metalloprotease	OtV1_014
Exonuclease	OtV1_005	33 kDa in vitro translation peptide	OtV1_134
Exonuclease	OtV1_128	33 kDa in vitro translation peptide	OtV1_150
RNase H	OtV1_170	Aminotransferase family protein	OtV1_041
VV A18 Helicase	OtV1_065	Zinc metallopeptidase	OtV1_090
NTPase/helicase	OtV1_114	3-dehydroquinate synthase	OtV1_038
Nucleotide transport and metabolism		Acetolactate synthase	OtV1_036
Ribonucleotide reductase, large subunit	OtV1_135	N-myristoyltransferase	OtV1_088
Ribonucleotide reductase, small subunit	OtV1_153	TPR domain-containing protein	OtV1_034
dUTP pyrophosphatase	OtV1_203	Proline dehydrogenase	OtV1_193
VV A32 virion packaging ATPase	OtV1_094	Patatin phospholipase	OtV1_118
Thymidine kinase	OtV1_200	Oxo-acyl carrier protein dehydrogenase	OtV1_039
FAD dependent thymidylate synthase	OtV1_057	Asparagine synthetase	OtV1_028
Transcription		Aspartyl/Asparaginyl beta-hydrolase	OtV1_182
RNA transcription factor TFIIB	OtV1_155	Aspartyl/Asparaginyl beta-hydrolase	OtV1_199
RNA transcription factor TFIIS	OtV1_024	**Capsid protein**	
RNase III	OtV1_127	Capsid protein 1	OtV1_066
SW1/SNF helicase domain protein	OtV1_131	Capsid protein 2	OtV1_072
TATA-box binding protein	OtV1_142	Capsid protein 3	OtV1_077
mRNA capping enzyme	OtV1_089	Capsid protein 4	OtV1_095
mRNA capping enzyme	OtV1_158	Capsid protein 5	OtV1_096
Sugar manipulation enzymes		Capsid protein 6	OtV1_167
Alpha-1,2-galactosyl transferase	OtV1_033	Capsid protein 7	OtV1_211
GDP-D-mannose dehydratase	OtV1_012	Capsid protein 8	OtV1_215
Glycosyltransferase	OtV1_013	**Miscellaneous**	
Glycosyltransferase	OtV1_021	Virus inclusion body	OtV1_117
Glycosyltransferase	OtV1_035	Rhodanese domain-containing protein	OtV1_209
Glycosyltransferase	OtV1_177	ABC1 domain protein	OtV1_060
Glycosyltransferase	OtV1_221	Fibronectin-binding protein	OtV1_207
Glycosyltransferase	OtV1_222	Hep_Hag haemagluttinin family protein	OtV1_075
6-phosphofructokinase	OtV1_174	Phosphate starvation inducible protein	OtV1_030
NAD-dependent epimerase/dehydratase	OtV1_040	Echinonectin	OtV1_115
dTDP-D-glucose 4,6-dehydratase	OtV1_042	**Signalling**	
DNA restriction/methylation		Serine/Threonine protein kinase	OtV1_147
Adenine-specific methyltransferase	OtV1_093		
DNA-cytosine methyltransferase	OtV1_056		
Adenine-specific methyltransferase	OtV1_126		
Methyltransferase FkbM	OtV1_015		

5.3.3. NCLDVs

5.3.3.1. Comparison of OtV-1 with other NCLDVs

OtV-1 is smaller than most other phycodnaviruses and sequenced NCLDVs (Table 5.4). A total of 15 Group I, II and III core genes was identified in the OtV-1 genome (Table 5.7). The presence/absence of these genes in five phycodnaviruses was subsequently determined. Recently, whole genome comparisons of NCLDVs (Iyer *et al.,* 2001; 2006) found nine genes (Group I) are shared by genomes from all NCLDV family members. A further eight genes (Group II) were found in all five NCLDV virus families but were missing in one or more lineages within those families. Finally, 14 Group III genes were conserved in three of the five families of NCLDVs and Group IV genes were conserved in two families of NCLDVs. However, fewer CDSs, including the core set of conserved virus genes for NCLDVs, were found in the OtV-1 genome when compared to the majority of NCLDVs sequenced to date, with the exception of the recently sequenced phaeovirus, FsV-158 (Schroeder *et al.,* 2009) (Table 5.7). Hence, homologues were identified for only 5 of the 9 Group I genes, 6 of the 8 Group II genes and 4 of the 14 Group III genes in the OtV-1 genome (Table 5.7).

Table 5.7.
Presence of NCLDV Core Genes (Groups 1, 2 and 3) in Various NCLDV Genomes

	1										2								3														
Family Species	Vaccinia virus D5-Type ATPase	DNA polymerase (*otv212f*)[a]	Vaccinia virus A32-Type ATPase (*otv94r*)	Vaccinia Virus A18-Type Helicase (*otv67f*)	Capsid protein (*otv66f/72f/77f/95f/96f/167f/211r/215r*)	Thiol-oxidoreductase	Vaccinia Virus D6R-Type Helicase	Serine/Threonine Protein Kinase (*otv147r*)	VLTF2-Like Transcription Factor	TFII-Like Transcription Factor (*otv124f*)	MuT-Like NTP Pyrophosphohydrolase	Myristyolated Virion Protein A	Proliferating Cell Nuclear Antigen (*otv109f*)	Ribonucleotide Reductase, Large Subunit (*otv135r*)	Ribonucleotide Reductase, Small Subunit (*otv153r*)	Thymidylate Kinase	dUTPase (*otv153f*)	A494R-Like Uncharacterised Protein	RuvC-Like Holliday Junction Resolvase	BroA-Like	Capping Enzyme (*otv94f/158f*)	ATP-dependent Ligase (*otv184r*)	RNA Polymerase, Subunit 1	RNA Polymerase, Subunit 2	Thioredoxin/Glutaredoxin	Serine/Threonine Phosphatase	BIR Domain	Virion-Associated Membrane Protein	Topoisomerase II (*otv216r*)	SW1/SNF2 Family Helicase (*otv131f*)	RNA Polymerase, Subunit 10
Phycodnaviridae																																	
OtV-1		✓	✓	✓	✓			✓		✓			✓				✓				✓	✓							✓	✓			
EhV-86	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓		✓	✓	✓		✓	✓			✓	✓	✓		✓		✓		✓	✓			✓
PBCV-1	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓			✓	✓		✓	✓		✓		✓	✓			
EsV-1	✓	✓	✓	✓	✓		✓	✓	✓	✓			✓	✓	✓		✓	✓	✓	✓				✓		✓							
FsV-158	✓	✓	✓		✓		✓		✓	✓			✓	✓	✓									✓									
Mimiviridae																																	
Mimivirus	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
Iridoviridae																																	
LCDV	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓				✓	✓					✓		✓			
IIV-6	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Asfarviridae																																	
ASFV	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓			✓	✓	✓			✓	
Poxviridae																																	
VACV	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓				✓
MOCV	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓				✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓				
AMEV	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
MSEV	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		✓			

NOTE – Based on data from this study and Bawden eta l (2000), Iyer, Aravind and Koonin (2001), and Raoult *et al.,* (2004). ASFV, African swine fever virus; NTP, nucleoside triphosphate; dUTPase, deoxyuridine triphosphatase; ATP, adenosine triphosphate. [a] Gene ID for OtV-1.

5.3.3.2. Redefinition of NCLDV core genes?

A high proportion of the CDSs of the OtV-1 genome bear close similarity to bacterial-type genes. It has been hypothesized that NCLDV genomes undergo successive accretions of bacterial genes, perhaps via bacteria-feeding hosts (Filee *et al.,* 2007). The recent sequencing of the complete genome of the phaeovirus FsV-1 (Schroeder *et al.,* 2009) and the prasinoviruses, OtV5 (Derelle *et al.,* 2008) and OtV-1 (this study), allows a review of the current identification and classification of groups of core NCLDV genes. The genes for Vaccinia Virus (VV) D5 type ATPase, thiol oxidoreductase, VV D6R-type helicase and VLFT2-like transcription factor are absent from the OtV-1 genome. Neither were they reported in the OtV5 genome. In addition, the genome of FsV-158 does not have a VV D6R-type helicase. These findings have implications for the present classification system of core genes. The Group I core genes absent from OtV-1 could possibly now be regarded as Group II, and not Group I, core NCLDV genes. Furthermore, the presence of genes for a Proliferating Cell Nuclear Antigen (PCNA) and ribonucleotide reductase large and small subunits, in all sequenced genomes of members of the *Phycodnaviridae* could reclassify these genes as Group I core genes instead of Group II core genes.

5.3.4. Genes unique to the OtV-1 genome

Among the 30 unique genes in OtV-1, i.e. those not found in OtV5 (Table 5.5), are genes which encode a potential exonuclease (OtV1_005) and DNA cytosine-methyltransferase (OtV1_056). The cytosine-methyltransferase shares close homology (~40% similarity at the amino acid level) with the equivalent protein in PBCV-1. With an additional two adenine-specific methyltransferases and one cytosine-specific methyltransferase, the OtV-1 genome encodes four methyltransferases in total (Table 5.6). The level of methylation of the OtV-1 genome is unknown. These DNA

methyltransferases lack obvious companion site-specific endonucleases within the OtV-1 genome.

5.3.5. Host related genes

A total of 11 CDSs in the OtV-1 genome share close similarity to genes found in the host genus, *Ostreococcus* (Table 5.8). High identity scores were seen for some of these, including genes that encode a putative dTDP-D-4,6-glucose dehydratase (OtV1_042), a proline dehydrogenase (OtV1_193) and a DNA topoisomerase II (OtV1_216) (Table 5.8). The OtV-1 genome also encodes a ribonucleotide reductase small subunit (OtV1_153), which shares 62% identity with the equivalent gene in *Ostreococcus lucimarinus.* One of the two mRNA capping enzymes encoded by OtV-1 (OtV1_089) shares 26% identity with a mRNA capping enzyme encoded by both *O. lucimarinus* and *O. tauri.* A second mRNA capping enzyme (OtV-1_158) shares 30% identity to that encoded by the chlorovirus PBCV-1, and therefore bears a closer homology to the ancestral NCLDV mRNA capping enzyme. In addition, a fibronectin-like binding protein (OtV1_207) shares homology (~40% amino acid identity) with a gene in both *O. lucimarinus* and *O. tauri.* Four of the eleven host-like genes are of unknown function.

Phylogenetic analysis of the aligned amino acid sequences of the active sites of DNA topoisomerase II genes of several viruses and eukaryotes, including *Ostreococcus* sp., revealed that *Ostreococcus tauri* viruses form a distinct cluster with strong bootstrap support (Fig. 5.4). The eukaryotic algal sequences, including host *Ostreococcus* genes, form a distinct cluster from all virus sequences (Fig. 5.4). This phylogeny suggests the DNA topoisomerase II gene arose in viruses in the evolutionary distant past, possibly prior to the divergence of the green algal lineage.

Table 5.8. Summary of putative CDSs with database homology to host genes.

OtV-1 CDS#	Hit to host[a]	Putative function	Identity	No. of identical amino acid residues	E-value
OtV1_042	OL	dTDP-D-4,6-glucose dehydratase	44%	135/305	$1e^{-65}$
OtV1_128	OT (OL)	PBCV-1 exonuclease	40% (31%)	74/223(66/213)	$2e^{-21}$ ($6e^{-19}$)
OtV1_152	OT	Unknown	32%	26/61	$1e^{-08}$
OtV1_153	OL	Ribonucleotide reductase, ss	62%	190/302	$1e^{-09}$
OtV1_158	OL (OT)	mRNA capping enzyme	26% (26%)	86/320(84/320)	$5e^{-20}$ ($2e^{-19}$)
OtV1_193	OL(OT)	Proline dehydrogenase	40% (49%)	112/276(75/153)	$1e^{-42}$ ($3e^{-27}$)
OtV1_194	OT(OL)	Unknown	43% (47%)	34/74(21/44)	$1e^{-5}$ (0.011)
OtV1_202	OL(OT)	Unknown	31% (50%)	50/159(32/78)	$1e^{-15}$ ($1e^{-08}$)
OtV1_207	OT(OL)	Fibronectin-binding protein	42% (40%)	54/128(52/129)	$2e^{-18}$ ($2e^{-17}$)
OtV1_216	OL(OT)	DNA topoisomerase II	44% (46%)	508/1126(485/1032)	0 (0)
OtV1_231	OT(OL)	Unknown	47% (39%)	82/198(68/172)	$4e^{-31}$ ($2e^{-30}$)

[a] OL = *Ostreococcus lucimarinus;* OT = *Ostreococcus tauri*

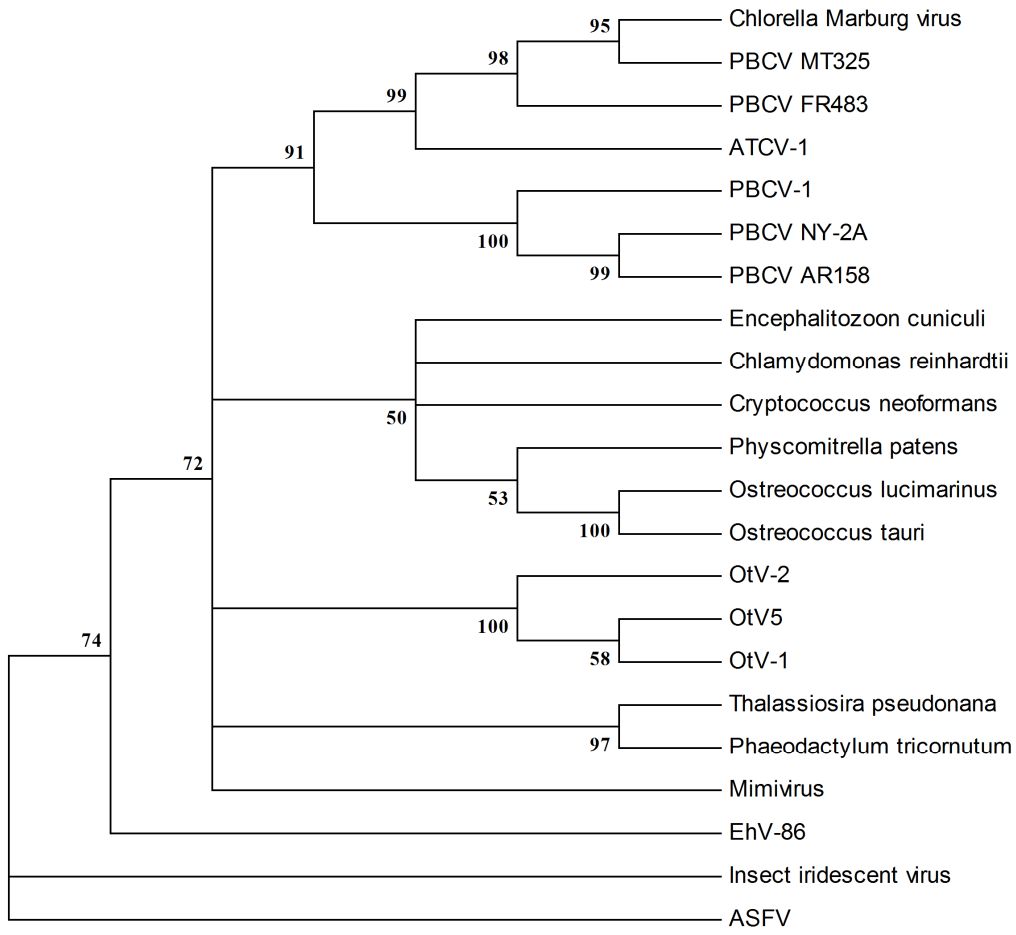


Figure 5.4. A phylogenetic inference based on a distance matrix algorithm between the DNA topoisomerase II gene fragments of *Ostreococcus tauri* viruses and other large dsDNA viruses and eukaryotic organisms [Neighbour in PHYLIP, version 3.68 (Felsenstein, 2005)]. The alignment was performed (ClustalW) on the regions spanning conserved domains of the DNA topoisomerase II gene encoding the active sites of the protein. Numbers at the nodes indicate bootstrap values. Abbreviations for virus names and GenBank accession numbers of DNA polymerase sequences used for phylogenetic analysis are given in Materials and Methods Section 5.2.8.

5.3.4. tRNAs

The OtV-1 genome was analysed for tRNAs using the tRNAscan-SE program (Lowe, 1997) (http://lowelab.ucsc.edu/tRNAscan-SE/) and is predicted to encode four tRNAs: encoding amino acids Ile, Gln, Asn, and a pseudo tRNA. All four tRNAs contain introns and are clustered in a region of the OtV-1 genome between 144,113 and 148,693 bp (Table 5.9).

Table 5.9. Summary of the features of the OtV-1 tRNA genes.

tRNA#	Start	End	tRNA Type	Anti codon	Intron Begins	Intron Ends	Intron length (nts)
1	144113	144209	Ile	TAT	144150	144174	24
2	146519	146605	Gln	TTG	146556	146570	14
3	146751	146847	Asn	GTT	146787	146812	25
4	148646	148728	Pseudo	AGT	148682	148693	11

5.3.5. Structural proteins

The OtV-1 genome encodes eight putative capsid proteins (Table 5.6). After a top BLASTP hit to OtV5, five of the capsid proteins have closest similarity to *Pyramimonas orientalis* virus PoV-1, two have closest similarity to *Heterosigma akashiwo* virus HaV-1, and one has closest similarity to a capsid protein encoded by PBCV NY-2A (Table 5.5). A summary of the features of each of the putative capsid protein genes of OtV-1 is shown in Table 5.10. Phylogenetic analysis of these proteins with other putative phycodnavirus major capsid protein (MCP) genes showed capsid protein 6 of the OtV-1 genome forms a distinct cluster with the phycodnavirus MCP gene sequences (Fig. 5.5), indicating capsid protein 6 is likely to be the MCP in OtV-1. This would require further experimental analysis to confirm.

Table 5.10. Summary of feature statistics of putative capsid protein genes of OtV-1.

	CDS #	Molecular weight (kDa)	Gene Length(amino acids)	G+C content (%)
Capsid 1	OtV1_066	43.0	377	39.0
Capsid 2	OtV1_072	38.0	344	44.83
Capsid 3	OtV1_077	38.5	349	55.52
Capsid 4	OtV1_095	43.8	383	41.75
Capsid 5	OtV1_096	49.0	427	43.53
Capsid 6	OtV1_167	46.7	422	56.34
Capsid 7	OtV1_211	40.0	366	48.60
Capsid 8	OtV1_215	41.6	358	40.48

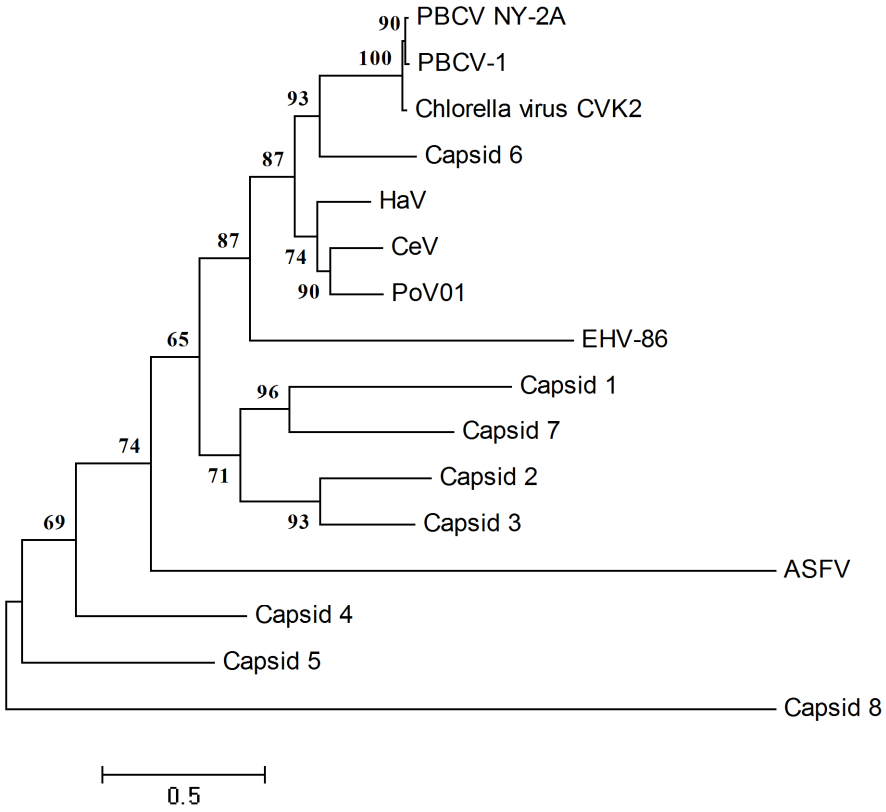


Figure 5.5. A phylogenetic inference based on a distance matrix algorithm between the 8 capsid protein gene fragments of OtV-1 and the major capsid protein genes of other large dsDNA viruses [Neighbour, version 3.68 (Felsenstein, 2005)]. Abbreviations for virus names and GenBank accession numbers of DNA polymerase sequences used for phylogenetic analysis were as follows (scientific name, with abbreviation in parentheses): African Swine Fever Virus (ASFV), ABL67020.1.; *Chlorella* virus CVK2, BAA76600.1; *Chrysochromulina ericina* virus (CeV), ABU23712; *Emiliania huxleyi* virus-86 (EhV-86), CAI65508.2; *Heterosigma akashiwo* virus (HaV), BAE06835.1; *Paramecium bursaria Chlorella virus* NY-2A (PBCV NY-2A), YP_001497781.1; *Paramecium bursaria Chlorella virus* (PBCV-1), AAA88828.1; *Pyramimonas orientalis* virus (PoV01), ABU23714.1; Capsids 1-8 of OtV-1 isolated in this study. The scale bar indicates a distance of 0.5 fixed mutations per amino acid.

5.3.6. DNA replication and repair-associated proteins

The OtV-1 genome encodes 10 proteins that are involved in DNA replication, recombination or repair (Table 5.6). Thus, there are proteins encoding both types of DNA topoisomerases i.e. Types I and II. Topoisomerases introduce temporary single- or double-stranded breaks in DNA, which help to resolve problems associated with DNA topology during its replication, transcription and recombination (Champoux, 2001). Type I topoisomerases (ATP- independent) function by passing one strand of the DNA through a break in the opposite strand, whilst type II topoisomerases are adenosine triphosphatases (ATPases) and introduce a double-stranded gap (Roca, 1995). With the notable exception of Poxviridae, many dsDNA viruses (including NCLDVs and phages) encode their own type II DNA topoisomerase. Accordingly, OtV-1 exhibits a CDS (OtV1_216) 1072 amino acids in length with a predicted molecular mass of 121 kDa, that shares 43% amino acid identity to several eukaryotic topoisomerases. The OtV-1 protein is similar in size to one found in the chlorovirus PBCV-1, which encodes one of the smallest topoisomerases, with a molecular mass of 120 kDa, compared to 160-180 kDa in most other eukaryotes (Lavrukhin *et al.,* 2000). Moreover, the PBCV-1 enzyme cleaves dsDNA approximately 30 times greater than the human type II DNA topoisomerase (Fortune *et al.,* 2001). However, the smallest type II topoisomerase, a polypeptide 1058 amino acids in length, is found in the *Chlorella* Pbi virus CVM-1. This enzyme has a DNA cleavage capacity 50-fold greater than the human-type II topoisomerase (Dickey and Osheroff, 2005).

The wide distribution of the DNA topoisomerase II enzyme, i.e. from *Chlorella* viruses to viruses EhV-86, APMV Mimivirus and now in all the OtV viral genomes sequenced to date (Derelle *et al.,* 2008; this study) suggests this enzyme plays a key functional role. The high DNA cleavage activity of the PBCV-1 DNA topoisomerase may indicate

a possible role in viral recombination (Dickey *et al.,* 2005). Interestingly, the OtV-1 type II DNA topoisomerase shows 50% amino acid identity to the host *Ostreococcus tauri* enzyme and 45% identity to the *Ostreococcus lucimarinus* ortholog (e value=0; Table 5.8). Phylogenetic analysis of the DNA topoisomerase II gene sequences of a number of eukaryotic algae and NCLDVs shows that the host algae form distinct clusters separate from the algal viruses (Fig. 5.4). It is possible then, that this gene has been acquired by horizontal gene transfer, but the direction remains unclear. The clustering of algal homologues into two distinct clades, suggests that this was no recent event and the topoisomerase gene must have transferred from host to virus, or vice versa, in the evolutionary distant past.

The OtV-1 genome also encodes a protein that resembles a Proliferating Cell Nuclear Antigen (PCNA) protein (OtV1_109). PCNA interacts with proteins not only involved in DNA replication but also DNA repair and post-replicative processing, such as DNA methyltransferases and DNA transposases (Warbrick, 2000). As the OtV-1 genome encodes proteins involved in both DNA repair and DNA methylation, this poses an interesting question as to whether the PCNA protein also interacts with the encoded methyltransferase proteins.

5.3.7. DNA polymerase gene of OtV-1

In the process of sequencing the genome of OtV-1, a putative DNA polymerase gene (OtV1_212) was identified by its similarity to other DNA polymerase genes. BLASTP searches in the databases revealed close similarities with the DNA polymerase gene of other phycodnaviruses, and more distantly related species (data not shown). Phylogenetic analysis of highly conserved regions common to viral DNA polymerases (Fig. 5.6) showed that most algal viruses are more closely related to each other than to

other groups of analysed viruses. Within this group, OtV-1 clusters with strong bootstrap support (100% bootstrap value for 1,000 trials) to the OtV5 and OtV-2 viruses, as well as the *Micromonas pusilla* virus. This microalgal cluster is distinct from viruses infecting macroalgae (EsV-1, FsV and FirrV-1) and from *Chlorella* viruses (PBCV-1 and NY-2A) (Fig. 5.6). Integrating these results with previous observations on the basic characteristics of this virus suggests OtV-1 should be assigned to the family *Phycodnaviridae.*

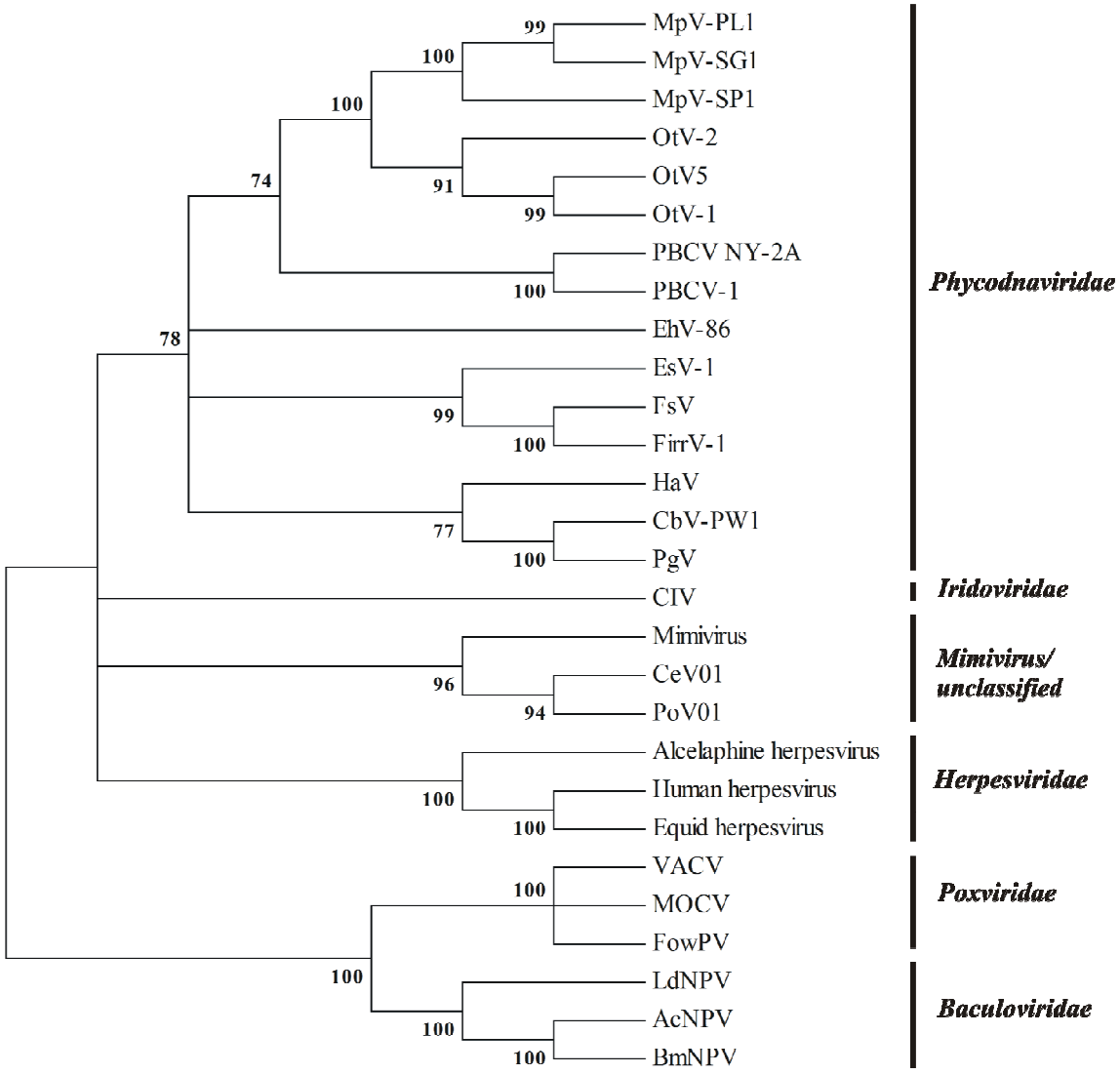


Figure 5.6. A phylogenetic inference based on a distance matrix algorithm between the DNA *pol* gene fragments of the family *Phycodnaviridae* and other large dsDNA viruses [Neighbour in PHYLIP, version 3.68 (Felsenstein, 2005)]. The alignment was performed (ClustalW) on the region spanning the highly conserved regions I and IV (Villarreal and DeFilippis, 2000) of the DNA *pol* genes. Abbreviations for virus names and GenBank accession numbers are described in the Materials and Methods section 5.2.8.

.

5.3.8. Characteristics of OtV-1 intein

The highly conserved motif I of the OtV-1 DNA polymerase domain, YGDTDS, which includes a catalytic-site residue, is interrupted at its centre (YGD/TDS) by a 329 amino-acid insert sequence (Fig. 5.7). The Intein Database and Registry (InBase) (http://www.neb.com/neb/inteins.html) was used to confirm this insertion as an intein. An intein (internal protein) is a protein sequence that is inserted in coding regions of other proteins and is translated with them. The intein is then autocatalytically spliced out to form the mature product, whilst simultaneously ligating the two flanking segments of the host protein (Perler *et al.,* 1997). Therefore, two proteins are derived from a single translational product, i.e. the free intein sequence and the mature form of the host protein, with the N-terminal and C-terminal external proteins, or exteins, ligated by a peptide bond. Although the excision of the intein from the precursor protein and the ligation of the flanking exteins are mediated primarily by the intein itself, the C-extein also plays an important role. Inteins are believed to be rarities as the majority of species do not carry any known inteins.

The location of the intein identified in OtV-1 is termed the Pol-c intein integration point and includes inteins in the type B polymerases of several archaeal DNA polymerase genes (YAD/TDG) (Niehaus, 1997; Waters *et al.,* 2003), as well as in the recently described type B polymerase of Mimivirus (Raoult *et al.,* 2004) and in the type B polymerase of the raphidovirus HaV (YGD/TDS) (Nagasaki *et al.,* 2005a). Inteins integrated at the same point in homologous proteins are termed intein alleles and are often more similar to each other than to other inteins, due to vertical transmission and horizontal transfer (Perler *et al.,* 1997; Pietrokovski, 2001).

Many inteins contain a homing endonuclease (EN). Inteins which include such a domain are referred to as full-length inteins. Most of the homing EN domains belong to the LAGLIDADG family (Chevalier, 2001; Gogarten *et al.,* 2002). Homing EN domains are believed to promote the spread of the intein through the gene pool of the host species via a recombination process or 'homing'. As most inteins have no known function, they are considered 'selfish' or parasitic elements (Pietrokovski, 2001). However, due to the efficiency of their removal from the host protein, their effect on the host phenotype is minimal (Liu *et al.,* 2003).

The first intein of viruses infecting eukaryotes was identified in the ribonucleotide reductase gene of Chilo iridescent virus (CIV), an icosahedral cytoplasmic insect iridovirus (Pietrokovski, 1998b). Very recently, a DNA polymerase intein was described in the huge dsDNA amoeba-infecting Mimivirus (Raoult *et al.,* 2004). Identical intein elements were also identified in the DNA polymerase genes of the raphidovirus, *Heterosigma akashiwo* virus (HaV), which infects the bloom-forming raphidophyte *H. akashiwo* (Nagasaki *et al.,* 2005a). The Pol-C allelic intein described in OtV-1 can now be included as one of the few identified inteins of viruses infecting eukaryotic algae.

Inteins seemingly undergo cycles of gaining and losing EN activity (Gimble, 2001; Matthew *et al.,* 2001). Inteins can exist without this activity but it does lend a major advantage, as it mediates the horizontal transfer of the elements to unoccupied insertion points or homing sites. The HaV intein was identified as a remnant of a typical EN homing domain and as such is unable to promote its own homing (Nagasaki *et al.,* 2005a). In contrast, this signature motif of the intein was present in OtV-1 and further

functional work could reveal the current potential of this intein to actively transfer to other members of the population.

The OtV-1 intein includes all of the motifs characterising a functional autocatalytic self-splicing domain (Perler, 2002), including the residues that provide the nucleophilic groups in the self-splicing reactions (Pietrokovski, 1998a; Perler, 2002) (Fig. 5.7). All of the identified protein-splicing domains are at typical positions within the intein, including the signature dodecapeptide motif (EN1, Fig. 5.7) found in endonucleases in the LAGLIDADG family (Chevalier, 2001). The presence of three additional conserved regions (EN2, EN3, EN4, Fig. 5.7) in the OtV-1 DNA *pol* intein, all indicate an active homing endonuclease.

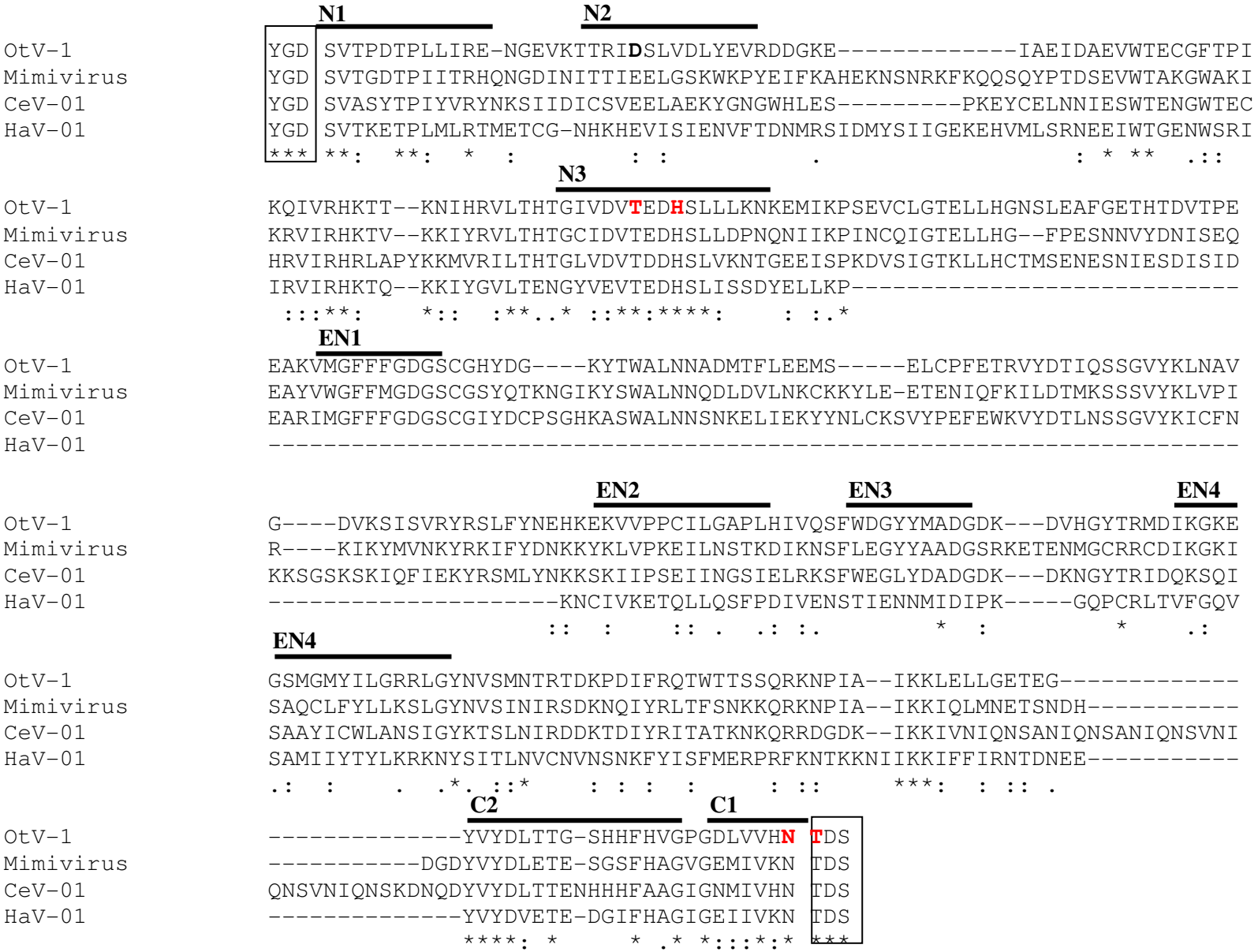


Figure 5.7. Amino acid sequence of the DNA polymerase B inteins and their immediate flanks of *Ostreococcus virus* 1 (OtV-1), Mimivirus, *Chrysochromulina ericina* virus - 01 (CeV-01) and *Heterosigma virus* 01 (HaV-01). Extein flanks are boxed. Sites corresponding to intein motifs (Pietrokovski 1998a) are overlined and named. Letters in red indicate active conserved residues essential for protein splicing.

Phylogenetic analysis of the OtV-1 intein, inteins from the InBase Registry and DNA polymerase I motif C inteins from viral and archaeal isolates resulted in three distinct clusters (Fig. 5.8). Inteins from OtV-1 and phycodnavirus-like DNA *pol* genes formed one cluster, while inteins from halophilic and thermophilic archaeal isolates each formed monophyletic clusters. OtV-1 is now the fourth dsDNA virus described to exhibit an intein. A nested PCR technique for amplifying a conserved region of the DNA polymerases of *Phycodnaviridae* designed by Chen & Suttle (1995a; b) did not work in OtV-1 (data not shown), as reported for a number of other algal viruses in previous studies (Chen and Suttle, 1995a; Sandaa *et al.,* 2001). In the case of OtV-1, this was presumably due to the intein insertion. No intein was observed in the DNA polymerase of OtV5. Perhaps more surprising though is the absence of any intein in a more extensive study recently performed on nearly 30 *Ostreococcus* virus isolates (Bellec, 2009). Despite this, it would appear inteins are possibly more commonplace in marine algal viruses than previously believed and have been identified in DNA polymerase, RNA polymerase and ribonucleotide reductase genes in more than one virus type (Raoult *et al.,* 2004; Nagasaki *et al.,* 2005a; present study; M. Allen pers. comm.; Matthias Fischer, pers. comm.). This raises the question how are DNA polymerase I motif C inteins acquired by members of the *Phycodnaviridae*? Possible modes of transmission include an intein inserting into a site in a genome from an alternative site within the same genome. Sequence analysis of inteins found in type B DNA polymerase I, however, shows that these inteins are specific for their insertion site (Ogata *et al.,* 2005). Another possibility is the exchange of inteins between viral and host genomes with similar insertion sites. Sequence data for a wider range of genes and viruses may reveal that viruses can act as reservoirs of intein elements and play a role in their transmission.

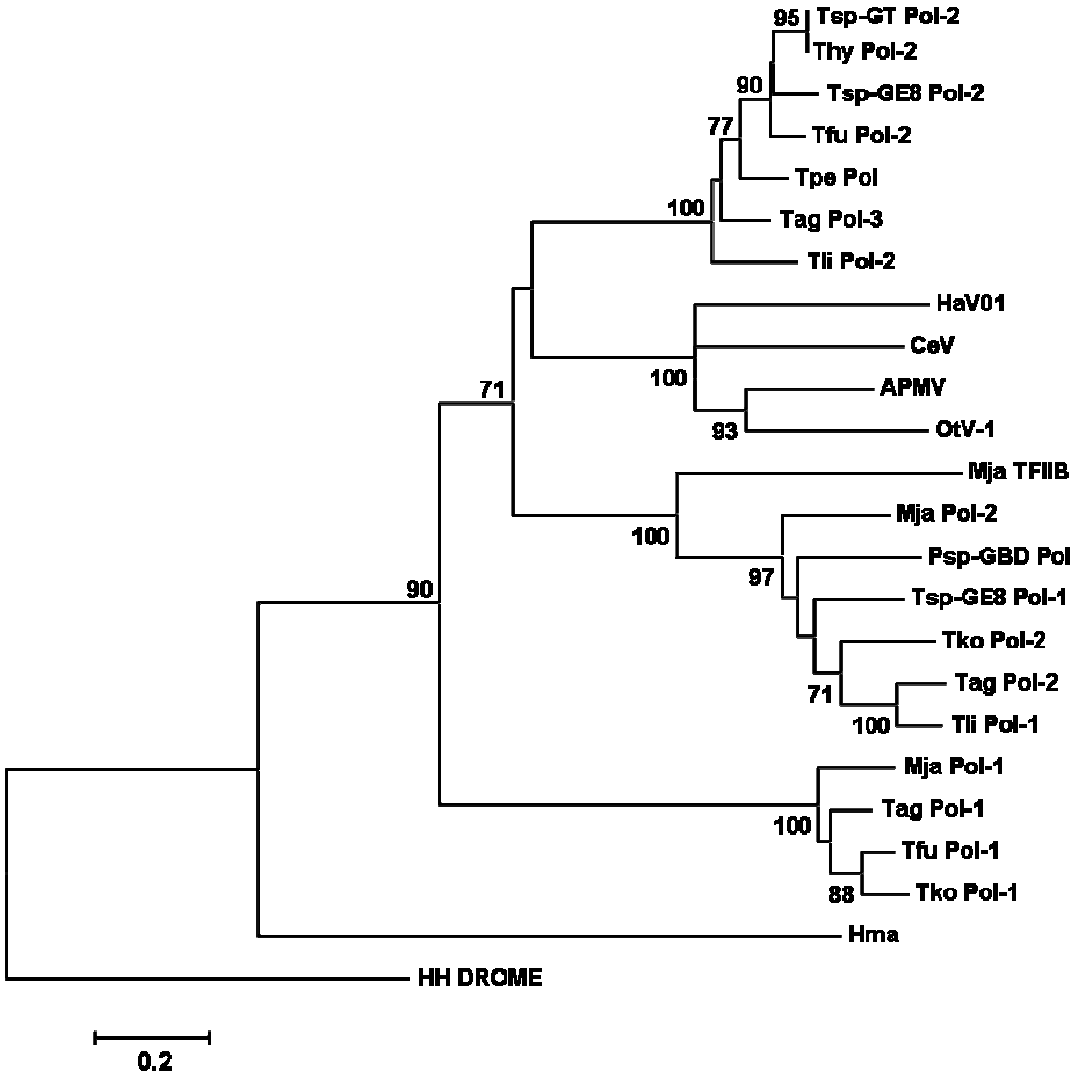


Figure 5.8. Phylogenetic neighbour-joining tree calculated from confidently aligned regions of the amino acid sequences of type B DNA polymerase intein alleles. The tree was rooted by using the *Drosophila* hedgehog HINT domain. Numbers at the nodes indicate bootstrap values below 60% were collapsed. The Gen Bank accession numbers of all sequences are provided in Materials and Methods Section 5.2.8. The scale bar indicates a distance of 0.2 fixed mutations per amino acid position.

5.3.9. Nucleotide metabolism-associated proteins

Due to their large genome size, NCLDVs usually encode several deoxynucleotide synthesis enzymes to ensure the provision of sufficient quantities of nucleotides for their replication. Viral DNA synthesis most likely requires higher concentrations of dNTPs than the host can supply, so large quantities of dNTPs must be synthesised *de novo* by both viral and host encoded proteins. Indeed, OtV-1 encodes six CDSs related to nucleotide metabolism enzymes (Table 5.6) including both ribonucleotide reductase subunits (OtV1_135 and OtV1_153), one VV A32 ATPase (OtV1_094), thymidine kinase (OtV1_200), dUTPase (OtV1_203) and a FAD-dependent thymidylate synthase (OtV1_057). Thymidine kinase is a universal enzyme that catalyses the ATP-dependent phosphorylation of thymidine. These enzymes, involved in DNA precursor metabolism, are known to be important in other members of the *Phycodnaviridae*, most notably the chloroviruses, which have been extensively characterised (Van Etten, 2003). For example, in PBCV-1, the concentration of DNA in the host cell has been reported to increase by 4-10 fold, four hours after infection due to viral DNA synthesis (Van Etten *et al.,* 1984; 2003). The enzyme dUTPase hydrolyses dUTP to dUMP and pyrophosphate. dUMP is a substrate for thymidylate synthase and is required for *de novo* synthesis of thymidylate (dTMP), an essential DNA precursor. This enzyme is also required to avoid the incorporation of deoxyuridine into newly synthesised DNA. The gene for dUTPase is ubiquitous in eukaryotes, eubacteria, and archaea and is also found in a number of retroviruses and DNA viruses, where viral dUTPases may help control local dUTP levels in order to enhance viral replication (Baldo and McClure, 1999).

OtV-1 lacks a traditional thymidylate synthetase A. Instead it encodes a protein that is a member of a new family of flavin-dependent thymidylate synthetases called ThyX

(Myllykallio *et al.,* 2002) (OtV1_056; Table 5.5). This enzyme is also found in chloroviruses but has not been reported in any other member of the *Phycodnaviridae*. Thymidylate (dTMP) is an essential DNA precursor. There are two pathways for thymidylate synthesis, both of which utilise a different thymidylate synthase, namely ThyA and ThyX (Myllykallio *et al.,* 2002). Both enzymes, through the methylation of dUMP, convert dUMP to dTMP. However, there is no sequence identity or structural similarity between these two enzymes. Although both ThyA and ThyX depend on methylenetetrahydrofolate for activity, their reductive mechanisms differ markedly. Only ThyX uses FAD as a cofactor and has a highly conserved sequence motif, RHRX_7S, or ThyX motif, as well as overall significant amino acid sequence similarity to a class of alternative synthases called ThyX. ThyX proteins have been found in approximately 25% of sequenced archaeal and bacterial genomes (Myllykallio *et al.,* 2002), including many pathogenic bacteria (Liu and Yang, 2004), a number of bacteriophages (Bhattacharya *et al.,* 2008) but only certain double-stranded DNA viruses (Graziani *et al.,* 2004; 2006). Numerous ThyA homologs have been analysed from viral sources but equivalent data has not been reported for viral ThyX proteins to date. Biochemical studies indicate that ThyX proteins bind FAD and act as NAD(P)H oxidases in the presence of bound dUMP. This would indicate that viral infection requires FAD from the host. As reported for the PBCV-1 genome, OtV-1 also lacks a canonical ThyA.

5.3.10. Transcription-associated proteins

OtV-1 encodes at least two putative transcription factor-like elements: TFIIB (OtV1_155) and TFIIS (OtV1_024). OtV-1 also encodes two proteins that are involved in creating a mRNA cap structure: both are mRNA capping guanylyltransferases (OtV1_089 and OtV1_158). The CDS OtV1_158 does not have a homolog in OtV5 but

has closest amino acid identity (30%) with the *Chlorella* virus PBCV-NY2A_B148R (Fitzgerald *et al.,* 2007b). The first two steps in the capping of cellular mRNAs are catalyzed by the enzymes RNA triphosphatase and RNA guanylyltransferase. The mRNA cap structure is a defining feature of eukaryotic mRNA and is required for mRNA stability and efficient translation. Expression of mRNA capping enzymes has been reported during immediate early infection (onset at 10 minutes post-infection) of chloroviruses (Kawasaki *et al.,* 2004).

OtV-1 does not encode any RNA polymerase genes, a feature in accordance with previous reported sequenced genomes of the PBCV-1 and EsV-1 lineages (Van Etten, 2003; Wilson *et al.,* 2009). In contrast, the genome of the coccolithovirus EhV-86 encodes five RNA polymerase subunits (Wilson *et al.,* 2005b). The smaller phycodnaviruses, chloroviruses and phaeoviruses, as well as the prasinoviruses described in this study, most likely underwent lineage-specific gene losses in their genomes, leading to the disappearance of a host-independent transcription apparatus. This would result in a major propagation change for the virus, which, through the loss of the RNA polymerase subunits, would become nuclear-dependent, not independent, for transcription (Allen *et al.,* 2006b).

In previous studies, the chlorovirus, PBCV-1, was reported to encode the smallest functional ATP-dependent DNA ligase from a eukaryotic system (Ho *et al.,* 1997) with a length of 298 amino acids. The OtV-1 ATP-dependent DNA ligase (OtV1_184) is similar in size to its homologue, OtV5, and has a length of 326 amino acids. It shares approximately 30% identity with DNA ligases from a wide range of eukaryotic organisms.

OtV-1 encodes a putative TATA-box binding protein (OtV1_142). A TATA binding protein transcription factor binds specifically to the TATA box sequence, which is usually found 25-30 bases upstream of the transcription start site in some eukaryotic gene promoters. It helps to position RNA polymerase II over the start of the transcription site. Since no recognisable RNA polymerase or RNA polymerase components have been detected in the OtV-1 genome, infectious viral DNA may target the nucleus utilising host RNA polymerase(s) to initiate viral transcription, possibly in conjunction with viral-encoded transcription factors.

The OtV-1 genome also encodes a RNase III enzyme (OtV1_127) and a RNase H (OtV1_170) that are presumably involved in processing viral mRNAs and/or tRNAs. The OtV-1 genome encodes a SW1/SNF family helicase (OtV1_131), which has been implicated in chromatin remodelling (Kim and Clark, 2002).

5.3.11. Sugar manipulation enzymes

The OtV-1 genome encodes several proteins with high identities to enzymes involved in either manipulating sugars, synthesizing polysaccharides or transferring sugars to proteins. Following the advent of '-omic' approaches, i.e. genomics, transcriptomics and proteomics as important biological research areas, the field of 'glycomics' is now emerging. The 'glycome' is defined as all the sugars a biological entity makes, including glycans fixed on proteins, lipids or DNA (Hirabayashi, 2004). Viruses can modify the glycome by two distinct mechanisms. Some affect expression of host glycosyltransferases and others express their own glycosyltransferases.

Glycosyltransferases form a complex group of enzymes involved in the biosynthesis of disaccharides, oligosaccharides and polysaccharides that are involved in the

posttranslational modification of proteins (N- and O- glycosylation), and the synthesis of lipopolysaccharides included in high molecular-weight cross-linked periplasmic or capsular material (Lovering *et al.,* 2007). Glycosyltransferase encoding genes have been reported in bacteriophages, baculoviruses, poxviruses, herpesviruses and phycodnaviruses (Markine-Goriaynoff *et al.,* 2004). In the phycodnaviruses, only two virus genera have been identified as encoding their own glycosyltransferases (Van Etten 2003). The chlorovirus ATCV-1, encodes six glycosyltransferases (Fitzgerald *et al.,* 2007b), whilst PBCV-1 encodes five putative glycosyltransferases (Zhang *et al.,* 2007). Here, we identified at least 6 putative glycosyltransferase genes in the OtV-1 genome, four from family 1, and one each from family 2 and 25 (Tables 5.5 and 5.6). Typically, viral proteins are glycosylated by host encoded glycosyltransferases located in the endoplasmic reticulum and Golgi Body (Knipe, 1996) but the PBCV-1 major capsid protein Vp54 is glycosylated by the virus itself (Wang *et al.,* 1993). No transmembrane domains or ER/Golgi body signal peptides were detected in the OtV-1 glycosyltransferases, suggesting these enzymes function in the cytoplasm. It is possible that the glycosyltransferases encoded by OtV-1 play a similar role to those in PBCV-1. Viral encoded glycosyltransferases may be crucial in the glycosylation of the OtV-1 capsid structure, particularly as there are eight putative major capsid genes in the OtV-1 genome. Further characterisation work is required to confirm this hypothesis.

OtV-1 also encodes a GDP-D-mannose dehydratase (GMD) (OtV1_012), an enzyme involved in the first step of biosynthetic pathways that lead to the formation of several deoxyhexoses e.g. GDP-L-fucose and GDP-D-rhamnose, in both prokaryotes and eukaryotes. GMD sequences are well conserved with three conserved amino acid residues Ser/Thr, Tyr and Lys involved in catalysis, and a glycine-rich motif at the N-

terminus involved in cofactor binding. GMD is widely represented across all taxa and has recently been identified in three chloroviruses (Tonetti *et al.,* 2003; Markine-Goriaynoff *et al.,* 2004; Fitzgerald, 2007a, b), OtV5 (Derelle *et al.,* 2008) and the *Prochlorococcus* phage P-SSM2 (Sullivan *et al.*, 2005). GDP-L-fucose is a donor substrate for fucosyltransferase activity. Fucose is found in many glycoconjugates of prokaryotes and eukaryotes, where it often plays a fundamental role in cell-cell adhesion and recognition (Luther and Haltiwanger, 2009).

OtV-1 also possesses 6-phosphofructokinase (OtV1_174), a key glycolytic enzyme. This enzyme converts fructose 6-phosphate to fructose-1,6-bisphosphate with concomitant ATP hydrolysis, an irreversible step which commits the glycolytic process to completion. This is only the second virus, aside from OtV5, reported to encode this enzyme. As glycolysis is the breakdown of sugars to release energy, this important reaction may be controlled by a viral-encoded 6-phosphofructokinase during infection.

5.3.12. Protein, amino acids and lipid synthesis, modification and degradation

The inability of viruses to perform protein synthesis independently of the host is one of the features that distinguish these biological entities from cellular or 'living' organisms. However, similar to OtV5, OtV-1 contains a gene (OtV1_036) encoding acetolactate synthase (ALS) a thiamine diphosphate-dependent enzyme, which catalyses the first step in the branched amino acid synthesis pathway i.e for synthesis of leucine, isoleucine and valine. ALS is found in plants, fungi and bacteria. Interestingly, the OtV-1 gene seems to have been acquired from a bacterial source, since all significant Blast results are to the equivalent gene in different bacterial species with an average amino acid identity of 33% and e-values in the region of $1e^{-83}$.

A putative proline dehydrogenase gene (OtV1_193) was also identified in the OtV-1 genome. This enzyme participates in arginine and proline metabolism. Interestingly, this gene may have originated from the host, since the virus gene shares 49% identity with the *O. tauri* host copy (Table 5.8). Acquiring genes encoding ALS and proline dehydrogenase presents obvious advantages for this virus, in terms of facilitating amino acid synthesis during replication. OtV-1 also encodes a putative asparagine synthase (glutamine-hydrolysing) (OtV1_028), an enzyme that plays a role in the generation of asparagine from aspartate. This enzyme is also found in the dsDNA viruses OtV5 (Derelle *et al.,* 2008) and APMV Mimivirus (Raoult et al 2004). 3-dehydroquinate synthase catalyses the formation of dehydroquinate (DHQ) and orthophosphate from 3-deoxy-D-arabino heptulosonic 7 phosphate (Bender *et al.,* 1989). This reaction is part of the shikimate pathway, which is involved in the biosynthesis of aromatic amino acids. CDS (OtV1_038) encodes a putative 3-dehydroquinate synthase. The 3-dehydroquinate synthase domain is present in isolation in various bacterial 3-dehydroquinate synthases and also present as a domain in the pentafunctional AROM polypeptide (Hawkins and Smith, 1991). The CDS in OtV-1 shares 36% identity with an equivalent gene in the bacterium *Magnetospirillum gryphiswaldense* (Table 5.5). As this enzyme is found widely in bacterial species, it would be unsurprising if this were the source of the gene found in the OtV-1 genome.

OtV-1 encodes a putative FtsH metalloendopeptidase (OtV1_014) that has not been reported in other characterised eukaryotic algal virus genomes, with the exception of OtV5 (Derelle *et al.,* 2008). FtsH is a membrane-bound ATP-dependent Zn^{2+} protease (Karata *et al.,* 1999). This cell division protein is present in bacteria and plants and the CDS in the OtV-1 genome bears close sequence homology to marine bacterial genes,

notably in *Prochlorococcus* (data not shown). Interestingly, this enzyme is involved in the early stages of photosystem II repair in the cyanobacterium *Synechocystis* sp. PCC6803 (Silva *et al.,* 2003). It is postulated here that the OtV-1 virus utilises this enzyme to perform a similar function to its cyanobacterial homologue, ultimately ensuring that the chloroplast functions normally during infection.

A putative N-myristoyltransferase (OtV1_088) is also present in the OtV-1 genome. Covalent modification with fatty acids is now a well established feature of several cellular and viral polypeptides (Cross, 1987). Two common fatty acid modifications are palmitylation and myristylation. Myristylation occurs pre-dominantly co-translationally and involves addition of the 14-carbon saturated fatty acid, myristic acid, usually via an amide bond to an amino-acid terminal glycine residue (Wilcox *et al.,* 1987). Although myristylated proteins are known to be widespread both in eukaryotic cell membranes and in viruses (Grand, 1989), this is the first report of a putative N-myristoyltransferase in a member of the *Phycodnaviridae.* However, there is a similar encoded protein reported for the NCLDV member, APMV Mimivirus (Raoult *et al.,* 2004). This enzyme has a specific requirement for myristoyl-CoA (Towler, 1986), produced from myristate by acyl-CoA synthetase in the presence of ATP. Myristolytransferase activity is associated with cytoplasmic and membrane fractions in both yeast and mammalian cells (Towler, 1986). Myristoylation is a process whereby essentially cytoplasmic proteins or enzymes can become membrane-bound, thus locating them at their site of action. It is therefore possible that myristate can play a role in mediating protein-protein interactions within the virus, such as in virus capsid assembly. The putative gene in OtV-1 shares approximately 36% identity with several eukaryotic organisms.

A putative prolyl-4-hydroxylase enzyme (OtV1_113) was also found in OtV-1. This is a procollagen-modifying enzyme and a key enzyme in the biosynthesis of collagens, a family of extracellular matrix proteins in higher organisms (Eriksson *et al.,* 1999). This enzyme has been reported in a small number of viruses, namely PBCV-1 (Eriksson *et al.,* 1999), APMV Mimivirus (Raoult *et al.,* 2004) and OtV5 (Derelle *et al.,* 2008). It had previously been thought that 4-hydroxyproline was restricted to certain plant and animal proteins only. The function of 4-hydroxyproline residues in all collagens and collagen-like proteins in animals is to stabilize their triple helical structures (Prockop, 1995). The functions of these residues in plant proteins are less well characterized but are also likely to involve stabilization of structures. The role of 4-hydroxyproline residues in viral proteins is likely to be similar to those in animal and plant proteins, but work is needed to elucidate these functions. Presumably, the OtV-1 virus utilises this enzyme in the production of structural bodies during its propagation.

The presence of a putative 2OG-Fe(II) oxygenase (OtV1_172) indicates the virus may encode additional enzymes involved in the modification of collagen-like proteins. The enzyme 2OG-Fe(II) oxygenase belongs to a class of enzymes that are widespread in eukaryotes and bacteria and catalyze a variety of reactions typically involving the oxidation of an organic substrate using a dioxygen molecule (Prescott, 1993). An extensively characterised reaction involving this enzyme is the hydroxylation of proline and lysine side chains in collagen (Aravind and Koonin, 2001). The presence of both a putative prolyl-4-hydroxylase and a putative 2OG-Fe(II) oxygenase in the OtV-1 genome indicates this virus may encode the stabilisation of complex structures, such as carbohydrate units, during its replication.

5.3.13. Restriction modification enzymes

Although the level of methylation of the OtV-1 genome is unknown, OtV-1 encodes four methyltransferases (Table 5.6). Methyltransferases have already been reported in the genomes of *Chlorella* viruses (Zhang, 1998; Fitzgerald, 2007a, b). OtV-1 encodes one DNA methyltransferase that is predicted to methylate cytosine and two DNA methyltransferases which methylate adenine, whilst a putative FkbM methyltransferase is also present. Although the exact biological function of the virus-encoded methyltransferases is unknown, methylation of the virus genome most likely protects viral DNA from endonucleolytic attack (Essani, 1987). In bacteria, restriction modification systems confer resistance to foreign DNA and viral DNA. It has been observed that the chlorovirus PBCV-1 encodes three cytosine and two adenine methyltransferases (Nelson *et al.,* 1998). A direct correlation has been reported between increasing 5-methylcytosine (m5C) concentrations in virus genomes and the sensitivity of virus replication to the cytidine methylation inhibitor, 5-azacytidine (Burbank *et al.,* 1990). OtV-1 encodes a FkbM methyltransferase (OtV1_015), with some similarity to an enzyme from several bacterial species. In addition to the presence of a putative FkbM methyltransferase in OtV-1, this enzyme is encoded by just 3 chlorovirus types (Fitzgerald *et al.,* 2007 a, b), OtV5 (Derelle *et al.,* 2008) and the *Prochlorococcus* phage P-SSM2 (Sullivan *et al.,* 2005).

5.3.14. Miscellaneous

A gene (OtV1_030) encoding a PhoH family protein is also present in OtV-1. PhoH is a cytoplasmic protein and predicted ATPase that is induced by phosphate starvation in various bacteria (Wanner, 1993; 1996). Phosphorus (P) is an essential macronutrient for the growth and development of living organisms. It is a constituent of key molecules such as ATP, nucleic acids, or phospholipids, and in forms such as

phosphate, pyrophosphate or ATP, plays a crucial role in energy transfer, metabolic regulation and protein activation. P is one of the key macronutrients potentially limiting phytoplankton growth and so it is not surprising that they have evolved various adaptive responses to cope with growth under conditions of limited P availability (Cembella, 1984). Biochemical and metabolic adaptations involve changes that increase the availability of endogenous and exogenous inorganic phosphate. The presence of a phosphate starvation inducible protein in the OtV-1 viral genome is of interest since the virus is likely to employ this protein to increase the phosphate concentration in the host cell to facilitate viral replication. The OtV-1 gene may have originated from a bacterial source, as it shares 41% identity with a similar protein encoded by the bacterium *Thermoanaerobacter pseudethanolicus* (Table 5.5), and shares similar homology to several other bacterial-encoded PhoH family proteins. The fact that a putative phosphate repressible phosphate permease (PPRPP) gene was previously identified in the coccolithovirus, EhV-86 (Wilson *et al.,* 2005b) reiterates the idea that the availability of P is critical for viral replication.

Echinonectin is a dimeric galactosyl-binding lectin believed to play a role in embryonic cell-extracellular matrix adhesion in sea urchin eggs and embryos (Alliegro and Alliegro, 2007). This protein has a marked specificity for galactose. The presence of a putative gene (OtV1_115) encoding this protein suggests it may act as an adhesion protein, as it has this function in cells of developing echinoderm embryos, where it is believed to play a role in cell anchoring. This enzyme may also be related in function to the sugar manipulation enzymes encoded by the OtV-1 genome.

OtV-1 has several genes encoding other putative proteins, including a rhodanese domain-containing protein (OtV1_209) and an ABC domain protein (OtV1_060). A large putative CDS (OtV1_117) has close homology to a virus-like inclusion body encoded by the protist *Trichomonas vaginalis* (Table 5.5). The CDS OtV1_117 is 3398 amino acids in length with a G+C content of 43.76%. This gene in OtV-1 aligns with three consecutive CDSs in the reference genome, OtV5. To ensure the single continuous ORF in OtV-1 was a true CDS, primers were designed for regions within this ORF to test for possible sequencing errors arising from the 454 process. The sequence obtained from these PCR checks confirmed the original sequencing was correct and the single large CDS was a complete and continuous open reading frame (results not shown). Virus inclusion bodies may be aggregations of mature virus particles or large centres for viral synthesis. For example, previous studies on the cauliflower mosaic virus have demonstrated inclusion bodies are the sites of accumulation of viral gene products and viral DNA synthesis within host cells (Pfeiffer and Hohn, 1983). Functional analysis of the putative gene in OtV-1 is required for further elucidation of its role in the virus lifestyle.

5.4. Conclusions

The 191, 761 bp OtV-1 genome is predicted to encode 232 proteins and 4 tRNAs. The putative protein-encoding genes are relatively evenly distributed on both strands and intergenic space is minimal. Aside from the core genes previously reported in members of the NCLDV group of viruses, there are several putative genes present in the OtV-1 genome that have either previously never been described in a virus or have only been reported in a few viruses at most. The data presented here has been a preliminary study and research into the viruses that infect *Ostreococcus* species is still very much in its infancy. The sequencing of the OtV-1 genome has revealed some surprising and

exciting features. Transcription profile analysis of the genes identified in the OtV-1 genome is required for further insights into their functions and their roles in the virus infection cycle. Functional characterisation of the genes reported in this study should help to gain a clearer insight into how these viruses operate during their replication cycle. Now that three OtV genomes have been recently sequenced, the information gleaned should afford a great opportunity for further insights into virus-host dynamics and the potential for exploitation of viral-encoded proteins.

Chapter 6

The genome sequence and annotation of

Ostreococcus tauri Virus-2, OtV-2

6.1. Introduction

Recently, the genomes of two viruses, OtV5 (Derelle *et al.,* 2008) and OtV-1 (Chapter 5; Weynberg *et al.,* 2009, in press), which infect the unicellular picoeukaryotic phytoplankton *Ostreococcus tauri,* were sequenced. OtV-1 and OtV5 are icosahedral, between 100 to 120 nm in diameter and were originally isolated by plaque assay from seawater samples collected from the English Channel (OtV-1) (Weynberg *et al.,* 2009, in press) and Mediterranean coastal waters (OtV5) (Derelle *et al.,* 2008). Phylogenetic analysis of the DNA polymerase genes of OtV-1 and OtV5 places these two viruses in the genus *Prasinovirus* within the family *Phycodnaviridae.* Their host, *Ostreococcus tauri,* is the smallest free-living eukaryote described to date, with a size of less than 1 µm . The cellular organisation of *O. tauri* is very simple with only a single chloroplast, a single mitochondrion, a single Golgi body and a very reduced cytoplasmic compartment (Henderson *et al.,* 2007). *O. tauri* also lacks flagella and there is no cell wall surrounding the cell membrane. The sequencing of the complete genomes of *O. lucimarinus* (Palenik *et al.,* 2007) (Chretiennotdinet *et al.,* 1995) and *O. tauri* (Derelle *et al.,* 2006) has revealed a size of 13.2 Mbp and 12.6 Mbp respectively, the latter being the smallest genome size of all known eukaryotes.

The *Ostreococcus* genus includes distinct genotypes physiologically adapted to high-or low-light environments, providing evidence of niche adaptation in eukaryotic picophytoplankton (Rodriguez *et al.,* 2005). Such adaptation has been well-characterised in recent studies on the diversity and ecophysiology of the cyanobacterium, *Prochlorococcus* (Moore *et al.,* 1998; Moore and Chisholm, 1999). The global success of this abundant primary producer has been in part attributed to distinct low- and high-light adapted ecotypes existing in different niches and utilising different resources (Rocap *et al.,* 2003). Strains of *O. tauri* were isolated in

geographically different locations and depths and have been shown to be genetically (based on rRNA 18S and internal transcribed spacer (ITS) sequencing) and physiologically (light-limited growth rates) different from one another (Rodriguez *et al.,* 2005). Growth rates of deep-sea strains were reported to display severe photoinhibition at high light intensities (and are thus commonly referred to as low-light-adapted strains), while strains isolated from surface waters have very slow growth rates at the lowest light intensities (and are thus commonly referred to as high-light-adapted strains). Genetic distances between isolates do not seem to reflect geographical trends but rather are due to contrasts in light and nutrient conditions experienced between surface and deep isolates, which drive their genetic divergence (Rodriguez *et al.,* 2005; Cardol *et al.,* 2008; Six *et al.,* 2008). Therefore, investigations of the viruses that infect different strains of the same species, which display distinct adaptations to environmental conditions, should provide insights into host-virus relationships within differing ecotypes.

This chapter describes the genome sequence of a third *Ostreococcus*-specific virus, OtV-2, that infects a single low-light-adapted strain of *O. tauri*, RCC393. The two sequenced viruses, OtV-1 (Weynberg *et al.,* 2009) and OtV5 (Derelle *et al.,* 2008) infect a high-light adapted strain, *O. tauri* OTH 95. In addition, a comparative analysis of all three OtV genomes is conducted. Recent work has shown that small-subunit rRNA sequences of *Ostreococcus* from cultures and environmental samples cluster into four different clades that are likely distinct enough to represent different species (Guillou *et al.,* 2004; Rodriguez *et al.,* 2005). The host strain, RCC393, belongs to clade B, based on analysis of the ITS region of the SSU ribosomal RNA operon (Guillou *et al.,* 2004; Rodriguez *et al.,* 2005). RCC393 is currently one of four members of Clade B strains of

Ostreococcus and was sampled at a depth of 90 m from the Mediterranean Sea (Rodriguez *et al.,* 2005). *O. tauri* strain RCC393 displays a different pigment composition than the high-light strains with a higher chlorophyll *b*: chlorophyll *a* ratio and the additional presence of the chlorophyll pigment chl c_{cs170}, which is absent in high-light adapted strains of the species (Rodriguez *et al.,* 2005). It is likely this particular chlorophyll pigment performs a light-harvesting function in the blue-green region of the spectrum, which is seen to dominate at low ambient light levels.

6.2. Materials and methods

6.2.1. Maintenance of algal host strain

The OtV-2 host, *Ostreococcus tauri* strain RCC393 was grown in Keller (K) medium (Keller *et al.,* 1987) under the conditions described in Section 2.2.1.

6.2.2. Virus isolation

The virus OtV-2 was isolated from surface seawater collected on 07/05/2007 at the L4 sampling station in the Western English Channel (50º15'N, 04º13'W) following the method described in Section 2.2.3.

6.2.3. Clonal virus isolation of OtV-2

In order to obtain clonal virus stocks, OtV-2 was purified by serial dilution to extinction as described in Section 2.2.4.2.2, as the host strain failed to grow successfully on agarose solid bottom plates preventing the use of plaque purification techniques.

6.2.4. DNA preparation and sequencing

For preparation of large quantities of viruses for genome sequencing, ten litre volumes of exponentially growing *O. tauri* strain RCC393 culture were inoculated with 10 ml of OtV-2 at a multiplicity of infection of one. Lysed cultures were passed sequentially through 0.8 µm and 0.2 µm filters to remove large cellular debris. Virus filtrates were concentrated by ultrafiltration to ~ 50 ml using a Quixstand benchtop system and hollow fibre cartridges with a 30, 000 pore size (MWC) (GE Healthcare Amersham Biosciences). The 50 ml concentrate was then further concentrated to ~10 ml using the Mid-Gee benchtop system and hollow fibre cartridge with a 30, 000 pore size (MWC) (GE Healthcare Amersham Biosciences). Aliquots (3 ml) of the concentrated OtV-2 lysate were purified by caesium chloride gradient purification as described in Section

2.2.5.2. Complete sequencing of the OtV-2 genome was performed at the Advanced Genomics Facility based at the University of Liverpool, UK.

6.2.5. Sequence assembly and finishing

The OtV-2 sequence was assembled and finished as described in Section 2.2.9.7.4.

6.2.6. Genome annotation

The OtV-2 genome was annotated as described in Section 2.2.9.7.5.

6.2.7. Analysis of repeat regions

To identify repetitive sequences within the OtV-2 genome, a dot-plot analysis was performed using the LBDotView Version 1.0 (http://www.lynnon.com/dotplot.html) (Huang and Zhang, 2004). This analytical tool can compare one genome on the x-axis with a second genome on the y-axis. Here, the OtV-2 genome is compared against itself and the OtV-1 and OtV5 genomes, enabling the precise location and orientation of homologous sequences to be located within the plot.

6.2.8. Phylogenetic analyses

Phylogenetic analyses were performed with the amino acid sequences of the DNA polymerase gene, DNA topoisomerase II gene, high affinity phosphate transporter gene and the major capsid genes as described in Section 2.2.9.7.3. The designations and GenBank accession numbers of the DNA polymerase and DNA topoisomerase II sequences used for phylogenetic analysis are described in Section 5.2.8. The designations and GenBank accession numbers of the high affinity phosphate transporter gene sequences are shown in Table 6.1.

Table 6.1. Details of high affinity phosphate transporter sequences used for phylogenetic analysis

Scientific name	Accession number
Alteromonas macleodii	NC_011138.1
Chlamydomonas reinhardtii	XP_001695776.1
Emiliania huxleyi	AF334403.1
Emiliania huxleyi virus-86 (EhV-86)	YP_293871.1
Ostreococcus lucimarinus	XP_001416412.1
Ostreococcus tauri	CAL52329
Tetraselmis chui	AAO47330.1
Thalassiosira pseudonana	XP_002281698.1
OtV-2	This study

The designations and GenBank accession numbers of the major capsid protein gene sequences used for phylogenetic analysis are shown in Table 6.2.

Table 6.2. Details of major capsid sequences used for phylogenetic analysis

Scientific name	Accession number
African Swine Fever Virus (ASFV)	ABL67020.1.;
Chlorella virus CVK2	BAA76600.1;
Chrysochromulina ericina virus (CeV01),	ABU23712
Emiliania huxleyi virus-86 (EhV-86)	CAI65508.2;
Heterosigma akashiwo virus (HaV)	BAE06835.1
Paramecium bursaria Chlorella virus NY-2A (PBCV NY-2A)	YP_001497781.1
Paramecium bursaria Chlorella virus (PBCV-1)	AAA88828.1
Pyramimonas orientalis virus (PoV01)	ABU23714.1
OtV-2	This study

6.3. Results and discussion

6.3.1. General description of the OtV-2 genome

Sequence analysis of the OtV-2 genome revealed a linear genome of 184,409 bp. This is approximately 6 kb shorter than the 190 kb predicted by pulsed field gel electrophoresis analysis of the OtV-2 genome (Chapter 4). General features of the OtV-

2 genome sequence include i) a nucleotide composition of 42.15% G+C ii) a total of 237 predicted coding sequences (CDSs) that were defined as having a start codon followed by at least 65 additional codons prior to a stop codon iii) CDSs equally distributed on both strands (54% on the positive strand and 46% on the negative strand) iv) an average gene length of 725 bp and v) a coding density of 1.285 genes per kbp (Table 6.3). These general features of the OtV-2 genome are compared to OtV-1 and OtV5 in Table 6.3.

Table 6.3. Comparison of the general characteristics of the genomes of OtV-2, OtV-1, OtV5, and *Ostreococus tauri* strain OTH 95.

	OtV-1	OtV-2	OtV-5	*O. tauri* OTH95
Genome size (bp)	191,761	184,409	186,234	12,560,000
% G + C	45.36	42.15	45.29	58
tRNA	4	5	5	-
Predicted CDSs (65 aa)	232	237	268	8,166
Mean gene size (bp/gene)	750	725	702	1300
Coding density (genes/kb)	1.303	1.285	1.424	-
% known genes	31	30.4	22	77
% hypothetical genes	13	18.5	20	-
% genes with no matches to genes of known function	56	51.0	57	23

Orientation of the OtV-2 genome revealed a high level of co-linearity with OtV-1 and OtV5 (Fig. 6.1). For example, 82.3% i.e. 195 of 237 CDSs in OtV-2 have orthologs in OtV5, with 42 CDSs being unique to OtV-2. A general comparison of OtV5 genome features with the OtV-2 genome also showed similar attributes, with OtV5 having a comparable size (186,234 bp) and slightly higher G+C composition (45.29%), though with a slightly smaller mean gene length (692 bp) and more predicted genes (268). The

OtV-1 genome is approximately 6 kb – 7kb larger than both genomes at 191,761 bp and has a larger mean gene size of 750 bp (Table 6.3).

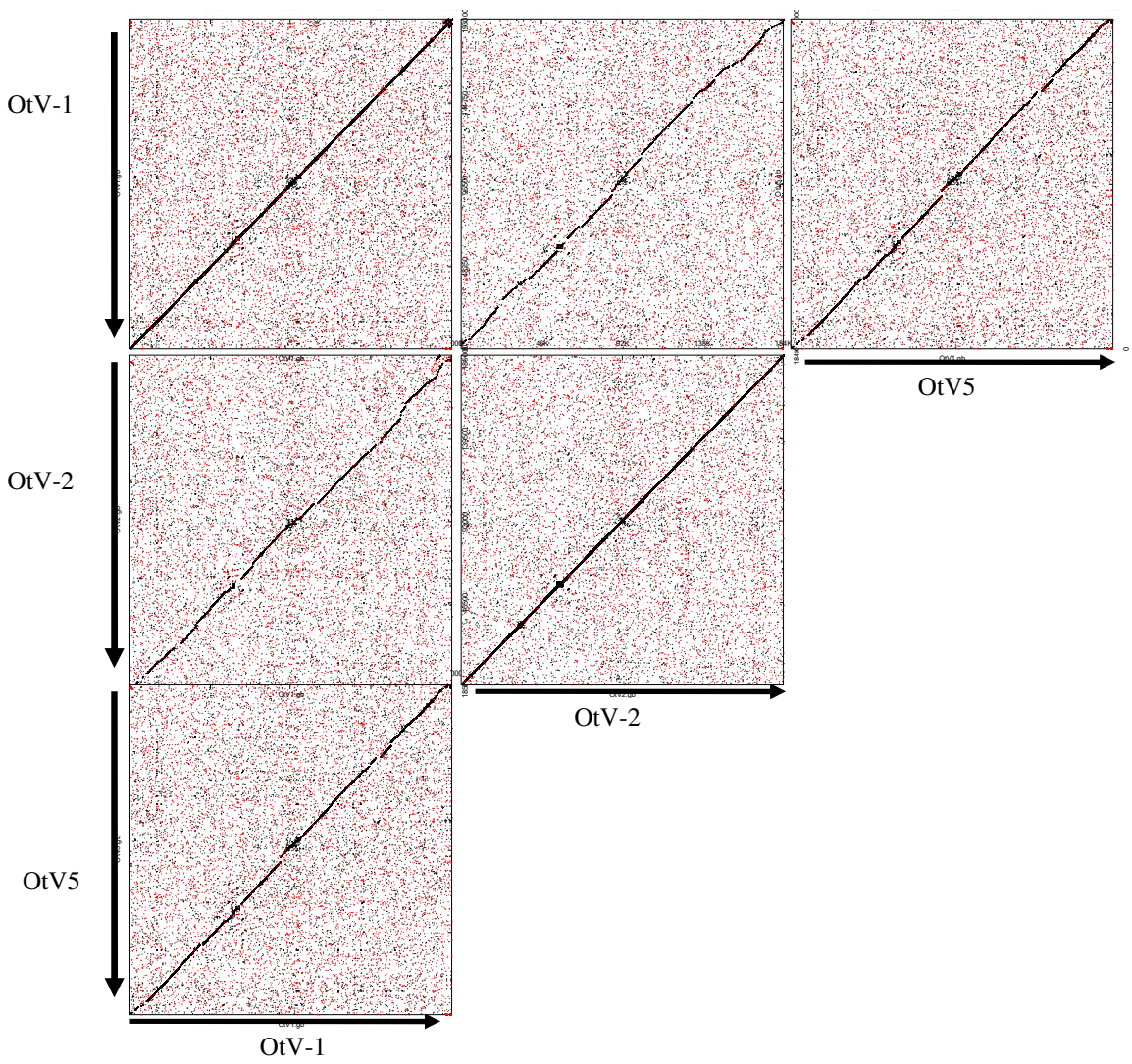


Figure 6.1. Comparison of the three sequenced *Ostreococcus* viruses, OtV-1, OtV-2 and OtV5 using dot plot analysis (LBDot). The dots represent ORF homology between two viruses with an e-value of less than 0.001.

6.3.2. Identity of putative CDSs

Of the 237 CDSs identified in the OtV-2 genome, 165 (69.6%) encode putative proteins where no function can be been attributed (Table 6.3). Of these unknown CDSs, 44 share close similarity with other hypothetical proteins in databases. Therefore, 121 (51%) of the CDSs do not correspond to known proteins outside of the three sequenced OtV genomes, highlighting a potentially high degree of functional novelty in this virus and its family. Predicted metabolic functions of the remaining 72 CDSs (30.4%), which have homology to known proteins, were split into 9 functional groups (Table 6.4). Ten CDSs could be assigned putative functions involved in DNA replication, recombination and repair; 7 in nucleotide metabolism and transport; 8 in transcription; 15 in protein and lipid synthesis, modification and degradation; one in signalling; six in DNA methylation; 9 in sugar metabolism; 8 encoding capsid proteins; and a further 8 with other miscellaneous functions.

Table 6.4. Putative proteins encoded on the OtV-2 genome grouped by function.

DNA, RNA, replication, recombination and repair	CDS#
DNA polymerase	OtV2_207
ATP dependent DNA ligase	OtV2_169
DNA topoisomerase I	OtV2_161
DNA topoisomerase II	OtV2_211
PCNA	OtV2_100
Exonuclease	OtV2_002
Exonuclease	OtV2_118
RNase H	OtV2_156
VV A18 Helicase	OtV2_056
NTPase/helicase	OtV2_107
Nucleotide transport and metabolism	
Ribonucleotide reductase, large subunit	OtV2_123
Ribonucleotide reductase, small subunit	OtV2_140
dUTP pyrophosphatase	OtV2_194
ATPase (VV A32 virion packaging)	OtV2_084
Thymidine kinase	OtV2_188
CMP/dCMP deaminase, Zinc-binding	OtV2_190
FAD dependent thymidylate synthase	OtV2_051
Transcription	
RNA transcription factor TFIIB	OtV2_142
RNA transcription factor TFIIS	OtV2_019
RNase III	OtV2_117
RNA polymerase sigma factor	OtV2_202
SW1/SNF helicase domain protein	OtV2_119
TATA-box binding protein	OtV2_130
mRNA capping enzyme	OtV2_079
mRNA capping enzyme	OtV2_145
Sugar manipulation enzymes	
GDP-D-mannose dehydratase	OtV2_007
Glycosyltransferase	OtV2_008
Glycosyltransferase	OtV2_028
Glycosyltransferase	OtV2_163
Glycosyltransferase	OtV2_216
6-phosphofructokinase	OtV2_159
NAD-dependent epimerase/dehydratase	OtV2_031
dTDP-D-glucose 4,6-dehydratase	OtV2_033
6-phosphogluconate dehydrogenase	OtV2_167
Signalling	
Serine/Threonine protein kinase	OtV2_135

DNA restriction/methylation	CDS#
Adenine-specific methyltransferase	OtV2_083
Adenine-specific methyltransferase	OtV2_116
DNA-cytosine methyltransferase	OtV2_050
DNA-cytosine methyltransferase	OtV2_003
Methyltransferase FkbM	OtV2_010
DNA methylase	OtV2_078
Protein and lipid synthesis/modification	
Prolyl 4-hydroxylase alpha-subunit	OtV2_106
2OG-Fe(II) oxygenase	OtV2_165
2OG-Fe(II) oxygenase	OtV2_198
2OG-Fe(II) oxygenase	OtV2_200
Procollagen-lysine,2-oxoglutarate 5-dioxygenase	OtV2_158
Ubiquitin C-terminal hydrolase	OtV2_146
ATP-dependent protease proteolytic subunit	OtV2_171
FtsH2 metalloprotease	OtV2_009
Aminotransferase family protein	OtV2_032
Acetolactate synthase	OtV2_030
N-myristoyltransferase	OtV2_076
TPR domain-containing protein	OtV2_027
Asparagine synthetase	OtV2_022
Aspartyl/Asparaginyl beta-hydroxylase	OtV2_166
3-methyl-2-oxobutanoate hydroxymethyltransferase	OtV2_029
Capsid protein	
Capsid protein 1	OtV2_057
Capsid protein 2	OtV2_062
Capsid protein 3	OtV2_066
Capsid protein 4	OtV2_085
Capsid protein 5	OtV2_086
Capsid protein 6	OtV2_154
Capsid protein 7	OtV2_206
Capsid protein 8	OtV2_210
Miscellaneous	
Virus inclusion body	OtV2_109
Rhodanese domain-containing protein	OtV2_203
ABC1 domain protein	OtV2_052
Fibronectin-binding protein	OtV2_199
Tail fibre assembly protein	OtV2_065
Cytochrome b_5	OtV2_201
Phosphate starvation inducible protein	OtV2_024
High affinity phosphate transporter	OtV2_222

6.3.3. General overview of conservation between Ostreococcus tauri viruses

6.3.3.1. Comparison with other NCLDVs

OtV-2 is smaller than most other phycodnaviruses and sequenced NCLDVs, including OtV-1 and OtV5 (Table 6.3). A total of 15 Group I, II and III core genes were identified in the OtV-2 genome (Table 6.5). However, fewer CDSs, including the core set of conserved virus genes for NCLDVs, were found in the OtV-2 genome when compared to the majority of NCLDVs sequenced to date, with the exception of the recently sequenced phaeovirus, FsV-158 (Schroeder *et al.,* 2009) and the two OtV viruses (Table 6.5). Hence, homologues were identified for only 5 of the 9 Group I genes, 6 of the 8 Group II genes and 4 of the 14 Group III genes in the OtV-2 genome (Table 6.5). This is also the case for the OtV-1 and OtV5 genomes. The genes for Vaccinia Virus (VV) D5 type ATPase, thiol oxidoreductase, VV D6R-type helicase and VLFT2-like transcription factor are absent from all three sequenced OtV genomes. In addition, the genome of FsV-158 does not have a VV D6R-type helicase. These findings have implications for the present classification system of core genes. The Group I core genes absent from all three OtV genomes could possibly now be regarded as Group II and not Group I core NCLDV genes. Furthermore, the presence of genes for a Proliferating Cell Nuclear Antigen (PCNA) and ribonucleotide reductase large and small subunits, in all sequenced genomes of members of the *Phycodnaviridae* could reclassify these genes as Group I core genes instead of Group II core genes. A redefinition of Group I core genes would then see a pattern emerging of predominantly genes with DNA/RNA replication, recombination and repair functions, nucleotide transport and metabolism functions, as well as the signalling function of serine/threonine protein kinase and the structural functions of the capsid proteins.

Table 6.5.
Presence of NCLDV Core Genes (Groups 1, 2 and 3) in Various NCLDV Genomes

	1										2									3												
Family Species	Vaccinia virus D5-Type ATPase	DNA polymerase (*otv2_207f*)[a]	Vaccinia virus A32-Type ATPase (*otv2_084r*)	Vaccinia Virus A18-Type Helicase (*otv2_056f*)	Capsid protein (*otv57f/62f/66f/85f/86f/154f/206r/210r*)	Thiol-oxidoreductase	Vaccinia Virus D6R-Type Helicase	Serine/Threonine Protein Kinase (*otv2_135r*)	VLTF2-Like Transcription Factor	TFII-Like Transcription Factor (*otv2_019f*)	MuT-Like NTP Pyrophosphohydrolase	Myristyolated Virion Protein A	Proliferating Cell Nuclear Antigen (*otv2_100f*)	Ribonucleotide Reductase, Large Subunit (*otv2_123r*)	Ribonucleotide Reductase, Small Subunit (*otv2_140r*)	Thymidylate Kinase	dUTPase (*otv2_194f*)	A494R-Like Uncharacterised Protein	RuvC-Like Holliday Junction Resolvase	BroA-Like	Capping Enzyme (*otv2_079f/otv2_145*)	ATP-dependent Ligase (*otv2_169r*)	RNA Polymerase, Subunit 1	RNA Polymerase, Subunit 2	Thioredoxin/Glutaredoxin	Serine/Threonine Phosphatase	BIR Domain	Virion-Associated Membrane Protein	Topoisomerase II (*otv2_211r*)	SW1/SNF2 Family Helicase (*otv2_119f*)	RNA Polymerase, Subunit 10
Phycodnaviridae																																
OtV-2		✓	✓	✓		✓			✓	✓			✓	✓		✓	✓	✓			✓	✓						✓	✓		
EhV-86	✓	✓	✓	✓		✓		✓	✓	✓	✓		✓	✓		✓	✓	✓			✓	✓	✓	✓	✓		✓	✓			✓
PBCV-1	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓			✓	✓			✓	✓		✓	✓		
EsV-1	✓	✓	✓	✓		✓		✓	✓				✓	✓		✓		✓	✓	✓					✓						
FsV-158	✓	✓	✓			✓		✓	✓				✓	✓		✓															
Mimiviridae																																
Mimivirus	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		✓		✓		✓	✓	✓	✓	✓	✓	✓	✓	✓				
Iridoviridae																																
LCDV	✓	✓	✓			✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓					✓	✓				✓		✓		
IIV-6	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Asfarviridae																																
ASFV	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓				✓	✓	✓	✓			✓	✓			✓
Poxviridae																																
VACV	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓				✓
MOCV	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓				✓	✓		✓		✓	✓	✓	✓						
AMEV	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓					✓	✓	✓	✓		✓	✓	✓	✓	✓	✓				
MSEV	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓					✓	✓	✓	✓		✓	✓	✓	✓		✓	✓		✓		

NOTE – Based on data from this study and Bawden et al (Bawden *et al.,* 2000), Iyer *et al.,* (Iyer *et al.,* 2001), Raoult et al (Raoult *et al.,* 2004), Allen et al (2006b), Schroeder et al (2009). ASFV, African swine fever virus; NTP, nucleoside triphosphate; dUTPase, deoxyuridine triphosphatase; ATP, adenosine triphosphate. [a] Gene ID for OtV-2.

6.3.3.2. tRNAs

The OtV-2 genome was analysed for tRNAs using the tRNAscan-SE program (Lowe, 1997) (http://lowelab.ucsc.edu/tRNAscan-SE/) and is predicted to encode five tRNAs: encoding amino acids Ile, Gln, Tyr, Asn, and Thr (Table 6.6). Two of the tRNAs (Ile and Tyr) contain introns and are clustered in a region of the OtV-2 genome between 133,756 and 136,498 bp (Table 6.6). The tRNAs present in the OtV-2 genome are the same as those identified in the OtV5 genome and occur in the same order (Table 6.6). Both OtV-2 and OtV5 have introns in Ile and Tyr. By contrast, the OtV-1 genome encodes just four of the five tRNAs (Tyr is missing) and all contain introns (Table 6.6). Considering there is a clear difference between the two host strains these viruses infect and the environments they occupy, the differences and similarities of their tRNAs provokes certain questions. A host range analysis was conducted (Chapter 4) that revealed a high strain specificity for the OtV-1 and OtV-2 viruses, as they were seen to infect their respective *O. tauri* host strains only. It is possible the OtV-1 tRNAs show marked differences in comparison to those in OtV-2 and OtV5 due to a larger dependency on the host tRNAs during infection. Horizontal gene transfer may also be an explanation for the loss of a tRNA Tyr in the OtV-1 genome and the presence of a pseudo tRNA. The viral tRNAs are possibly transcribed as a large precursor RNA and processed via intermediates to mature RNAs, as previously reported in the chlorovirus CVK2 (Nishida *et al.,* 1999).

Table 6.6. Comparison of the OtV-1, OtV-2 and OtV5 tRNA genes.

tRNA	Anticodon	OtV-1			OtV-2			OtV5		
		tRNA #	Start	End	tRNA #	Start	End	tRNA #	Start	End
Ile	TAT	1	144,113	144,209	1	133,756	133,852	1	141,936	142,033
Gln	TTG	2	146,519	146,605	2	136,164	136,235	2	144,339	144,409
Tyr	GTA	-	-	-	3	136,240	136,344	3	144,413	144,509
Asn	GTT	3	146,751	146,847	4	136,349	136,421	4	144,514	144,586
Thr	AGT	Pseudo	148,646	148,728	5	136,427	136,498	5	144,591	144,662

6.3.3.3. Structural proteins

As seen in the OtV-1 and OtV5 genomes, the OtV-2 genome encodes eight putative capsid proteins (Table 6.4). Aside from the OtV viruses, five of the capsid proteins have closest similarity to *Pyramimonas orientalis* virus PoV-1, two have closest similarity to *Heterosigma akashiwo* virus HaV-1, and one has closest similarity to a capsid protein encoded by PBCV NY-2A. Phylogenetic analysis of these proteins with other putative phycodnavirus major capsid protein (MCP) genes showed capsid protein 6 of the OtV-2 genome (OtV2_154) forms a distinct cluster with the phycodnavirus MCP gene sequences (Fig. 6.2), indicating capsid protein 6 is likely to be the 'major' MCP in OtV-2. This will require further experimental analysis to confirm.

The presence of multiple genes encoding capsid proteins in all three OtV genomes is intriguing. The herpes simplex virus type 1 has seven genes which encode seven capsid proteins (Davison *et al.,* 1992). These proteins have various roles at different stages of assembly and help build the complex capsid structure composed of hexamer and pentamer components. One of these proteins encodes a protein which forms a single portal in one vertex of the capsid structure in the herpesvirus particle (Cardone *et al.,* 2007). The same level of complexity in the OtV capsid structures may also warrant the inclusion of several capsid proteins encoded by their genomes. The possible presence of a portal in the capsid for delivery and encapsidation of viral DNA may also be encoded by one or more of these genes.

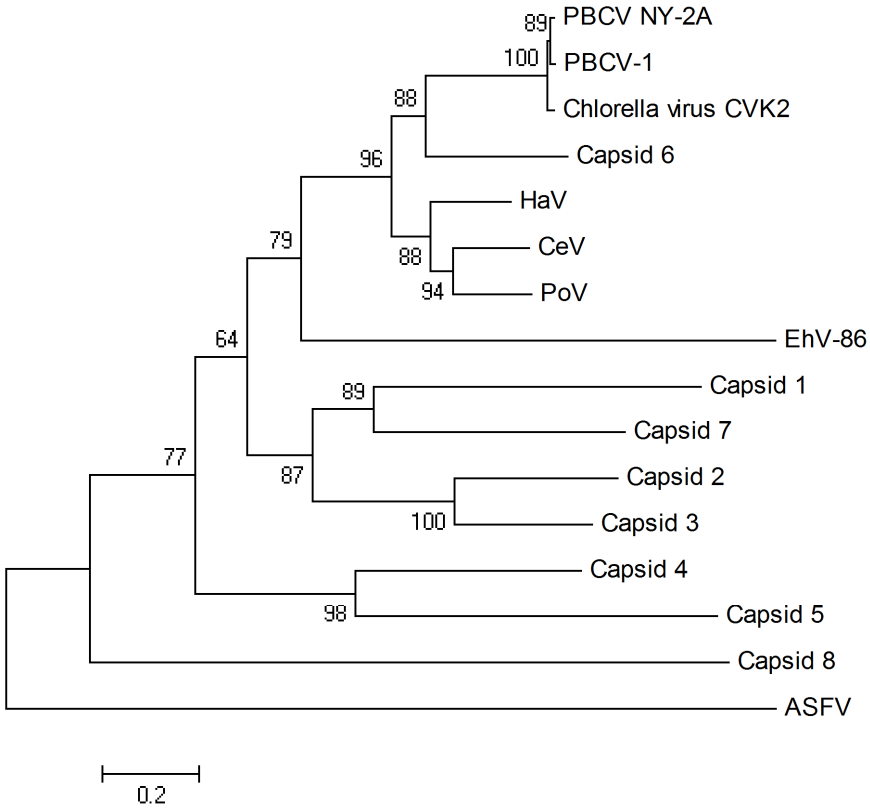


Figure 6.2. A phylogenetic inference based on a distance matrix algorithm between the 8 capsid protein gene fragments of OtV-2 and the major capsid protein genes of other large dsDNA viruses [Neighbour in PHYLIP, version 3.61 (Felsenstein, 2005)]. Abbreviations for virus names and GenBank accession numbers of DNA polymerase sequences used for phylogenetic analysis were as follows (scientific name, with abbreviation in parentheses): African Swine Fever Virus (ASFV), ABL67020.1.; *Chlorella* virus CVK2, BAA76600.1; *Chrysochromulina ericina* virus (CeV01), ABU23712; *Emiliania huxleyi* virus-86 (EhV-86), CAI65508.2; *Heterosigma akashiwo* virus (HaV), BAE06835.1; *Paramecium bursaria Chlorella virus* NY-2A (PBCV NY-2A), YP_001497781.1; *Paramecium bursaria Chlorella virus* (PBCV-1), AAA88828.1; *Pyramimonas orientalis* virus (PoV01), ABU23714.1; Capsids 1-8 of OtV-2 isolated in this study. Scale bar bar indicates a distance of 0.1 fixed mutations per amino acid position.

6.3.3.4. DNA replication, recombination and repair

OtV-2 has ten CDSs that are involved in either DNA replication, recombination or repair (Table 6.4), such as DNA polymerase (OtV-2_207), ATP-dependent DNA ligase (OtV-2_169) and a proliferating cell nuclear antigen (PCNA) (OtV-2_100). All ten CDSs are found in the OtV-1 and OtV5 genomes, with the exception of NTPase/helicase which is not found in OtV5. Unlike the DNA polymerase gene of OtV-1, no intein was identified in the OtV-2 gene. Phylogenetic analysis of the DNA polymerase gene in OtV-2 strengthens the clustering of OtV viruses, which forms a significant (100% bootstrap value for 1,000 trials) monophyly with viruses infecting *Micromonas pusilla* (Chapter 5, Fig. 5.2). Integrating these results with previous observations on the basic characteristics allow a putative assignment of the OtV viruses to the Prasinovirus family of the *Phycodnaviridae.*

6.3.3.5. Restriction modification enzymes

Although the level of methylation of the OtV-2 genome is unknown, OtV-2 does encode six methyltransferases (Table 6.4). At least four predicted methyltransferases are encoded for by OtV-1 and OtV5. OtV-2 encodes two DNA methyltransferases that methylate cytosines (OtV-2_003 and OtV-2_050) and two DNA methyltransferases which methylate adenines (OtV-2_083 and OtV-2_116). As in the other sequenced OtV genomes, a putative FkbM methyltransferase is also present (OtV-2_010), with some similarity to an enzyme from several bacterial species. This enzyme is encoded by just 3 chlorovirus types (Fitzgerald, 2007b; Fitzgerald *et al.,* 2007a) and the *Prochlorococcus* phage P-SSM2 (Sullivan *et al.,* 2005). Methylation of the virus genome possibly protects viral DNA from self-digestion or endonucleolytic attack (Essani *et al.,* 1987). In addition there is a putative DNA methylase (OtV-2_078) which is absent from the other sequenced OtV viruses.

The presence of these enzymes has been reported in other phycodnaviruses (Van Etten, 2003; Fitzgerald *et al.*, 2007a;b). Unlike some other previously characterised phycodnaviruses (Xia *et al.,* 1988; Zhang *et al.,* 1998; Delaroque *et al.,* 2001; 2003; Wilson *et al.,* 2005b; Agarkova *et al.,* 2006), there is no evidence of site-specific endonucleases in any of the sequenced OtV genomes and it is predicted the methyltransferases lack companion site-specific endonucleases. The exact biological function of the virus-encoded methyltransferases is unknown. In bacteria, restriction modification systems confer resistance to foreign DNA and viral DNA. It has been observed that if a chlorovirus, such as PBCV-1, encodes a functional cytosine methyltransferase, then the expression of this gene will be necessary for successful viral propagation. A direct correlation has been reported between increasing 5-methylcytosine (m5C) concentrations in virus genomes and the sensitivity of virus replication to the cytidine methylation inhibitor, 5-azacytidine (Burbank *et al.,* 1990).

6.3.4. How does OtV-2 differ from OtV-1 and OtV5?

6.3.4.1. Host related genes

A total of 14 CDSs in the OtV-2 genome share close similarity to genes found in the host genus, *Ostreococcus* (Table 6.7). In comparison, a total of 11 CDSs in the OtV-1 genome and six CDSs in the OtV5 genome share homology to genes found in the host genus. Five of the host-like CDSs in the OtV-2 genome are of unknown function. The remaining seven generate low e-values and high identity scores predicting CDSs that encode (among others) a putative mRNA capping enzyme (OtV-2_145) and a DNA topoisomerase II (OtV-2_211) (Table 6.7). Interestingly, the OtV-2 type II DNA topoisomerase gene has an amino acid identity of 42% to the DNA topoisomerase II genes in *Ostreococcus* host genomes (e values = 0; Table 6.7). The OtV-2 genome also encodes a ribonucleotide reductase small subunit, which shares 64% identity with

Ostreococcus lucimarinus (Table 6.7). All the OtV virus genomes sequenced to date encode a lambda-type exonuclease, also encoded by the chlorovirus PBCV-1, which is also found in the host *Ostreococcus* genomes. The presence of this gene in the host raises the possibility of gene acquisition via horizontal transfer from virus to host or host to virus, this transfer event presumably occurring long enough ago for the *Chlorella* and *Ostreococcus* lineages to diverge. All three OtV genomes encode the following host-like genes: an exonuclease; ribonucleotide reductase, small subunit; mRNA capping enzyme; fibronectin-binding protein; DNA topoisomerase II; and three unknown protein products.

Table 6.7. Summary of putative OtV-2 CDSs with database matches to host genes.

OtV-2 CDS#	Putative function	Identity[a]	E-value[b]	Present in OtV-1	Present in OtV5
OtV-2_042	dTDP-D-glucose 4,6-dehydratase	44%	$6e^{-67}$	Y	Y
OtV2_118	PBCV-1 exonuclease	41%	$9e^{-24}$	Y	Y
OtV2_140	Ribonucleoside reductase, ss	64%	$1e^{-111}$	Y	Y
OtV2_145	mRNA capping enzyme	26%	$4e^{-21}$	Y	Y
OtV2_183	Unknown	47%	$6e^{-06}$	Y	Y
OtV2_192	Unknown	34%	$9e^{-14}$	Y	Y
OtV2_195	Unknown	36%	$2e^{-05}$	Y	Y
OtV2_199	Fibronectin-binding protein	41%	$6e^{-19}$	Y	Y
OtV2_201	Cytochrome $\boldsymbol{b_5}$	60%	$5e^{-31}$	N	N
OtV2_202	RNA polymerase sigma factor	42%	$1e^{-09}$	N	N
OtV2_211	DNA topoisomerase II	45%	e= 0	Y	Y
OtV2_222	High affinity phosphate transporter	63%	e= 0	N	N
OtV2_225	Unknown	57%	e^{-160}	N	N
OtV2_229	Unknown	57%	$7e^{-09}$	N	N

[a]Amino acid identity score to host gene
[b]E-value of BLAST hit to host gene

OtV-2 is the only one of the three viruses that encodes a cytochrome b_5 (OtV-2_201), a RNA polymerase sigma factor (OtV-2_202), a high affinity phosphate transporter (OtV-2_222) and an additional two host-like genes with unknown functions (Table 6.5). A notable feature of most of the host-like genes, including those unique to the OtV-2

genome, is their position in the OtV-2 genome. They occur mostly in the second half region of the genome between 147,332 bp (start of OtV-2_183) to 180,613 bp (end of OtV-2_229). Most intriguingly is the clustering of the genes encoding a putative fibronectin-binding protein (OtV-2_199), a putative cytochrome b_5 (OtV-2_201) and a putative RNA polymerase sigma factor (OtV-2_202). Indeed the latter two genes are adjacent to one another. This clustering of host-like genes raises the possibility of a 'hotspot' region within the OtV-2 genome, with a propensity to acquire genes from the host. Previously, phycodnaviruses were reported to largely acquire genes from bacterial sources, whilst metazoan viruses, such as poxviruses, predominantly acquire genes from their hosts (Filee *et al.*, 2008). This study also reported a strong tendency for these genetic transfers to result in the positioning of laterally acquired genes at the tips of genomes. Accretion of host genes is believed to play a role in the evolution of viruses (Iyer *et al.,* 2006). The evolution of bacteriophages is believed to be driven in part by the acquisition of more and more foreign DNA, the so-called 'moron' hypothesis (Hendrix *et al.,* 2000). The existence of several host-like genes in the OtV-2 genome provides strong evidence of horizontal gene transfer and functional analysis of the genes involved should enable a greater insight into this close interaction.

All three OtV genomes encode a dTDP-d-glucose 4,6-dehydratase which shares 44% amino acid identity with the *Ostreococcus* host genus. A putative proline dehydrogenase gene with 49% amino acid identity to a host gene is also present in all three OtV genomes but is in a modified form in the OtV-2 genome. There is a proline dehydrogenase gene in the OtV-2 genome (annotated as OtV2_182, but is only 175 amino acids in length in contrast to the 'full length' 253 amino acid version found in both OtV-1 and OtV5). A single nucleotide insertion causing a frameshift appears to

have truncated the OtV-2 version prematurely. Resequencing of this region suggests this genetic alteration is real and not an artefact of the sequencing process. However, the high conservation of both DNA and translated amino acids sequence between OtV-2 and OtV-1/OtV5 in the slipped frame suggests either a strong selection pressure for this region in OtV-2 (despite the truncation) or that the alteration is a very recent event.

6.3.4.2. Nucleotide transport and metabolism

The OtV-2 genome encodes a putative deoxycytidylate deaminase (dCD) (OtV2_190), not found in OtV-1 or OtV5. This enzyme converts deoxycytidine 5'-monophosphate (dCMP) to deoxyuridine 5'-monophosphate (Maley and Maley, 1999) and is a major supplier of the substrate for thymidylate synthase, an important enzyme in DNA synthesis (Carreras and Santi, 1995). This is of significance as the OtV-2 virus also encodes a thymidylate synthase, ThyX (OtV2_051), which is also found in OtV-1 and OtV5. Both enzymes have been shown to be simultaneously elevated in rapidly dividing cells and have minimal activity in non-dividing cells (Maley and Maley, 1999). The enzyme dCD is present in most eukaryotic and bacterial organisms, with those present in humans and T4-bacteriophage being the most extensively studied. Viruses known to encode a dCD include certain bacteriophage, e.g. T4 (Maley and Maley, 1999), the chloroviruses and the Mimivirus. Members of the cytidine deaminase superfamily of enzymes have been investigated extensively in eukaryotes, as they play a role in antibody production in the immune system. As dCD is a major provider of dUMP and thymidylate synthase is the only *de novo* source of dTMP in most biological systems, these enzymes have also become potential targets for anticancer therapy (Rose *et al.,* 2002). Functional analysis of the enzyme in OtV-2 should help to determine its precise nature and activity.

6.3.4.3. Transcription

The OtV-2 genome encodes eight CDSs involved in transcription (Table 6.2). A putative RNA polymerase sigma factor (OtV-2_202) has been identified in the OtV-2 genome, which is not found in OtV-1 or OtV5. Genes encoding for RNA polymerase sigma factors have previously only been reported in bacteriophages and this is believed to be the first finding of such a gene in a virus infecting eukaryotes. The synthesis of mRNA by DNA-dependent RNA polymerases often requires one or more accessory proteins. These factors assist the polymerase in binding to the promoter selectively and initiating transcription (Burgess *et al.,* 1969). In prokaryotes, this process is regulated by the sigma (σ) - factor protein family (Paget and Helmann, 2003). RNA polymerase sigma factor is a well-characterised prokaryotic transcription factor that enables specific binding of RNA polymerase to gene promoters (Paget and Helmann, 2003). For example, in *E. coli,* the holoenzyme version of RNA polymerase is required for initiation of transcription (Bujord, 1980) and it is the sigma subunit of the enzyme that is responsible for promoter site recognition (Ring *et al.,* 1996; Barne *et al.,* 1997). The initiation of transcription from promoter elements is triggered by the reversible association of sigma factors to the complex to form a holoenzyme. The sigma-70 (σ^{70}) factors are known as the primary or major sigma factors and can initiate the transcription of a wide variety of genes. Members of the σ^{70} family are components of the RNA polymerase holoenzyme that direct bacterial or plastid core RNA polymerase to specific promoter elements that are situated 10 and 35 base-pairs upstream of transcription-initiation points (Paget and Helmann, 2003). The primary family sigma factor, which is essential for general transcription in exponentially growing cells, is reversibly associated with RNA polymerase. Sequence alignments of the σ^{70} family members reveal that they have four conserved regions, the highest conservation being found in regions 2 and 4 which are involved in binding to RNA polymerase recognising

promoters and separating DNA strands (DNA 'melting') (Paget and Helmann, 2003). Bacteriophages, such as T4, have been seen to encode a σ^{70} factor (Travers, 1969).

BLASTP analysis revealed the CDS OtV-2_202 in OtV-2 shared homology (amino acid identity of 42% and an E-value of $1e^{-09}$) with a RNA polymerase σ^{70} factor gene in the host *O. tauri* genome. An alignment of the two genes was performed and conserved domains were identified between the two genes (Fig. 6.3). All members of the σ^{70} factor superfamily exhibit region 2, the most conserved domain of this protein. The region 2 of this gene is highly conserved between organisms as it contains both the -10 promoter recognition helix and the primary core RNA *pol*-binding determinant. Analysis of the putative gene in the OtV-2 genome using the Pfam database confirmed this gene contains a recognisable region 2 (Fig. 6.3) as does the host gene, thus enabling the identification of this as a putative σ^{70} factor. The virus gene has deletions at both ends when compared to the host gene, indicating the virus has undergone deletions in its sequence over time. Moreover, the virus gene has a GC content of 37.95% but the host gene has a GC content of 60%. Therefore, if this gene was acquired by the virus from the host it is unlikely to have been a recent event. An alternative source of acquisition of this gene by the virus is from a bacterium. Aside from the closest BLASTP hit to a partial region of the gene in the host, all closest BLASTP hits are with equivalent genes from bacterial species, e.g. *Rickettsia* and *Synechococcus.* The acquisition of genes by phycodnaviruses from a bacterium has been reported (Filee *et al.,* 2007; 2008). Another route for the transfer of this gene to the virus may have been from a bacterial source to the host and then by horizontal gene transfer from host to virus. Both BLASTP and Pfam searches of the host gene gave closest hits to similar genes in cyanobacteria species, with approximate E-values of $3e^{-31}$ and amino acid identities of 33%. The only

close hit to a similar gene in a eukaryote species is to one in *Micromonas pusilla,* which is also a prasinophyte species. This finding could then lead to the formation of an alternative hypothesis to all previously outlined. The *Prasinophyceae* diverged early at the base of the Chlorophyta and consequently of the green lineage (Bhattacharya and Medlin, 1998). As members of the *Prasinophyceae*, both *Micromonas* and particularly *Ostreococcus (*with its small cell size (less than 1 µm), lack of flagella and simple cellular organisation) hold key phylogenetic positions in the eukaryotic tree of life. As a primitive species, *Ostreococcus* may have this gene from a bacterial evolutionary ancestor. A further point of interest is the absence of this gene in the OtV-1 and OtV5 genomes. If the gene was acquired by the host from a cyanobacterial source then perhaps only OtV-2 has undergone a further horizontal gene transfer event that has not occurred in the two viruses infecting the high-light host strain.

```
O.tauri          MDAEDGRRTGGDGRGTASAQAGRRAARRRRKARSEGGESGVEPGADATARSERSEETTSRTTGRGTRGGSSAERWARERALARGTSRDPS
OtV-2            ---------------------------------------------------------------------------------------

O.tauri          VRLNVDEIRTLSDALQRLLTIERVEEELRDEEEAKRLELKATLGVVEALALHEREAKDAKARFEQALARRAGFGEDSTKMRAVMRQGKIA
OtV-2            -----------------MITINATNIPIKVDEHKK-----------------------------------------------YKMNLY
                                  ::**: .:  :: :*. *                                                : ::

O.tauri          RQKLVSANIELAAKVATDVFHTFSPMERRSISRQDMIHEGVTGLVRASEKFDPALGCAFSTYAYNWTKQAVTRGVLNNGRTIRVPEQVLY
OtV-2            KKELIANHVKLAYKVSNDVYFKTYPRQRGIHTRKDINSVGLHGLVRAAQKFNPELGFKFSTYAYPWIYWSCKN----------------
                 :::*:: :::** **:.**:..  * :*   :*:*:   *: *****::**:* **  ****** *   : ..

O.tauri          ERREAHELSKRMEIELGRRPTLHEIAARSQKFSVDKLQLLVGIAKQHPLSLDILGVGDEERDGDETGLSGTWENEMGSTDEITRTIELEF
OtV-2            ---------------CLRRTTIHEELQ-------------------------------------------------------FYETPP
                               **.*:**                                                              *

O.tauri          LRDDIKKTLNELSSDEAYVLKHRFGLVGGEFMTNAEIGSRIGKSAEGVRYIEKRALERLKSTPDFSDLAQHLKGDRHAAAKKLDIVDSND
OtV-2            YYDKEPDILYGLDDVSRYILENYY----GKHLTLKDLAVELGVTVYTVTKWRDKALLHLKI----------------------------
                   *.  . *  *.. . *:*:: :    *:.:*  ::. .:* :.  *   ..:** :**

O.tauri          GTERKSAPKKPKRTKVKASKPSKSAIDLVRSIDLA
OtV-2            ----------------------------------
```

Figure 6.3. Clustal alignment of the RNA polymerase sigma factor putative genes of *O. tauri* and OtV-2. The highly conserved sigma 70_r2 region 2 in both genes is highlighted in yellow.

6.3.4.4. Protein and lipid synthesis/modification

The OtV-2 genome encodes a putative procollagen-lysine, 2-oxoglutarate 5-dioxygenase (PLOD) (OtV-2_158), an enzyme which exists in eukaryotes but has not been reported in prokaryotes or viruses until this study. This enzyme has been extensively characterised in animals where its role is to hydroxylate lysine residues in collagens leading to the resulting hydroxylysines serving as attachment sites for carbohydrate units, which are essential for the stability of intermolecular collagen crosslinks (Kivirikko *et al.,* 1992). It is also similar in its catalytic activity to prolyl-4-hydroxylase (OtV-2_106), an enzyme encoded by all three OtV virus genomes. The putative PLOD gene in OtV-2 shares approximately 30% amino acid identity with a similar gene in several eukaryotic organisms. The functional presence of this gene in the OtV-2 genome may be to ensure the stability of carbohydrate units in structures such as the capsid protein. Without functional analysis, this hypothesis remains speculative.

A putative 3-methyl-2-oxobutanoate hydroxymethyltransferase is encoded by OtV-2 (OtV-2_029). This enzyme is the first in the pantothenate biosynthesis pathway and catalyses the committing step in the synthesis of pantothenate, or vitamin B5 (Teller, 1976). Pantothenate is a necessary precursor to coenzyme A and phosphopantetheine, the prosthetic group of the acyl carrier protein, both of which are vital to a multitude of metabolic processes. Coenzyme A assists in transferring fatty acids from the cytoplasm to mitochondria (Ottenhof *et al.,* 2004). Pantenthoate is synthesized by microorganisms, i.e. bacteria and fungi, and plants but not animals, which require it as part of their diet. This biosynthesis pathway has been well studied in bacteria such as *E. coli* (Lobley *et al.,* 2003).

The putative 3-methyl-2-oxobutanoate hydroxymethyltransferase in the OtV-2 genome has closest similarity to a gene encoded by several bacterial species (amino acid identity of approximately 45% and E-value of $1e^{-60}$). Although it is not known if this gene is functional in OtV-2, it does exhibit the Rossmann fold (Rao and Rossmann, 1973), which is found in most but not all of the very large group of dehydrogenases using NAD(P)(H) as a substrate. The active site domains are also conserved in the OtV-2 gene.

Within the OtV-2 genome are three putative 2OG-Fe(II) oxygenase genes (OtV-2_165, 198, 200) and a putative prolyl 4-hydroxylase (OtV-2_106). The additional presence of two putative 2OG-Fe(II) oxygenase enzymes indicates the following reaction may be initiated by the virus:

Procollagen L-proline + 2-oxoglutarate + O_2 = procollagen trans-4-hydroxy-L-proline + succinate + CO_2.

The enzyme 2OG-Fe(II) oxygenase belongs to a class of enzymes that are widespread in eukaryotes and bacteria and catalyze a variety of reactions typically involving the oxidation of an organic substrate using a dioxygen molecule (Prescott, 1993). An extensively characterised reaction involving this enzyme is the hydroxylation of proline and lysine sidechains in collagen (Aravind and Koonin, 2001). The presence of a putative prolyl-4-hydroxylase, three putative 2OG-Fe(II) oxygenases and a procollagen-lysine, 2-oxoglutarate 5-dioxygenase in the OtV-2 genome indicates this virus may be able to stabilise complex structures, such as carbohydrate units, during its replication. Both the OtV-1 and OtV5 genomes encode only one putative 2OG-Fe(II) oxygenase each. Therefore, this begs the question why a virus infecting a low-light strain would

require three of these genes? Perhaps the assembly and stabilisation requirements of the virus capsid necessitate the involvement of these enzymes.

Of note, OtV-2 does not encode two enzymes involved in lipid metabolism, namely a patatin-like phospholipase and an oxo-acyl carrier dehydrogenase or a 3-dehydroquinate synthase, which is only encoded by the viruses infecting the high-light strain. This may be due to the OtV-2 virus utilising host genes involved in these metabolic processes, thus negating the requirement for these genes or similar in the virus genome. Alternatively, the infection process of OtV-2 may differ somewhat from other OtV viruses and thus exclude the need for genes involved in lipid synthesis.

6.3.4.5. Sugar manipulation enzymes

The OtV-2 genome encodes several proteins with high identities to enzymes involved in either manipulating sugars, synthesizing polysaccharides or transferring sugars to proteins. Some viruses affect expression of host glycosyltransferases and others express their own glycosyltransferases. The OtV-2 genome encodes at least four glycosyltransferases: two from family 1, and one each from family 2 and 25 (Table 6.4). This is two fewer than encoded by the OtV-1 genome, which also encodes an alpha galactosyltransferase not encoded by OtV5 or OtV-2. All three genomes encode a dTDP-d-glucose 4,6-dehydratase. The OtV-2 genome also encodes a putative 6-phosphogluconate dehydrogenase (OtV-2_167), not found in viruses infecting the high-light *O. tauri* strain. This enzyme catalyses the decarboxylating reduction of 6-phosphogluconate into ribulose 5-phosphate, in the presence of NADP (Wood, 1986). This reaction is part of the hexose mono-phosphate shunt and pentose phosphate pathways. Virus DNA can be synthesised from the intermediates of the reductive pentose phosphate pathway during photosynthesis, from the intermediates of the

oxidative pentose phosphate pathway, and also from nucleotide precursors degraded from host DNA (Sindelar *et al.,* 1999). The oxidative pentose phosphate pathway metabolises glucose-6-phosphate to ribose-5-phosphate, which is necessary for the *de novo* biosynthesis of purine and pyrimidine nucleotides of viral DNA (Sindelar, 1986). Prokaryotic and eukaryotic 6-phosphogluconate dehydrogenase enzymes are proteins of approximately 470 amino acids with highly conserved sequences. The gene in the OtV-2 genome has approximately 30% amino acid identity and e-values of approximately $2e^{-50}$ with several bacterial species. It is possible the virus acquired this gene from a bacterial source. The only other virus reported to encode this enzyme is *Synechococcus* phage syn9 (Weigele *et al.,* 2007) and this is therefore only the second report of this enzyme in a virus. Presumably, the majority of viruses which do not encode this gene hijack the host's biosynthesis pathway for nucleotide synthesis during infection. The OtV-2 genome may possess this gene due to the niche the host occupies making a viral-encoded version a beneficial acquisition.

6.3.4.6. Miscellaneous

OtV-2 has several genes encoding other putative proteins, including a rhodanese domain-containing protein (OtV-2_203) that may function as a sulphurtransferase, as rhodaneses are widespread enzymes that catalyse the transfer of a sulphur atom from thiosulphate to cyanide (Westley *et al.,* 1983). Similar to OtV-1 and OtV5, OtV-2 encodes an ABC domain protein (OtV-2_052), which may play a role in translation inititation (Andersen and Leevers, 2007). A large putative CDS in the OtV-2 genome has close homology to a virus-like inclusion body encoded by the protist *Trichomonas vaginalis* (OtV-2_109). The OtV-2 CDS is 2037 amino acids in length with a G+C content of 41.75%. This gene in OtV-2 aligns with three consecutive CDSs in the reference genome, OtV-5 and a single large CDS (3398 amino acids long and G+C

content of 43.76%) in the OtV-1 genome. Resequencing of the large gene in OtV-1 and OtV-2 confirmed these genes are single CDSs kept in frame. No conserved domains were identified within these CDSs and further functional work is required.

A putative tail assembly fibre protein (OtV-2_065) is encoded by OtV-2 but not seen in the other OtV genomes. A PSI-BLAST of this gene against the public databases resulted in close hits to bacteriophage tail fibre assembly protein genes with an average amino acid identity of 30% and an E-value of $2e^{-32}$. TEM images of the OtV-2 particles indicate the possible presence of tail-like stubs (Fig. 6.4). This now requires further structural characterisation work.

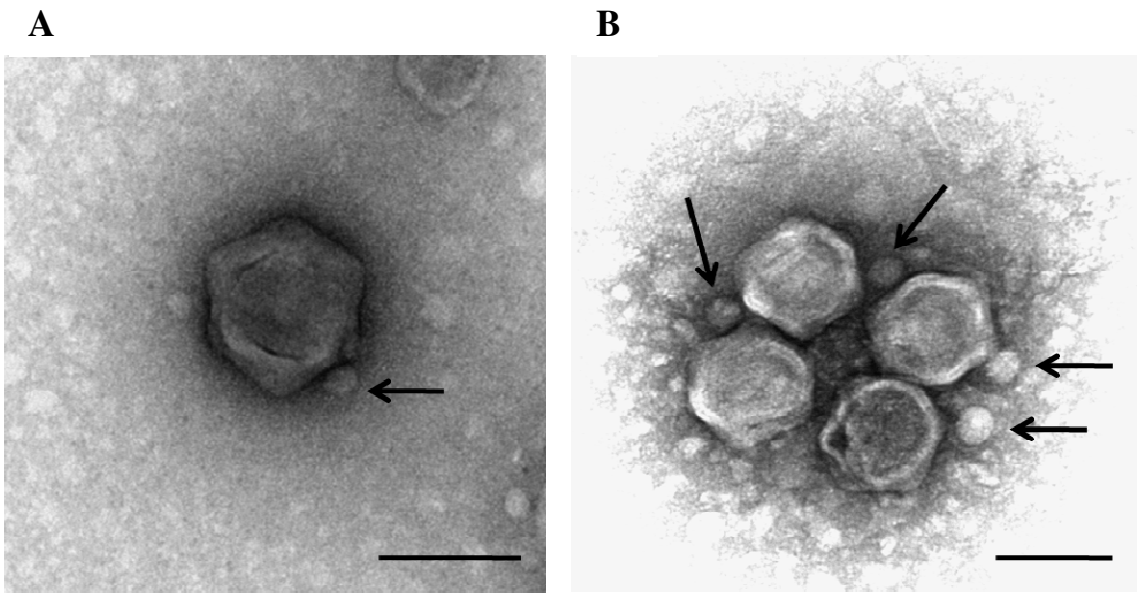


Figure 6.4. Transmission electron micrographs of negatively stained whole mounts of *Ostreococcus tauri* virus–2 (OtV-2). A tail-like stub is indicated by arrows in A and B. Scale bars represent 100 nm.

A putative phosphate starvation inducible protein of the PhoH protein family was identified in the OtV-2 genome (OtV-2_024) and has now been identified in all three genomes. PhoH is a predicted ATPase protein located in the cytoplasm that is induced by phosphate starvation (Kim *et al.,* 1993). The phosphate starvation inducible gene in OtV-2 does not have a homologue in the host and may have originated from a bacterial source, as it shares 47% identity with a similar protein encoded by the bacterium *Thermosipho melanesiensis* and shares similar homology to several other bacterial encoded PhoH family proteins. Phosphorus is a required macronutrient in the marine environment, as it is incorporated into nucleic acids, phospholipids and glycolytic intermediates. Phosphorus is a key element in energy transfer reactions and helps regulate many signal transduction cascades (Alberts *et al.,* 2002). Therefore, it is unsurprising that phytoplankton have evolved the ability to utilise phosphorus from diverse sources to cope with low nutrient availability in environments such as oligotrophic oceans (Cembella, 1984; Antia, 1991).

Genes of the phosphate (*pho*) regulon in *Escherichia coli* are activated as a response to phosphate limitation. The products of these genes are involved in transport and use of various forms of combined phosphates or free phosphate (Torriani, 1990; Kim *et al.,* 1993). Homologues of PstS the phosphate binding protein in *E. coli.,* were identified in two marine cyanobacterial genera, *Synechococcus* and *Prochlorococcus* (Scanlan *et al.,* 1993; 1997). In addition to a putative PhoH gene, a surprising finding was the presence of a putative high affinity phosphate transporter gene (OtV-2_222) in the OtV-2 genome. A database BLASTP search for other viruses with homology to the gene in OtV-2 gave only one other match to a putative phosphate-repressible phosphate permease (PPRPP) in the coccolithovirus, EhV-86. The putative high affinity phosphate

transporter gene reported in OtV-2 encodes a PHO4 family protein. With an amino acid identity of 57% to the host gene, it is likely the virus has acquired this gene from its *Ostreococcus* host. This gene has been identified and characterised in other phytoplankton species, such as *Tetraselmis chui* (Chung *et al.,* 2003). A phylogenetic analysis of the high affinity phosphate transporter gene in OtV-2 with equivalent genes in *Ostreococcus* species, other algal species and a 'deep ecotype' bacterium *Alteromonas macleodii* was performed (Fig. 6.5). The OtV-2 gene forms a distinct cluster with the host genes (bootstrap value of 100% for 1,000 trials). From these results, the gene appears to have been acquired by the virus through horizontal transfer from the host. A ClustalW alignment of the putative high affinity phosphate transporter gene in the virus and the equivalent genes in the host species' *O. tauri* and *O. lucimarinus* shows a degree of conservation between the three genes (Fig. 6.6). However, the virus gene appears to have undergone a 162 bp in-frame deletion in the middle area of the gene. The region in the host genes corresponding to this deleted region does not contain any conserved domains and it is unknown if this region plays a key functional role in P uptake. As all three sequenced OtV genomes to date encode a PhoH family protein, the OtV viruses may thus be able to increase phosphate availability if one or both of these phosphate metabolism-related genes are activated and expressed during infection. This would then ensure sufficient levels of phosphate turnover in the host cell to facilitate transcription and production of progeny virions.

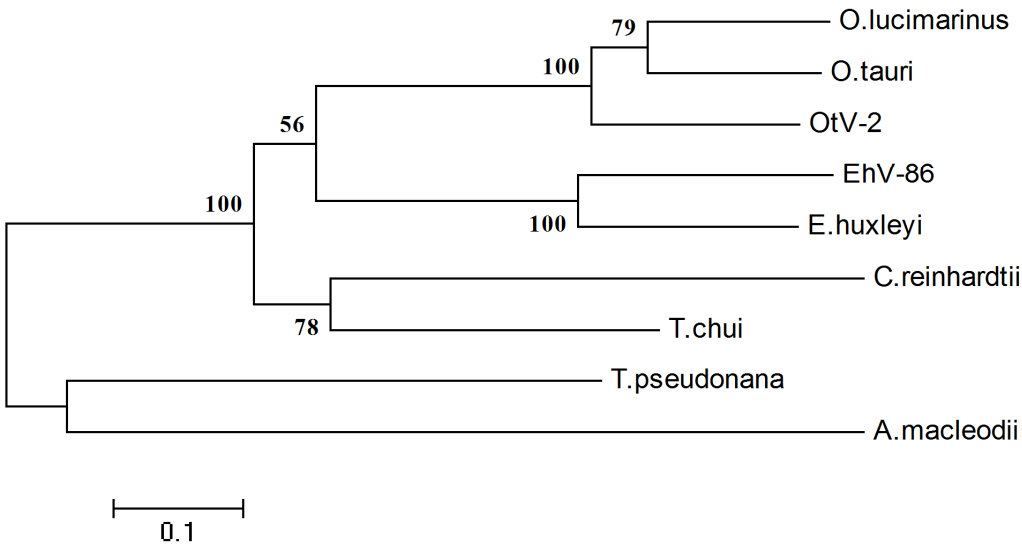


Figure 6.5. Phylogenetic tree of OtV-2 high affinity phosphate transporter gene. GenBank accession numbers are as follows: *Alteromonas macleodii,* NC_011138.1*; Chlamydomonas reinhardtii,* XP_001695776.1; *Emiliania huxleyi,* AF334403.1; *Emiliania huxleyi* virus-86 (EhV-86), YP_293871.1; *Ostreococcus lucimarinus,* XP_001416412.1; *Ostreococcus tauri,* CAL52329; *Tetraselmis chui,* AAO47330.1; *Thalassiosira pseudonana,* XP_002281698.1; OtV-2, this study. Scale bar indicates a distance of 0.1 fixed mutations per amino acid position.

Figure 6.6. ClustalW alignment of high affinity phosphate transporter genes of OtV-2, *O.tauri* and *O. lucimarinus.* Dots (.) represent consensus sequences. Dashes (-) represent gaps.

Unexpectedly, a putative cytochrome b_5 gene (OtV-2_201) was identified in the OtV-2 genome, which shares homology with a similar gene encoded by the host *O. tauri* (Table 6.7). Cytochrome b_5 is a small hemoprotein and a ubiquitous electron transport carrier found in animals, yeasts and plants. The role cytochrome b_5 plays in the synthesis of unsaturated fatty acids has been well-characterised in animals and plants (Smith, 1998). Cytochrome b_5 consists of two domains, namely a hydrophobic tail that anchors the protein to the membrane and a hydrophilic portion, the heme-binding domain, which is active in redox reactions (Vergères and Waskell, 1995). In animals, the membrane-bound cytochrome b_5 is located in the endoplasmic reticulum where it can accept an electron from either NADH-cytochrome b_5 reductase or NADPH-cytochrome b_5 P450 reductase. Reduced cytochrome b_5 then provides reducing equivalents for the biosynthesis of selected lipids and donates the second electron in the reaction cycle to certain members of the superfamily of mixed function oxidases known as cytochrome P450 (Peterson, 1986). Hence, cytochrome b_5 is involved in redox reactions such as the desaturation of acyl-coA substrates and cholesterol biosynthesis (Ozols, 1989). In higher plants, cytochrome b_5 has been clearly shown to be involved in the desaturation of acyl-complex lipids in oilseeds (Smith *et al.,* 1990) and the related process of fatty acid hydroxylation (Smith *et al.,* 1992).

BLASTP analysis of the OtV-2 gene against the public databases resulted in an amino acid identity of 60% and an e-value of $2e^{-29}$ to a cytochrome b_5 gene in *O. tauri.* An alignment of the putative cytochrome b_5 gene in OtV-2 and the similar gene in *O. tauri* shows the virus gene aligns with a partial region of the host gene. Across most of this partial alignment, there is a high level of conservation between the two genes. However, there is an apparent deletion in the virus gene near the 5' end of the virus

gene (Fig. 6.7). The core of the heme-binding domain is formed by the characteristic cytochrome b_5 motif His-Pro-Gly-Gly (HPGG) (Smith, 1998). This diagnostic cytochrome b_5 motif 2 of the heme-binding domain was identified in the OtV-2 cytochrome b_5 gene, at residues 36-39 of the 91 aa polypeptide and the *O. tauri* host cytochrome b_5 gene at residues 545-548 of the 586 aa polypeptide. From the evidence presented, there is a possibility the cytochrome b_5 gene in OtV-2 was acquired from the algal host, particularly as the adjacent gene encodes a host-like RNA polymerase sigma factor. The only other identification of a putative cytochrome b_5 gene in a virus is in *Acanthamoeba polyphaga* Mimivirus (Raoult *et al.,* 2004). This is therefore only the second virus reported to encode a putative cytochrome b_5.

```
O._tauri        MPDARESSTNDRVVTRRRRSRSIAMAEQVIVIGGGLAGLSAAHTVLEHGARVVLLDKCSFLGGNSTKATSGINGALTRTQARLNIPDSAD
OtV-2           ----------------------------------------------------------------------------------------
O._tauri        KFEDDIIKGAAGVGHTEAPAHTIPLAKVLAQGSGSSVDWLCEKFKLDLSLVAQLGGHSYPRTHRGKERFPGFTITYALMEGLEKVMEDSN
OtV-2           ----------------------------------------------------------------------------------------
O._tauri        GETARIITKAEAKRLLTDGSGTVIGVEYEKDGVLNQEYGPVVIATGGFGADYKPDSLLKKYRPDLQALPTTNGDHCTGDGIKMAMAVGAD
OtV-2           ----------------------------------------------------------------------------------------
O._tauri        TIDMTSVQVHPTGLVNPAEPDSKVKFLAAEALRGVGGILLDANGNRFADELGRRDYVSGEMNRNKGPFRLILNGKASTEIEWHCKHYVGR
OtV-2           ----------------------------------------------------------------------------------------
O._tauri        GIMKRHDSGAEVAKELGISPQKLADTFAKYNEAARTKNCPFGKKFFTNAPFEMNDFFHSAIVCTVVHYTMGGLAINTDSQIVGPRGPIPG
OtV-2           ----------------------------------------------------------------------------------------
O._tauri        LFGAGEVVGGIHGRNRLGGNSLLDCVVFGRVAGSAVSRHLMSTAIRALRSGQTTAMNRVANLNDKINPPAMSAAPAAASAASGGGSRALT
OtV-2           ------------------------------------------------------MNRIKTINDHINP-----------------RDLS
                                                                        ***: .:**:***                 * *:

O._tauri        MDEINKHNTEGDLWVIIEGNVYDLTKFLPDHPGGKKAIMLFAGKDATEEFNMLHPPNVLKKYLSPDAKIGTVLG
OtV-2           LTEIAKHNTEEDCWVIIKDIVYDLTKFLPDHPGGKKAIILFAGKDATEEFDMLHPPNVLKKYLTPEVVLGPVKK
                : ** ***** * ****:. ******************:***********:************:*:. :*.*
```

Figure 6.7. Alignment of the cytochrome b_5 genes encoded by *Ostreococcus tauri* and the virus OtV-2. Conserved amino acids are identified by an asterisk underneath the corresponding residue. The conserved heme-binding domain is highlighted in red.

6.4. Conclusions

The 184,409 bp OtV-2 genome is predicted to encode 237 proteins and 5 tRNAs. The putative protein-encoding genes are relatively evenly distributed on both strands and intergenic space is minimal. Approximately 30% of the gene products have been identified, some of which resemble prokaryotic proteins whilst others have eukaryotic characteristics. Of the latter, 14 predicted genes share close similarity with genes found in the host *Ostreococcus tauri* genome. Interesting highlights include a putative RNA polymerase sigma factor, a cytochrome b_5 and at least two genes predicted to increase uptake of phosphate into the cell. All the genes identified in the OtV-2 genome will require transcriptional analysis to further investigate their putative functions. The current analysis of genes in the OtV-2 genome has revealed some of these host genes may have had bacterial origins either through horizontal gene transfer events or a bacterial ancestor in the host lineage. This is an ideal system for future study as it is an interesting viral/eukaryotic system but with strong roots in bacteriology.

Chapter 7
Summary and future work

7.1. Summary

Prior to reports of high virus abundances in natural seawater samples (Bergh *et al.,* 1989; Suttle *et al.,* 1990), the concentration of viruses in the marine environment was generally perceived to be low and of no significant consequence. Since these findings 20 years ago, the importance of viruses as agents of mortality and the controlling influence they exert on population structure and diversity of phytoplankton communities has been demonstrated (Cottrell and Suttle, 1991; Bratbak *et al.,* 1993; Brussaard *et al.,* 1996b). Moreover, the biogeochemical and ecological effects of viruses in the sea are now generally accepted (reviewed by Fuhrman, 1999; Brussaard, 2004b; Weinbauer, 2004). Since viruses were first isolated that infect the marine phytoplankton *Micromonas pusilla* (Mayer and Taylor, 1979), an ever expanding literature has reported the isolation and characterisation of other marine eukaryotic algal viruses (Table 1.1) revealing a broad genetic diversity (Dunigan *et al.,* 2006). Many aspects of the underlying biology of phytoplankton-virus interactions remain unknown or poorly understood. Understanding the dynamics of host-virus relationships and their interactions in the natural environment is important in terms of both ecological and biogeochemical implications. In order to gain such insights, more host-virus systems need to be identified and isolated for study under laboratory conditions.

The work conducted during this PhD was mainly concerned with the isolation and characterisation of novel viruses infecting phytoplankton in the size range 0.2 µm to 20 µm, i.e. the nano- and picophytoplankton assemblages. In particular, the thesis looked in depth at viruses infecting *Ostreococcus tauri* and *Micromonas pusilla.* The use of traditional characterisation techniques coupled with molecular tools have led to important findings that improve our understanding of the relationships between

microalgal hosts and viruses, including horizontal gene transfer, evolutionary history, and propagation strategies.

7.2.. The detection and isolation of viruses infecting marine nano- and picophytoplankton

The screening of natural seawater sampled in the Western English Channel resulted in the isolation of 8 viruses that infected six strains of *O. tauri* and four viruses that infected a single *M. pusilla* species. Although lysis was observed in certain other algal species, clonal isolation of the lytic agent was not possible. Lysogenic infections were not detected in this study. All viral infections of microalgae described to date are lytic and lysogeny has yet to be described in microalgal hosts.

The relative success of isolating viruses for *O. tauri* and *M. pusilla* in this study may be in part due to their ubiquitous occurrence in marine waters (Zhu *et al.,* 2005; Worden, 2006). Both of these phytoplankton species are clade II Prasinophytes, which are more abundant in coastal waters, as found at the sampling site in this study. Detection and clonal isolation of viruses was most successfully achieved using concentrated seawater samples mostly sampled during the spring and summer months, rather than in the winter season. Although no metadata was associated with this study, the greater success in virus isolation at these times may be attributed to increased temperature, daylight and stratification leading to increased total phytoplankton biomass. Virus numbers have been seen to increase in correlation with increased host numbers related to seasonal changes (Cottrell and Suttle, 1991; Zingone *et al.,* 1999).

7.3. Characterisation of viruses infecting *Ostreococcus tauri* and *Micromonas pusilla*

The viruses detected and successfully isolated in Chapter 3 were characterised, using molecular and traditional techniques. Lower predicted burst sizes seen for OtV viruses in comparison to those seen in infections by MpV may be explained in terms of the smaller host size and restricted space of the cytoplasmic compartment. TEM analysis also revealed slightly smaller virus particle sizes of OtV isolates compared to MpV strains. Aside from a small difference in size the OtV and MpV viruses are morphologically similar, with an icosahedral capsid with regular symmetry and an electron dense core, as reported for members of the family of algal viruses, the *Phycodnaviridae*. The host range analysis showed a very high strain specificity of all isolates in this study, infecting only a single strain of each host for which they had been initially isolated for. Assessing the results of host range study with morphological characteristics and the phylogenetic analysis of the DNA polymerase genes of virus isolated from this study, these OtV and MpV isolates can be putatively assigned to the prasinovirus genus of the *Phycodnaviridae* family. Data are now becoming increasingly available for this genus, which has previously been poorly understood and somewhat overlooked. The predicted genome sizes for the OtV and MpV isolates are not the largest reported for phycodnaviruses and can be placed in the smaller size bracket for members of this family.

Analysis of the adsorption mechanism of OtV-1 in this study revealed the virus fuses with the host membrane and the empty capsid shell remains irreversibly attached during the remainder of the infection process. The increase in size of the empty capsid may be a result of conformational changes during ejection of the nucleic acid contents. The possible 'stargate'-like structure seen in TEM images of the MpV-3 virus (Fig. 4.13 e)

also hints at a possible conformational change in the capsid structure to deliver the DNA via a portal into the host.

The observed re-growth of the *O. tauri* host RCC 501 raises many questions as to the underlying causes of this phenomenon. These include resistance emerging in a strain or sub-population of a sample. In another system, namely the infection of the coccolithophorid *Emiliania huxleyi* by EhV-86, the production of reactive oxygen species (ROS) by the infected host may lead to apoptosis of infected cells only, leaving a population of uninfected cells remaining (Evans *et al.,* 2006). Metacaspase expression is also implicated in this host-virus system, as a resistant algal population may not activate these enzymes whose expression is required by viruses to facilitate their replication process (Bidle *et al.,* 2007). Until the underlying mechanisms of algal resistance are determined, such hypotheses are only speculative and the signs of re-growth reported in the *O. tauri* strain RCC 501 in this study require further investigation.

7.4. The complete genome sequence and annotation of *Ostreococcus tauri* Virus-1, OtV-1

The OtV-1 virus that infects *O. tauri* strain OTH 95 was sequenced and annotated. The identification of only five of the Group I core genes of the classified NCLDVs raises questions about the current grouping of these genes, which are shared by viruses believed to have a common evolutionary ancestor. The fact that certain Group II genes are shared amongst all members may require these to be reclassified as Group I with the genes lacking in the OtV-1 genome belonging to Group II.

Several genes have been identified in the OtV-1 genome that have not been reported in a virus previously, or only in very few viruses, raising important questions regarding their function and origins. It is apparent from this study, horizontal gene transfer events have occurred between the host and virus, as well as between bacterial and other eukaryotic sources and OtV-1. With a number of virus genes having close homology to genes in the algal host, an exchange of genetic material is occurring and possibly in either direction. The unexpected presence of an intein in the DNA polymerase gene of OtV-1 highlights an intriguing phenomenon. The identification of domains within this intein indicates it is active and could potentially insert into other genomes. This type of event may also contribute to gene transfer and transduction. The presence of several genes encoding capsid proteins indicates a possible complex capsid structure and adds weight to the presence of a portal via which viral DNA becomes encapsidated and released. A putative PhoH family protein gene in OtV-1 would enable the virus to increase phosphate turnoverduring infection and, this would make intuitive sense given the high P demand during virus propagation.

7.5. The genome sequence and annotation of the *Ostreococcus tauri* Virus-2, OtV-2

A second OtV virus (OtV-2) was sequenced and annotated, then compared with OtV-1. These two viruses infect different host strains which exhibit light regime niche partitioning; OtV-1 infects a high-light adapted strain and OtV-2 infects a low-light adapted strain. The information yielded by the sequencing of these two virus genomes in addition to the genome of a third virus, OtV5, which infects the same host strain as OtV-1, has revealed some exciting similarities and differences. All three viruses encode 8 capsid proteins with close genetic homology, hinting at some commonality in capsid structure. The OtV-2 genome shares close homology to 14 genes in the host, which is the most shared by any of the three viruses. OtV-2 appears to have a stronger

propensity for acquisition of genes from the host. OtV-2 also encodes proteins not identified in the other two virus genomes and these may reflect the differences in environmental conditions experienced by the different host strains. Only OtV-1 contains an intein in the DNA polymerase gene. OtV-1 also has the greatest number of glycosyltransferases, which may reduce its reliance on host enzymes for glycosylation of the capsid proteins. All three virus genomes encode for important enzymes involved in DNA replication and repair, transcription and nucleotide metabolism and transport. There are between three and six methyltransferases encoded by these viruses, whose biological function is unknown. In bacteria, restriction-modification systems serve to confer a level of protection from endonucleolytic attack by foreign DNA and DNA viruses (Essani *et al.,* 1987). It has been speculated that methyltransferases encoded by other phycodnaviruses may protect virus DNA from similar enzyme activity (Van Etten, 2003).

The results arising from this study offer a greater insight than previously held into the prasinoviruses and their associations with their hosts. Functional analysis of features of their genomes should address some key questions that have emerged during this study.

7.6. Future work

Intriguing observations of infection strategies and capsid structure using TEM imagery require further scrutiny. Structural studies of the capsid and its adsorption could be conducted using, for example, cryo-EM techniques and X-ray crystallography. Further studies on the active status of the OtV-1 intein are warranted. These would include experiments to assess the potential for the intein to insert into other virus genomes, initially testing this with other OtV isolates. The nature of resistance seen in *O. tauri* strain RCC 501 requires more in-depth enquiry, as do the underlying causes of this phenomenon in algal hosts. The viral genome sequence data generated in this study can now be used to design novel primers for use in further investigations, for example the enumeration of the level of infection *in situ* or the assessment of genetic diversity within this group of viruses. Molecular tools, such as microarrays, could be employed to aid such studies, particularly to analyse the transcriptional profile of these viruses. The importance of these viruses can then be evaluated in the natural environment and clearer insights gained of the diversity in viral communities. Confirming the function of CDSs in the virus genomes with identified putative functions is an obvious first step in the functional analysis of genome features. The following major challenge will be the attribution of function to those CDSs for which no function has been identified in the preliminary analysis of these viral genomes. This will necessitate the use of techniques such as proteomics and transcriptional analysis. Exploitable traits could then be screened for. Subsequent results will also help reveal the evolutionary history of the viruses isolated here and their interactions with algal hosts and the marine environment.

Chapter 8

References

REFERENCES

Ackermann, H.W., and DuBow, M.S. (1987) Viruses of Prokaryotes. In *General Properties of Bacteriophages*. Boca Raton: CRC Press, p. 202.

Agarkova, I.V., Dunigan, D.D., and Van Etten, J.L. (2006) Virion-associated restriction endonucleases of chloroviruses. *Journal of Virology* **80**: 8114-8123.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (eds) (2002) *Molecular biology of the cell*. New York: Garland Science.

Allen, M.J., and Wilson, W.H. (2006) The coccolithovirus microarray: an array of uses. *Briefings in Functional Genomics and Proteomics* **5**: 273-279.

Allen, M.J., and Wilson, W.H. (2008) Aquatic virus diversity accessed through omic techniques: A route map to function. *Current Opinion in Microbiology* **11**: 226-232.

Allen, M.J., Schroeder, D.C., and Wilson, W.H. (2006a) Preliminary characterisation of repeat families in the genome of EhV-86, a giant algal virus that infects the marine microalga *Emiliania huxleyi*. *Archives of Virology* **151**: 525-535.

Allen, M.J., Schroeder, D.C., Holden, M.T.G., and Wilson, W.H. (2006b) Evolutionary history of the Coccolithoviridae. *Molecular Biology and Evolution* **23**: 86-92.

Allen, M.J., Schroeder, D.C., Donkin, A., Crawfurd, K., and Wilson, W.H. (2006c) Genome comparison of two Coccolithoviruses. *Virology Journal* **3**: 15.

Allen, M.J., Forster, T., Schroeder, D.C., Hall, M., Roy, D., Ghazal, P., and Wilson, W.H. (2006d) Locus-specific gene expression pattern suggests a unique propagation strategy for a giant algal virus. *Journal of Virology* **80**: 7699-7705.

Allen, M.J., Martinez-Martinez, J., Schroeder, D.C., Somerfield, P.J., and Wilson, W.H. (2007) Use of microarrays to assess viral diversity: from genotype to phenotype. *Environmental Microbiology* **9**: 971-982.

Alliegro, M.C., and Alliegro, M.A. (2007) Echinonectin is a Del-1-like molecule with regulated expression in sea urchin embryos. *Gene Expression Patterns* **7**: 651-656.

Andersen, D.S., and Leevers, S.J. (2007) The essential *Drosophila* ABC domain protein, pixie, binds the 40S ribosome in an ATP-dependent manner and is required for translation initiation. *Journal of Biological Chemistry* **282**: 14752-14760.

Angly, F.E., Felts, B., Breitbart, M., Salamon, P., Edwards, R.A., Carlson, C., Chan, A.M., Haynes, M., Kelley, S., Liu, H., Mahaffy, J.M., Mueller, J.E., Nulton, J., Olson, R., Parsons, R., Rayhawk, S., Suttle, C.A. and Rohwer, F. (2006) The marine viromes of four oceanic regions. *Plos Biology* **4**: 2121-2131.

Antia, N.J., Harrison, P.J., Oliveira, L. (1991) The role of organic dissolved nitrogen in phytoplankton nutrition, cell biology and ecology. *Phycologia* **30**: 1-89.

Aravind, L., and Koonin, E.V. (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. *Genome Biol* **2**: Research 0007.

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyerreil, L.A., and Thingstad, F. (1983) The ecological role of water column microbes in the sea. *Marine Ecology-Progress Series* **10**: 257-263.

Baldo, A.M., and McClure, M.A. (1999) Evolution and horizontal transfer of dUTPase-encoding genes in viruses and their hosts. *Journal of Virology* **73**: 7710-7721.

Bamford, D.H., Caldentey, J., and Bamford, J.K.H. (1995) Bacteriophage PRD1 - A broad host-range dsDNA Tectivirus with an internal membrane. In *Advances in Virus Research, Vol 45*. San Diego: Academic Press Inc, pp. 281-319.

Bamford, D.H., Burnett, R.M., and Stuart, D.I. (2002) Evolution of viral structure. *Theoretical Population Biology* **61**: 461-470.

Baranowski, E., Ruiz-Jarabo, C.M., and Domingo, E. (2001) Evolution of cell recognition by viruses. *Science* **292**: 1102-1105.

Barne, K., Bown, J., Busby, S., and Minchin, S. (1997) Region 2.5 of the *Escherichia coli* RNA polymerase sigma 70 subunit is responsible for the recognition of the 'extended-10' motif at promoters. *EMBO Journal* **16**: 4034 - 4040.

Baudoux, A.C., and Brussaard, C.P.D. (2005) Characterization of different viruses infecting the marine harmful algal bloom species *Phaeocystis globosa*. *Virology* **341**: 80-90.

Baudoux, A.C., Noordeloos, A.A.M., Veldhuis, M.J.W., and Brussaard, C.P.D. (2006) Virally induced mortality of *Phaeocystis globosa* during two spring blooms in temperate coastal waters. *Aquatic Microbial Ecology* **44**: 207-217.

Bawden, A.L., Glassberg, K.J., Diggans, J., Shaw, R., Farmerie, W., and Moyer, R.W. (2000) Complete genomic sequence of the *Amsacta moorei* entomopoxvirus: analysis and comparison with other poxviruses. *Virology* **274**: 120-139.

Beijerinck, M.W. (1898) Concerning a contagium vivum fluidum as a cause of the spot-disease of tobacco leaves. *Verh Akad Wet Amsterdam* **II**: 3-21.

Bellec, L., Grimsley, N., Moreau, H., Desdevises, Y. (2009) Phylogenetic analysis of new Prasinoviruses (*Phycodnaviridae)* that infect the green unicellular algae *Ostreococcus, Bathycoccus* and *Micromonas. Environmental Microbiology Reports* **1**: 114-123.

Bender, S.L., Mehdi, S., and Knowles, J.R. (1989) Dehydroquinate synthase - the role of divalent metal cations and of Nicotinamide Adenine Dinucleotide in catalysis. *Biochemistry* **28**: 7555-7560.

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

Bernal, J.D., and Fankuchen, I. (1941) X-ray and crystallographic studies of plant virus preparations. *Journal of General Physiology* **25**: 147.

Best, S.M., and Bloom, M.E. (2004) Caspase activation during virus infection: more than just the kiss of death? *Virology* **320**: 191-194.

Bettarel, Y., Kan, J., Wang, K., Williamson, K.E., Cooney, S., Ribblett, S., Chen, F., Wommack, K.E., and Coats, D.W. (2005) Isolation and preliminary characterisation of a small nuclear inclusion virus infecting the diatom *Chaetoceros* cf. *gracilis*. *Aquatic Microbial Ecology* **40**: 103-114.

Bhattacharya, B., Giri, N., Mitra, M., and Das Gupta, S.K. (2008) Cloning, characterization and expression analysis of nucleotide metabolism-related genes of mycobacteriophage L5. *FEMS Microbiology Letters* **280**: 64-72.

Bhattacharya, D., and Medlin, L. (1998) Algal phylogeny and the origin of land plants. *Plant Physiology* **116**: 9-15.

Bidle, K.D., Haramaty, L., Ramos, J.B.E., and Falkowski, P. (2007) Viral activation and recruitment of metacaspases in the unicellular coccolithophore *Emiliania huxleyi*. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 6049-6054.

Black, L.W. (1988) DNA packaging in dsDNA Bacteriophages. In *The Bacteriophages*. Calendar, R. (ed). New York: Plenum Press, pp. 321 - 373.

Blanchot, J., and Rodier, M. (1996) Picophytoplankton abundance and biomass in the western tropical Pacific Ocean during the 1992 El Nino year: Results from flow cytometry. *Deep-Sea Research Part I-Oceanographic Research Papers* **43**: 877-895.

Bratbak, G., Heldal, M., Thingstad, T.F., Riemann, B., and Haslund, O.H. (1992) Incorporation of viruses into the budget of microbial C Transfer - a 1st approach. *Marine Ecology-Progress Series* **83**: 273-280.

Bratbak, G., Egge, J.K., and Heldal, M. (1993) Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal blooms. *Marine Ecology Progress Series* **93**: 39-48.

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

Bratbak, G., Jacobsen, A., and Heldal, M. (1998a) Viral lysis of *Phaeocystis pouchetii* and bacterial secondary production. *Aquatic Microbial Ecology* **16**: 11-16.

Bratbak, G., Jacobsen, A., Heldal, M., Nagasaki, K., and Thingstad, F. (1998b) Virus production in *Phaeocystis pouchetii* and its relation to host cell growth and nutrition. *Aquatic Microbial Ecology* **16**: 1-9.

Brautigam, M., Klein, M., Knippers, R., and Muller, D.G. (1995) Inheritance and meiotic elimination of a virus genome in the host *Ectocarpus siliculosus* (Phaeophyceae). *Journal of Phycology* **31**: 823-827.

Breitbart, M., Miyake, J.H., and Rohwer, F. (2004a) Global distribution of nearly identical phage-encoded DNA sequences. *FEMS Microbiology Letters* **236**: 249-256.

Breitbart, M., Felts, B., Kelley, S., Mahaffy, J.M., Nulton, J., Salamon, P., and Rohwer, F. (2004b) Diversity and population structure of a near-shore marine-sediment viral community. *Proceedings of the Royal Society of London Series B-Biological Sciences* **271**: 565-574.

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

Brown, C.W., and Yoder, J.A. (1994) Coccolithophorid blooms in the global ocean. *Journal of Geophysical Research-Oceans* **99**: 7467-7482.

Brussaard, C.P.D., Gast, G.J., vanDuyl, F.C., and Riegman, R. (1996a) Impact of phytoplankton bloom magnitude on a pelagic microbial food web. *Marine Ecology-Progress Series* **144**: 211-221.

Brussaard, C.P.D., Kempers, R.S., Kop, A.J., Riegman, R., and Heldal, M. (1996b) Virus-like particles in a summer bloom of *Emiliania huxleyi* in the North Sea. *Aquatic Microbial Ecology* **10**: 105-113.

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

Brussaard, C.P.D. (2004b) Viral control of phytoplankton populations - a review. *Journal of Eukaryotic Microbiology* **51**: 125-138.

Brussaard, C.P.D., Noordeloos, A.A.M., Sandaa, R.A., Heldal, M., and Bratbak, G. (2004) Discovery of a dsRNA virus infecting the marine photosynthetic protist *Micromonas pusilla*. *Virology* **319**: 280-291.

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

Brussaard, C., Bratbak, G., Baudoux, A.-C., and Ruardij, P. (2007) *Phaeocystis* and its interaction with viruses. *Biogeochemistry* **83**: 201-215.

Brussaard, C.P.D., and Martinez Martinez, J. (2008) Algal bloom viruses. *Plant viruses* **2**: 1-13.

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

Bubeck, D., Filman, D.J., and Hogle, J.M. (2005) Cryo-electron microscopy reconstruction of a poliovirus-receptor-membrane complex. *Nature Structural & Molecular Biology* **12**: 615-618.

Bubeck, J.A., and Pfitzner, A.J.P. (2005) Isolation and characterization of a new type of chlorovirus that infects an endosymblotic *Chlorella* strain of the heliozoon *Acanthocystis turfacea*. *Journal of General Virology* **86**: 2871-2877.

Bujord, H. (1980) The interaction of *E. coli* RNA polymerase with promoters. *Trends in Biochemical Sciences* **5**: 274-278.

Burbank, D.E., Shields, S.L., Schuster, A.M., and Van Etten, J.L. (1990) 5-azacytidine-resistant mutants of *Chlorella* virus IL-3A. *Virology* **176**: 311-315.

Burgess, R., Travers, A., Dunn, J., and Bautz, E. (1969) Factor stimulating transcription by RNA polymerase. *Nature* **221**: 43 - 46.

Butcher, R. (1952) Contributions to our knowledge of the smaller marine algae. *Journal of the Marine Biological Association of the United Kingdom* **31**: 175-191.

Butler, P.J.G., and Klug, A. (1971) Assembly of the particle of TMV from RNA and disk of protein. *Nature New Biology* **229**: 47.

Cardol, P., Bailleul, B., Rappaport, F., Derelle, E., Beal, D., Breyton, C. Bailey, S., Wollman, F.A., Grossman, A., Moreau, H., and Finazzi, G. (2008) An original adaptation of photosynthesis in the marine green alga *Ostreococcus*. *Proceedings of the National Academy of Sciences of the United States of America* **105**: 7881-7886.

Cardone, G., Winkler, D.C., Trus, B.L., Cheng, N.Q., Heuser, J.E., Newcomb, W., Brown, J.C., and Steven, A.C. (2007) Visualization of the herpes simplex virus portal in situ by cryo-electron tomography. *Virology* **361**: 426-434.

Carreras, C.W., and Santi, D.V. (1995) The catalytic mechanism and structure of thymidylate synthase. *Annual Review of Biochemistry* **64**: 721-762.

Carver, T.J., Rutherford, K.M., Berriman, M., Rajandream, M.A., Barrell, B.G., and Parkhill, J. (2005) ACT: the Artemis comparison tool. *Bioinformatics* **21**: 3422-3423.

Castberg, T., Thyrhaug, R., Larsen, A., Sandaa, R-A., Heldal, M., Van Etten, J. L., and Bratbak, G. (2002) Isolation and characterization of a virus that infects *Emiliania huxleyi* (Haptophyta). *Journal of Phycology* **38**: 767-774.

Castelnovo, M., and Evilevitch, A. (2007) DNA ejection from bacteriophage: towards a general behavior for osmotic-suppression experiments. *European Physical Journal E* **24**: 9-18.

Castenholz, R.W., and Waterbury, J.B. (1989) Oxygenic photosynthetic bacteria. Group I. Cyanobacteria. Preface. In *Bergey's Manual of Systematic Bacteriology*. Staley, J.T., Bryant, M.P., Pfenning, N., and Holt, J.G. (eds). Baltimore: Williams and Wilkins, pp. 1710-1727.

Cembella, A.D., Antia, N.J., Harrison, P.J. (1984) The utilisation of organic and inorganic phosphorus compounds as nutrients by eukaryotic microalgae: a multidisciplinary perspective: part 1. *Critical Review of Microbiology* **10**: 317-391.

Champoux, J.J. (2001) DNA topoisomerases: Structure, function, and mechanism. *Annual Review of Biochemistry* **70**: 369-413.

Charlson, R.J., Lovelock, J.E., Andreae, M.O., and Warren, S.G. (1987) Oceanic phytoplankton, atmospheric sulfur, cloud albedo and climate. *Nature* **326**: 655-661.

Chen, F., and Suttle, C.A. (1995a) Amplification of DNA polymerase gene fragments from viruses infecting microalgae. *Applied and Environmental Microbiology* **61**: 1274-1278.

Chen, F., and Suttle, C.A. (1995b) Nested PCR With 3 highly degenerate primers for amplification and identification of DNA from related organisms. *Biotechniques* **18**: 609-612.

Chen, F., Suttle, C.A., and Short, S.M. (1996) Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. *Applied and Environmental Microbiology* **62**: 2869-2874.

Chevalier, B.S., Stoddard, B.L. (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Research* **29**: 3757-3774.

Chiura, H.X. (1997) Generalized gene transfer by virus-like particles from marine bacteria. *Aquatic Microbial Ecology* **13**: 75-83.

Chretiennotdinet, M.J., Courties, C., Vaquer, A., Neveux, J., Claustre, H., Lautier, J., and Machado, M.C. (1995) A new marine picoeukaryote - *Ostreococcus tauri* gen. et sp. nov. (Chlorophyta, *Prasinophyceae*). *Phycologia* **34**: 285-292.

Chung, C.-C., Hwang, S.-P.L., and Chang, J. (2003) Identification of a high affinity phosphate transporter gene in a Prasinophyte alga, *Tetraselmis chui*, and its expression under nutrient limitation. *Applied Environmental Microbiology* **69**: 754-759.

Cottrell, M.T., and Suttle, C.A. (1991) Wide-spread occurrence and clonal variation in viruses which cause lysis of a cosmopolitan, eukaryotic marine phytoplankter, *Micromonas pusilla. Marine Ecology-Progress Series* **78**: 1-9.

Cottrell, M.T., and Suttle, C.A. (1995) Dynamics of a lytic virus infecting the photosynthetic marine picoflagellate *Micromonas pusilla*. *Limnology and Oceanography* **40**: 730-739.

Countway, P.D., and Caron, D.A. (2006) Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Applied and Environmental Microbiology* **72**: 2496-2506.

Courties, C., Perasso, R., Chretiennot-Dinet, M.J., Gouy, M., Guillou, L., and Troussellier, M. (1998) Phylogenetic analysis and genome size of *Ostreococcus tauri* (Chlorophyta, *Prasinophyceae*). *Journal of Phycology* **34**: 844-849.

Cross, G.A.M. (1987) Eukaryotic protein modification and membrane attachment via phosphatidylinositol. *Cell* **48**: 179-181.

Culley, A.I., Lang, A.S., and Suttle, C.A. (2006) Metagenomic analysis of coastal RNA virus communities. *Science* **312**: 1795-1798.

d'Herelle, F.H. (1917) Sur un microbe invisible antagoniste des bacilles dysenteriques. *CR Academie Science* **165**: 373-375.

Daugbjerg, N., and Moestrup, O. (1993) 4 new species of *Pyramimonas* (*Prasinophyceae*) from Arctic Canada including a light and electron-microscopic description of *Pyramimonas quadrifolia* sp. nov. *European Journal of Phycology* **28**: 3-16.

Davison, M.D., Rixon, F.J., and Davison, A.J. (1992) Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes simplex virus type 1. *Journal of General Virology* **73**: 2709-2713.

Delaroque, N., Muller, D.G., Bothe, G., Pohl, T., Knippers, R., and Boland, W. (2001) The complete DNA sequence of the *Ectocarpus siliculosus* virus EsV-1 genome. *Virology* **287**: 112-132.

Delaroque, N., Boland, W., Muller, D.G., and Knippers, R. (2003) Comparisons of two large phaeoviral genomes and evolutionary implications. *Journal of Molecular Evolution* **57**: 613-622.

Delcher, A.L., Bratke, K.A., Powers, E.C., and Salzberg, S.L. (2007) Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**: 673-679.

Derelle, E., Ferraz, C., Rombauts, S., Rouze, P., Worden, A.Z., Robbens, S. Partensky, F., Degroeve, S., Echeynie, S., Cooke, R., Saeys, Y., Wuyts, J., Jabbari, K., Bowler, C., Panaud, O., Piegu, B., Ball, S., Ral, J.P., Bouget, F.-Y., Piganeau, G., De Baets, B., Picard, A., Delseny, M., Demaille, J., Van de Peer, Y., and Moreau, H.. (2006)

Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 11647-11652.

Derelle, E., Ferraz, C., Escande, M.-L., Eychenie, S., Cooke, R., Piganeau, G., Desdevises, Y., Bellec, L., Moreau, H., and Grimsley, N. (2008) Life-cycle and genome of OtV5, a large DNA virus of the pelagic marine unicellular green alga *Ostreococcus tauri*. *PLoS ONE* **3**: e2250.

Dickey, J.S., and Osheroff, N. (2005) Impact of the C-terminal domain of topoisomerase II alpha on the DNA cleavage activity of the human enzyme. *Biochemistry* **44**: 11546-11554.

Dickey, J.S., Van Etten, J.L., and Osheroff, N. (2005) DNA methylation impacts the cleavage activity of *Chlorella* virus topoisomerase II. *Biochemistry* **44**: 15378-15386.

Diez, B., Pedros-Alio, C., and Massana, R. (2001) Study of genetic diversity of eukaryotic picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Applied and Environmental Microbiology* **67**: 2932-2941.

Dimmock, N.J., Easton, A., and Leppard, K. (2001) *Introduction to Modern Virology.*: Wiley-Blackwell.

DiTullio, G.R., Grebmeier, J.M., Arrigo, K.R., Lizotte, M.P., Robinson, D.H., Leventer, A. Barry, J.P., Van Woert, M.L., and Dunbar, R.B. (2000) Rapid and early export of *Phaeocystis antarctica* blooms in the Ross Sea, Antarctica. *Nature* **404**: 595-598.

Dundas, I., Johannessen, O.M., Berge, G., and Heimdal, B. (1989) Toxic algal bloom in Scandinavian waters, May-June 1988. *Oceanography* **2**: 9-14.

Dunigan, D.D., Fitzgerald, L.A., and Van Etten, J.L. (2006) Phycodnaviruses: A peek at genetic diversity. *Virus Research* **117**: 119-132.

Eikrem, W., and Throndsen, J. (1990) The ultrastructure of *Bathyococcus* gen. nov. and *B. prasinos* sp. nov., a non-motile picoplanktonic alga (Chlorophyta, *Prasinophyceae*) from the Mediterranean and Atlantic. *Phycologia* **29**: 344-350.

Eissler, Y., Sahlsten, E., and Quinones, R.A. (2003) Effects of virus infection on respiration rates of marine phytoplankton and microplankton communities. *Marine Ecology-Progress Series* **262**: 71-80.

Eriksson, M., Myllyharju, J., Tu, H.M., Hellman, M., and Kivirikko, K.I. (1999) Evidence for 4-hydroxyproline in viral proteins - Characterization of a viral prolyl 4-hydroxylase and its peptide substrates. *Journal of Biological Chemistry* **274**: 22131-22134.

Essani, K., Goorha, R., and Granoff, A. (1987) Mutation in a DNA-binding protein reveals an association between DNA-methyltransferase activity and a 26,000-Da polypeptide in Frog Virus-3-infected cells. *Virology* **161**: 211-217.

Estep, K.W., and MacIntyre, F. (1989) Taxonomy, life cycle, distribution and dasmotrophy of *Chrysochromulina*: a theory accounting for scales, haptonema, muciferous bodies and toxicity. *Marine Ecology-Progress Series* **57**: 11-21.

Evans, C., Archer, S.D., Jacquet, S., and Wilson, W.H. (2003) Direct estimates of the contribution of viral lysis and microzooplankton grazing to the decline of a *Micromonas* sp. population. *Aquatic Microbial Ecology* **30**: 207-219.

Evans, C., Malin, G., Mills, G.P., and Wilson, W.H. (2006) Viral infection of *Emiliania huxleyi* (*Prymnesiophyceae*) leads to elevated production of reactive oxygen species. *Journal of Phycology* **42**: 1040-1047.

Evilevitch, A., Lavelle, L., Knobler, C.M., Raspaud, E., and Gelbart, W.M. (2003) Osmotic pressure inhibition of DNA ejection from phage. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 9292-9295.

Evilevitch, A., Castelnovo, M., Knobler, C.M., and Gelbart, W.M. (2004) Measuring the force ejecting DNA from phage. *Journal of Physical Chemistry B* **108**: 6838-6843.

Evilevitch, A., Gober, J.W., Phillips, M., Knobler, C.M., and Gelbart, W.M. (2005) Measurements of DNA lengths remaining in a viral capsid after osmotically suppressed partial ejection. *Biophysical Journal* **88**: 751-756.

Falkowski, P.G., Barber, R.T., and Smetacek, V. (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**: 200-206.

Felsenstein, J. (2005) PHYLIP (Phylogeny Inference Package) version 3.6. In, pp. *Distributed by the author. Department of Genome Sciences. University of Washington, Seattle*.

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

Figueroa, R.I., and Rengefors, K. (2006) Life cyle and sexuality of the freshwater raphidophyte *Gonyostomum semen* (Raphidophyceae). *Journal of Phycology* **42**: 859-871.

Filee, J., Siguier, P., and Chandler, M. (2007) I am what I eat and I eat what I am: acquisition of bacterial genes by giant viruses. *Trends in Genetics* **23**: 10-15.

Filee, J., Pouget, N., Chandler, M. (2008) Phylogenetic evidence for extensive lateral acquisition of cellular genes by Nucleocytoplasmic large DNA viruses. *BMC Evolutionary Biology* **8**: 320.

Fischer, D., and Eisenberg, D. (1999) Finding families for genomic ORFans. *Bioinformatics* **15**: 759-762.

Fischer, U.R., Wieltschnig, C., Kirschner, A.K.T., and Velimirov, B. (2006) Contribution of virus-induced lysis and protozoan grazing to benthic bacterial mortality estimated simultaneously in microcosms. *Environmental Microbiology* **8**: 1394-1407.

Fitzgerald, L.A., Graves, M.V., Li, X., Feldblyum, T., Hartigan, J., and Van Etten, J.L. (2007a) Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi. *Virology* **358**: 459-471.

Fitzgerald, L.A., Graves, M.V., Li, X., Feldblyum, T., Nierman, W.C., and Van Etten, J.L. (2007b) Sequence and annotation of the 369-kb NY-2A and the 345-kb AR158 viruses that infect *Chlorella* NC64A. *Virology* **358**: 472-484.

Fitzgerald, L.A., Graves, M.V., Li, X., Hartigan, J., Pfitzner, A.J.P., Hoffart, E., and Van Etten, J.L. (2007c) "Sequence and annotation of the 288-kb ATCV-1 virus that infects an endosymbiotic *Chlorella* strain of the heliozoon *Acanthocystis turfacea*" (vol 362, pg 350, 2007). *Virology* **366**: 226-226.

Flint, S.J., Enquist, L.W., R.M., K., Racaniello, V.R., and Skalka, A.M. (2004) *Principles of virology - molecular biology, pathogenesis and control of animal viruses.* Washington D.C.: ASM Press.

Fortune, J.M., Lavrukhin, O.V., Gurnon, J.R., Van Etten, J.L., Lloyd, R.S., and Osheroff, N. (2001) Topoisomerase II from *Chlorella* virus PBCV-1 has an exceptionally high DNA cleavage activity. *Journal of Biological Chemistry* **276**: 24401-24408.

Fraenkel-Conrat, H., Singer, B., and Williams, R.C. (1957) Infectivity of viral nucleic acid. *Biochimica Et Biophysica Acta* **25**: 87.

Frohns, F., Kasmann, A., Kramer, D., Schafer, B., Mehmel, M., Kang, M., Van Etten, J.L., Gazzarrini, S., Moroni, A., and Thiel, G. (2006) Potassium ion channels of *Chlorella* viruses cause rapid depolarization of host cells during infection. *Journal of Virology* **80**: 2437-2444.

Fuhrman, J.A. (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541-548.

Fuller, N.J., Wilson, W.H., Joint, I.R., and Mann, N.H. (1998) Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. *Applied and Environmental Microbiology* **64**: 2051-2060.

Fuller, N.J., Tarran, G.A., Cummings, D.G., Woodward, E.M.S., Orcutt, K.M., Yallop, M., Le Gall, F., and Scanlan, D.J. (2006a) Molecular analysis of photosynthetic picoeukaryote community structure along an Arabian Sea transect. *Limnology and Oceanography* **51**: 2502-2514.

Fuller, N.J., Campbell, C., Allen, D.J., Pitt, F.D., Zwirglmaierl, K., Le Gall, F., Vaulot, D., and Scanlan, D.J. (2006b) Analysis of photosynthetic picoeukaryote diversity at open ocean sites in the Arabian Sea using a PCR biased towards marine algal plastids. *Aquatic Microbial Ecology* **43**: 79-93.

Garry, R.T., Hearing, P., and Cosper, E.M. (1998) Characterization of a lytic virus infectious to the bloom-forming microalga *Aureococcus anophagefferens* (Pelagophyceae). *Journal of Phycology* **34**: 616-621.

Gasol, J.M., Pedros-Alio, C., and Vaque, D. (2002) Regulation of bacterial assemblages in oligotrophic plankton systems: results from experimental and empirical approaches. *Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology* **81**: 435-452.

Gastrich, M.D., Anderson, O.R., Benmayor, S.S., and Cosper, E.M. (1998) Ultrastructural analysis of viral infection in the brown-tide alga, *Aureococcus anophagefferens* (Pelagophyceae). *Phycologia* **37**: 300-306.

Gastrich, M.D., Anderson, O.R., and Cosper, E.M. (2002) Viral-like particles (VLPS) in the alga, *Aureococcus anophagefferens* (Pelagophyceae), during 1999-2000 brown tide blooms in Little Egg Harbor, New Jersey. *Estuaries* **25**: 938-943.

Gastrich, M.D., Leigh-Bell, J.A., Gobler, C.J., Anderson, O.R., Wilhelm, S.W., and Bryan, M. (2004) Viruses as potential regulators of regional brown tide blooms caused by the alga, *Aureococcus anophagefferens*. *Estuaries* **27**: 112-119.

Georgiou, T., Yu, Y.T.N., Ekunwe, S., Buttner, M.J., Zuurmond, A.M., Kraal, B.., Kleanthous, C., and Snyder, L. (1998) Specific peptide-activated proteolytic cleavage of *Escherichia coli* elongation factor Tu. *Proceedings of the National Academy of Sciences of the United States of America* **95**: 2891-2895.

Gimble, F.S. (2001) Degeneration of a homing endonuclease and its target sequence in a wild yeast strain. In, pp. 4215-4223.

Gobler, C.J., Hutchins, D.A., Fisher, N.S., Cosper, E.M., and Sanudo-Wilhelmy, S.A. (1997) Release and bioavailability of C, N, P, Se, and Fe following viral lysis of a marine chrysophyte. *Limnology and Oceanography* **42**: 1492-1504.

Gobler, C.J., Anderson, O.R., Gastrich, M.D., and Wilhelm, S.W. (2007) Ecological aspects of viral infection and lysis in the harmful brown tide alga *Aureococcus anophagefferens*. *Aquatic Microbial Ecology* **47**: 25-36.

Gogarten, J.P., Senejani, A.G., Zhaxybayeva, O., Olendzenski, L., and Hilario, E. (2002) Inteins: Structure, Function, and Evolution. In, pp. 263-287.

Graham, L.K.E., and Wilcox, L.W. (2000) The origin of alternation of generations in land plants: a focus on matrotrophy and hexose transport. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **355**: 757-766.

Grand, R.J.A. (1989) Acylation of viral and eukaryotic proteins. *Biochemistry Journal* **258**: 625-638.

Graziani, S., Xia, Y.N., Gurnon, J.R., Van Etten, J.L., Leduc, D., Skouloubris, S. Myllykallio, H., and Liebl., U. (2004) Functional analysis of FAD-dependent

thymidylate synthase ThyX from *Paramecium bursaria Chlorella* virus-1. *Journal of Biological Chemistry* **279**: 54340-54347.

Graziani, S., Bernauer, J., Skouloubris, S., Graille, M., Zhou, C.Z., Marchand, C., Decottignies, P., van Tilbeurgh, H., Myllykallio, H., and Liebl, U. (2006) Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. *Journal of Biological Chemistry* **281**: 24048-24057.

Guillard, R.R.L. (1975) Culture of phytoplankton for feeding marine invertebrates. In *Culture of marine invertebrate animals*. Smith, W.L., and Chanley, M.H. (eds). New York: Plenum Press, pp. 29-60.

Guillou, L., Eikrem, W., Chretiennot-Dinet, M.J., Le Gall, F., Massana, R., Romari, K., Pedros-Alio, C., and Vaulot, D. (2004) Diversity of picoplanktonic prasinophytes assessed by direct nuclear SSU rDNA sequencing of environmental samples and novel isolates retrieved from oceanic and coastal marine ecosystems. *Protist* **155**: 193-214.

Haaber, J., and Middelboe, M. (2009) Viral lysis of *Phaeocystis pouchetii*: Implications for algal population dynamics and heterotrophic C, N and P cycling. *Isme Journal* **3**: 430-441.

Hajdu, S., Larsson, U., and Moestrup, O. (1996) Seasonal dynamics of *Chrysochromulina* species (*Prymnesiophyceae*) in a coastal area and a nutrient-enriched inlet of the northern Baltic proper. *Botanica Marina* **39**: 281-295.

Han, G., Gable, K., Yan, L., Allen, M.J., Wilson, W.H., Moitra, P., Harmon, J.M., and Dunn, T.M. (2006) Expression of a novel marine viral single-chain serine palmitoyltransferase and construction of yeast and mammalian single-chain chimera. *J Biol Chem* **281**: 39935-39942.

Hawkins, A.R., and Smith, M. (1991) Domain structure and interaction within the pentafunctional arom polypeptide. *European Journal of Biochemistry* **196**: 717-724.

Hawser, S.P., O'Neil, J.M., Roman, M.R., and Codd, G.A. (1992) Toxicity of blooms of the cyanobacterium *Trichodesmium* to zooplankton. *Journal of Applied Phycology* **4**: 79-86.

Hedges, J.I. (1992) Global Biogeochemical Cycles - Progress and Problems. *Marine Chemistry* **39**: 67-93.

Henderson, G.P., Gan, L., and Jensen, G.J. (2007) 3-D ultrastructure of *O. tauri:* electron cryotomography of an entire eukaryotic cell. *PLoS ONE* **2**: e749.

Hendrix, R.W., Lawrence, J.G., Hatfull, G.F., and Casjens, S. (2000) The origins and ongoing evolution of viruses. *Trends in Microbiology* **8**: 504-508.

Hennes, K.P., and Suttle, C.A. (1995) Direct counts of viruses in natural waters and laboratory cultures by epifluorescence microscopy. *Limnology and Oceanography* **40**: 1050-1055.

Hennes, K.P., Suttle, C.A., and Chan, A.M. (1995) Fluorescently labeled virus probes show that natural virus populations can control the structure of marine microbial communities. *Applied and Environmental Microbiology* **61**: 3623-3627.

Henry, E.C., and Meints, R.H. (1992) A persistent virus infection in *Feldmannia* (Phaeophyceae). *Journal of Phycology* **28**: 517-526.

Hewat, E.A., Neumann, E., and Blaas, D. (2002) The concerted conformational changes during human rhinovirus 2 uncoating. *Molecular Cell* **10**: 317-326.

Hewson, I., Govil, S.R., Capone, D.G., Carpenter, E.J., and Fuhrman, J.A. (2004) Evidence of *Trichodesmium* viral lysis and potential significance for biogeochemical cycling in the oligotrophic ocean. *Aquatic Microbial Ecology* **36**: 1-8.

Hirabayashi, J. (2004) On the origin of glycome and saccharide recognition. *Trends in Glycoscience and Glycotechnology* **16**: 63-85.

Hiroishi, S., Okada, H., Imai, I., and Yoshida, T. (2005) High toxicity of the novel bloom-forming species *Chattonella ovata* (Raphidophyceae) to cultured fish. *Harmful Algae* **4**: 783-787.

Ho, C.K., Van Etten, J.L., and Shuman, S. (1997) Characterization of an ATP-dependent DNA ligase encoded by *Chlorella* virus PBCV-1. *Journal of Virology* **71**: 1931-1937.

Hobbie, J.E., Daley, R.J., and Jasper, S. (1977) Use of nucleopore filters for counting bacteria by fluorescence microscopy. *Applied and Environmental Microbiology* **33**: 1225-1228.

Holligan, P.M., and Harbour, D.S. (1977) The vertical distribution and succession of phytoplankton in the western English Channel in 1975 and 1976. *Journal of the Marine Biological Association of the United Kingdom* **57**: 1075-1093.

Holligan, P.M., Fernandez, E., Aiken, J., Balch, W.M., Boyd, P., Burkill, P.H.Finch, M., Groom, S.B., Malin, G., Muller, K., Purdie, D.A., Robinson, C., Trees, C.C., Turner, S.M., and van der Wal, P. (1993) A biogeochemical study of the coccolithophore, *Emiliania huxleyi,* in the North Atlantic. *Global Biogeochemical Cycles* **7**: 879-900.

Honjo, T. (1993) Overview on bloom dynamics and physiological ecology of *Heterosigma akashiwo*. In *Toxic phytoplankton blooms in the sea*. Smayda, T.J., and Shimizu, Y. (eds). New York: Elsevier, pp. 33-41.

Horiguchi, T. (1997) *Heterocapsa arctica* sp nov (Peridiniales, Dinophyceae), a new marine dinoflagellate from the arctic. *Phycologia* **36**: 488-491.

Huang, Y., and Zhang, L. (2004) Rapid and sensitive dot-matrix methods for genome analysis. *Bioinformatics* **20**: 460-466.

Ivanovsky, D. (1892) Concerning the mosaic disease of tobacco. *Z Pflanzenkr Pflanzenpathol Pflanzenschutz* **13**: 1-41.

Iyer, L.M., Aravind, L., and Koonin, E.V. (2001) Common origin of four diverse families of large eukaryotic DNA viruses. *Journal of Virology* **75**: 11720-11734.

Iyer, L.A., Balaji, S., Koonin, E.V., and Aravind, L. (2006) Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Research* **117**: 156-184.

Jacobsen, A., Bratbak, G., and Heldal, M. (1996) Isolation and characterization of a virus infecting *Phaeocystis pouchetii* (*Prymnesiophyceae*). *Journal of Phycology* **32**: 923-927.

Jacquet, S., Heldal, M., Iglesias-Rodriguez, D., Larsen, A., Wilson, W., and Bratbak, G. (2002) Flow cytometric analysis of an *Emiliana huxleyi* bloom terminated by viral infection. *Aquatic Microbial Ecology* **27**: 111-124.

Jacquet, S., Domaizon, I., Personnic, S., and Sime-Ngando, T. (2007) Do small grazers influence virus-induced mortality of bacteria in Lake Bourget (France)? *Fundamental and Applied Limnology* **170**: 125-132.

Jiang, S.C., and Paul, J.H. (1998) Gene transfer by transduction in the marine environment. *Applied and Environmental Microbiology* **64**: 2780-2787.

Joint, I.R., Owen, N.J.P., and Pomroy, A.J. (1986) Seasonal production of photosynthetic picoplankton and nanoplankton in the Celtic sea. *Marine Ecology-Progress Series* **28**: 251-258.

Jones, A.M., and Dangl, J.L. (1996) Logjam at the Styx: Programmed cell death in plants. *Trends in Plant Science* **1**: 114-119.

Karata, K., Inagawa, T., Wilkinson, A.J., Tatsuta, T., and Ogura, T. (1999) Dissecting the role of a conserved motif (the second region of homology) in the AAA family of ATPases - Site-directed mutagenesis of the ATP-dependent protease FtsH. *Journal of Biological Chemistry* **274**: 26225-26232.

Kausche, G.A., and Ruska, H. (1939) Die Struktur dur "kristallinen Aggregate" des Tabakmosaikvirus-proteins. *Biochem Z* **303**: 211.

Kawakami, H., and Kawakami, N. (1978) Behaviour of a virus in a symbiotic system, *Paramecium bursaria ZooChlorella*. *Journal of Protozoology* **25**: 217-225.

Kawasaki, T., Tanaka, M., Fujie, M., Usami, S., and Yamada, T. (2004) Immediate early genes expressed in chlorovirus infections. *Virology* **318**: 214-223.

Keller, M.D., Selvin, R.C., Claus, W., and Guillard, R.R.L. (1987) Media for the culture of marine ultraphytoplankton. *Journal of Phycology* **23**: 633-638.

Khan, S., Arakawa, O., and Onoue, Y. (1997) Neurotoxins in a toxic red tide of *Heterosigma akashiwo* (Raphidophyceae) in Kagoshima Bay, Japan. *Aquaculture Research* **28**: 9-14.

Kim, S.K., Makino, K., Amemura, M., Shinagawa, H., and Nakata, A. (1993) Molecular analysis of the phoH gene, belonging to the phosphate regulon in *Escherichia coli*. *Journal of Bacteriology* **175**: 1316-1324.

Kim, Y., and Clark, D.J. (2002) SWI/SNF-dependent long-range remodeling of yeast HIS3 chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 15381-15386.

Kindt, J., Tzlil, S., Ben-Shaul, A., and Gelbart, W.M. (2001) DNA packaging and ejection forces in bacteriophage. *Proceedings National Academy of Sciences USA* **98**: 13671-13674.

Kivirikko, K.I., Myllyla, R., and Pihlajaniemi, T. (1992) Hydroxylation of proline and lysine residues in collagens and other animal and plant proteins. . In *Focus on post-translational modifications of proteins*. Harding, J.J., and Crabbe, M.J.C. (eds). Boca Raton, FL: CRC Press, pp. 1-51.

Knipe, D.M. (1996) Virus-host cell interactions. In *Fields Virology*. Fields, B.N., Knipe, D.M., Howley, P.M., Chanock, R.M., Melnick, J.L. (ed). Philadelphia: Lipincott-Raven.

Kurtz, S., Phillippy, A., Delcher, A.L., Smoot, M., Shumway, M., Antonescu, C., and Salzberg, S.L. (2004) Versatile and open software for comparing large genomes. *Genome Biology* **5**: 9.

La Scola, B., Audic, S., Robert, C., Jungang, L., de Lamballerie, X., Drancourt, M., Birtles, R., Claverie, J.M., and Raoult, D. (2003) A giant virus in amoebae. *Science* **299**: 2033-2033.

Lancelot, C., Billen, G., Sournia, A., Weisse, T., Colijn, F., Veldhuis, M.J.W., Davies, A., and Wassman, P. (1987) *Phaeocystis* blooms and nutrient enrichment in the continental coastal zones of the North Sea. *Ambio* **16**: 38-46.

Landry, M.R., Kirshtein, J., and Constantinou, J. (1996) Abundances and distributions of picoplankton populations in the central equatorial Pacific from 12 degrees N to 12 degrees S, 140 degrees W. *Deep-Sea Research Part Ii-Topical Studies in Oceanography* **43**: 871-890.

Lang, A.S., Culley, A.I., and Suttle, C.A. (2004) Genome sequence and characterization of a virus (HaRNAV) related to picoma-like viruses that infects the marine toxic bloom-forming alga *Heterosigma akashiwo*. *Virology* **320**: 206-217.

Lanka, S.T.J., Klein, M., Ramsperger, U., Muller, D.G., and Knippers, R. (1993) Genome structure of a virus infecting the marine brown alga *Ectocarpus siliculosus*. *Virology* **193**: 802-811.

Larsen, A., Castberg, T., Sandaa, R.A., Brussaard, C.P.D., Egge, J., Heldal, M., Paulino, A. Thyrhaug, R., van Hannen, E.J. and Bratbak, G. (2001) Population dynamics and diversity of phytoplankton, bacteria and viruses in a seawater enclosure. *Marine Ecology-Progress Series* **221**: 47-57.

Larsen, J.B., Larsen, A., Bratbak, G., and Sandaa, R.A. (2008) Phylogenetic analysis of members of the *Phycodnaviridae* virus family, using amplified fragments of the major capsid protein gene. *Applied and Environmental Microbiology* **74**: 3048-3057.

Lavrukhin, O.V., Fortune, J.M., Wood, T.G., Burbank, D.E., Van Etten, J.L., Osheroff, N., and Lloyd, R.S. (2000) Topoisomerase II from *Chlorella* virus PBCV-1 - Characterization of the smallest known type II topoisomerase. *Journal of Biological Chemistry* **275**: 6915-6921.

Lawrence, J.E., Chan, A.M., and Suttle, C.A. (2001) A novel virus (HaNIV) causes lysis of the toxic bloom-forming alga *Heterosigma akashiwo* (Raphidophyceae). *Journal of Phycology* **37**: 216-222.

Lawrence, J.E., Chan, A.M., and Suttle, C.A. (2002) Viruses causing lysis of the toxic bloom-forming alga *Heterosigma akashiwo* (Raphidophyceae) are widespread in coastal sediments of British Columbia, Canada. *Limnology and Oceanography* **47**: 545-550.

Lawrence, J.E., Brussaard, C.P.D., and Suttle, C.A. (2006) Virus-specific responses of *Heterosigma akashiwo* to infection. *Applied and Environmental Microbiology* **72**: 7829-7834.

Lee, A.M., Ivey, R.G., Henry, E.C., and Meints, R.H. (1995) Characterization of a repetitive DNA element in a brown algal virus. *Virology* **212**: 474-480.

Lewis, L.A., and McCourt, R.M. (2004) Green algae and the origin of land plants. *American Journal of Botany* **91:** 1535-1556.

Li, W.K.W. (1994) Primary production of prochlorophytes, cyanobacteria, and eukaryotic ultraphytoplankton - Measurements from flow cytometric sorting. *Limnology and Oceanography* **39**: 169-175.

Li, Y., Lu, Z.Q., Sun, L.W., Ropp, S., Kutish, G.F., Rock, D.L., and Van Etten, J.L. (1997) Analysis of 74 kb of DNA located at the right end of the 330-kb *Chlorella* virus PBCV-1 genome. *Virology* **237**: 360-377.

Li, L.Y., Huang, Q.J., Wu, S.H., Lin, D., Chen, J.H., and Chen, Y.Q. (2008) The spatial and temporal distribution of microalgae in the South China Sea: evidence from GIS-based analysis of 18S rDNA sequences. *Science in China Series C-Life Sciences* **51**: 1121-1128.

Lindell, D., Sullivan, M.B., Johnson, Z.I., Tolonen, A.C., Rohwer, F., and Chisholm, S.W. (2004) Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 11013-11018.

Liss, P.S., Malin, G., Turner, S.M., and Holligan, P.M. (1994) Dimethyl sulfide and *Phaeocystis* - a review. *Journal of Marine Systems* **5**: 41-53.

Liu, X.-Q., Yang, J., and Meng, Q. (2003) Four inteins and three group II introns encoded in a bacterial ribonucleotide reductase gene. *Journal of Biological Chemistry* **278:** 46826-46831.

Liu, X.Q., and Yang, J. (2004) Bacterial thymidylate synthase with intein, group II intron and distinctive ThyX motifs. *Journal of Bacteriology* **186**: 6316-6319.

Lobley, C.M.C., Schmitzberger, F., Kilkenny, M.L., Whitney, H., Ottenhof, H.H., Chakauya, E., Webb, M.E., Birch, L.M., Tuck, K.L., Abell, C., Smith, A.G., and Blundell, T.L. (2003) Structural insights into the evolution of the pantothenate-biosynthesis pathway. *Biochemical Society Transactions* **31**: 563-571.

Loeffler, F., and Frosch, P. (1898) Report of the commission for research on foot-and-mouth disease. *Zentralbl Bakteriologie Parasitenkunde Infektionskrankh Hyg Abt 1 Orig* **23**: 371-391.

Lopez-Garcia, P., Moreira, D., and Rodriguez-Valera, F. (2001a) Diversity of free-living prokaryotes from a deep-sea site at the Antarctic Polar Front. *FEMS Microbiology Ecology* **36**: 193-202.

Lopez-Garcia, P., Rodriguez-Valera, F., Pedros-Alio, C., and Moreira, D. (2001b) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature* **409**: 603-607.

Lovering, A.L., de Castro, L.H., Lim, D., and Strynadka, N.C.J. (2007) Structural insight into the transglycosylation step of bacterial cell wall biosynthesis. *Science* **315**: 1402-1405.

Lowe, T.M.E., S.R. (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Research* **25**: 955-964.

Luther, K.B., and Haltiwanger, R.S. (2009) Role of unusual O-glycans in intercellular signaling. *International Journal of Biochemistry & Cell Biology* **41**: 1011-1024.

Mackinder, L.C.M., Worthy, C.A., Biggi, G., Hall, M., and Schroeder, D.C. (in press) A unicellular algal virus, *Emiliania huxleyi* virus 86, exploits an animal-like infection strategy. *Journal of General Virology.*

Maier, I., Wolf, S., Delaroque, N., Muller, D.G., and Kawai, H. (1998) A DNA virus infecting the marine brown alga *Pilayella littoralis* (*Ectocarpales*, Phaeophyceae) in culture. *European Journal of Phycology* **33**: 213-220.

Maley, F., and Maley, G.F. (1999) Structure, function analysis of T4-phage dexoycytidylate deaminase and its role in the phage metabolic pathway. *Paths Pyrimidines* **7**: 1-7.

Malin, G., and Kirst, G.O. (1997) Algal production of dimethyl sulfide and its atmospheric role. *Journal of Phycology* **33**: 889-896.

Malin, G., Wilson, W.H., Bratbak, G., Liss, P.S., and Mann, N.H. (1998) Elevated production of dimethylsulfide resulting from viral infection of cultures of *Phaeocystis pouchetii. Limnology and Oceanography* **43**: 1389-1393.

Mann, N.H., Cook, A., Millard, A., Bailey, S., and Clokie, M. (2003) Marine ecosystems: Bacterial photosynthesis genes in a virus. *Nature* **424**: 741-741.

Manton, I., and Leadbeater, B.S.C. (1974) Fine structural observations on six species of *Chrysochromulina* from wild Danish marine nanoplankton, including a description of *C. campanulifera* sp. nov. and a preliminary summary of the nanoplankton as a whole. *Det Kongelige Danske Videnskabernes Selskab Biologiske Skrifter* **20**: 1- 26.

Markine-Goriaynoff, N., Gillet, L., Van Etten, J.L., Korres, H., Verma, N., and Vanderplasschen, A. (2004) Glycosyltransferases encoded by viruses. *Journal of General Virology* **85**: 2741-2754.

Martin, E.L., and Benson, R. (1988) Phages of cyanobacteria. In *The Bacteriophages*. Calendar, R. (ed). New York: Plenum Press, pp. 607-645.

Massana, R., Guillou, L., Diez, B., and Pedros-Alio, C. (2002) Unveiling the organisms behind novel eukaryotic ribosomal DNA sequences from the ocean. *Applied and Environmental Microbiology* **68**: 4554-4558.

Massana, R., Balague, V., Guillou, L., and Pedros-Alio, C. (2004) Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiology Ecology* **50**: 231-243.

Massana, R., and Pedros-Alio, C. (2008) Unveiling new microbial eukaryotes in the surface ocean. *Current Opinion in Microbiology* **11**: 213-218.

Matthew, R.G., Duncan, G., and Austin, B. (2001) Outcrossed sex allows a selfish gene to invade yeast populations. *Proceedings of the Royal Society B: Biological Sciences* **268**: 2537-2542.

Mayer, J.A., and Taylor, F.J.R. (1979) A virus which lyses the marine nanoflagellate *Micromonas pusilla*. *Nature* **281**: 299-301.

Mehmel, M., Rothermel, M., Meckel, T., Van Etten, J.L., Moroni, A., and Thiel, G. (2003) Possible function for virus encoded K+ channel Kcv in the replication of *Chlorella* virus PBCV-1. *Febs Letters* **552**: 7-11.

Meints, R.H., Van Etten, J.L., Kuczmarski, D., Lee, K., and Ang, B. (1981) Viral infection of the symbiotic *Chlorella*-like alga present in *Hydra viridis*. *Virology* **113**: 698-703.

Meints, R.H., Lee, K., Burbank, D.E., and Van Etten, J.L. (1984) Infection of a *Chlorella*-Like Alga With the Virus, PBCV-1 - Ultrastructural Studies. *Virology* **138**: 341-346.

Meints, R.H., Lee, K., and Van Etten, J.L. (1986) Assembly Site of the Virus PBCV-1 in a *Chlorella*-Like Green- Alga - Ultrastructural Studies. *Virology* **154**: 240-245.

Meints, R.H., Burbank, D.E., Van Etten, J.L., and Lamport, D.T.A. (1988) Properties of the *Chlorella* Receptor For the Virus PBCV-1. *Virology* **164**: 15-21.

Middelboe, M., and Lyck, P.G. (2002) Regeneration of dissolved organic matter by viral lysis in marine microbial communities. *Aquatic Microbial Ecology* **27**: 187-194.

Middelboe, M., Jorgensen, N.O.G., and Kroer, N. (1996) Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. *Applied and Environmental Microbiology* **62**: 1991-1997.

Miki, T., and Jacquet, S. (2008) Complex interactions in the microbial world: underexplored key links between viruses, bacteria and protozoan grazers in aquatic environments. *Aquatic Microbial Ecology* **51**: 195-208.

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 of the United States of America* **101**: 11007-11012.

Miller, R.V., and Ripp, S. (1998) The importance of pseudolysogeny to *in situ* bacteriophage-host interactions. In *Horizontal Gene Transfer*. Syvanen, M., and Kado, C.C. (eds). San Diego, CA: Academic Press, pp. 81-91.

Miller, R.V., and Ripp, S. (2002) Pseudolysogeny: a bacteriophage strategy for increasing longevity *in situ.* In *Genetic interaction among microorganisms in the natural environment*. Wellington, E.M., and Van Elsa, S.J.D. (eds). Oxford: Pergamon Press, pp. 176-193.

Milligan, K.L.D., and Cosper, E.M. (1994) Isolation of virus capable of lysing the brown tide microalga, *Aureococcus anophagefferens*. *Science* **266**: 805-807.

Mizumoto, H., Tomaru, Y., Takao, Y., Shirai, Y., and Nagasaki, K. (2008) Diverse responses of the bivalve-kiving dinoflagellate *Heterocapsa circularisquama* to infection by a single-stranded RNA virus. *Applied and Environmental Microbiology* **74**: 3105-3111.

Moon-van der Staay, S., DeWachter, R., and Vaulot, D. (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. *Nature* **409**: 607-610.

Moore, L.R., Rocap, G., and Chisholm, S.W. (1998) Physiology and molecular phylogeny of coexisiting *Prochlorococcus* ecotypes. *Nature* **393**: 464-467.

Moore, L.R., and Chisholm, S.W. (1999) Photophysiology of the marine cyanobacterium *Prochlorococcus*: Ecotypic differences among cultured isolates. *Limnology and Oceanography* **44**: 628-638.

Moreira, D., and López-García, P. (2009) Ten reasons to exclude viruses from the tree of life. *Nature Reviews Microbiology* **7**: 306-311.

Mousseau, L., Fortier, L., and Legendre, L. (1998) Annual production of fish larvae and their prey in relation to size-fractionated primary production (Scotian Shelf, NW Atlantic). *Ices Journal of Marine Science* **55**: 44-57.

Mühling, M., Fuller, N.J., Millard, A., Somerfield, P.J., Marie, D., Wilson, W.H., Scanlan, D.J., Post, A.F., Joint, I. and Mann, N.H. (2005) Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. *Environmental Microbiology* **7**: 499-508.

Müller, D.G., and Stache, B. (1992) Worldwide occurrence of virus infections in filamentous marine brown algae. *Helgolander Meeresuntersuchungen* **46**: 1-8.

Müller, D.G., Kapp, M., and Knippers, R. (1998) Viruses in marine brown algae. In *Advances in Virus Research, Vol 50*. San Diego: Academic Press Inc, pp. 49-67.

Müller, D.G., Westermeier, R., Morales, J., Reina, G.G., del Campo, E., Correa, J.A., and Rometsch, E. (2000) Massive prevalence of viral DNA in *Ectocarpus* (Phaeophyceae, *Ectocarpales*) from two habitats in the North Atlantic and South Pacific. *Botanica Marina* **43**: 157-159.

Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297**: 105-107.

Nagasaki, K., and Yamaguchi, M. (1997) Isolation of a virus infectious to the harmful bloom causing microalga *Heterosigma akashiwo* (Raphidophyceae). *Aquatic Microbial Ecology* **13**: 135-140.

Nagasaki, K., and Yamaguchi, M. (1998) Intra-species host specificity of HaV (*Heterosigma akashiwo* virus) clones. *Aquatic Microbial Ecology* **14**: 109-112.

Nagasaki, K., Tarutani, K., and Yamaguchi, M. (1999) Cluster analysis on algicidal activity of HaV clones and virus sensitivity of *Heterosigma akashiwo* (Raphidophyceae). *Journal of Plankton Research* **21**: 2219-2226.

Nagasaki, K., Tomaru, Y., Katanozaka, N., Shirai, Y., Nishida, K., Itakura, S., and Yamaguchi, M. (2004) Isolation and characterization of a novel single-stranded RNA

virus infecting the bloom-forming diatom *Rhizosolenia setigera*. *Applied and Environmental Microbiology* **70**: 704-711.

Nagasaki, K., Shirai, Y., Tomaru, Y., Nishida, K., and Pietrokovski, S. (2005a) Algal viruses with distinct intraspecies host specificities include identical intein elements. *Applied Environmental Microbiology* **71**: 3599-3607.

Nagasaki, K., Shirai, Y., Takao, Y., Mizumoto, H., Nishida, K., and Tomaru, Y. (2005b) Comparison of genome sequences of single-stranded RNA viruses infecting the bivalve-killing dinoflagellate *Heterocapsa circularisquama*. *Applied and Environmental Microbiology* **71**: 8888-8894.

Nagasaki, K., Tomaru, Y., Shirai, Y., Takao, Y., and Mizumoto, H. (2006) Dinoflagellate-infecting viruses. *Journal of the Marine Biological Association of the United Kingdom* **86**: 469-474.

Nandhagopal, N., Simpson, A.A., Gurnon, J.R., Yan, X.D., Baker, T.S., Graves, M.V., Van Etten, J.L., and Rossman, M. (2002) The structure and evolution of the major capsid protein of a large, lipid-containing DNA virus. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 14758-14763.

Nelson, M., Burbank, D.E., and Van Etten, J.L. (1998) *Chlorella* viruses encode multiple DNA methyltransferases. *Biological Chemistry* **379**: 423-428.

Niehaus, F., Frey, B., and Antranikian, G. (1997) Cloning and characterisation of a thermostable alpha-DNA polymerase from the hyperthermophilic archaeon *Thermococcus* sp. TY. *Gene* **204**: 153-158.

Nishida, K., Kawasaki, T., Fujie, M., Usami, S., and Yamada, T. (1999) Aminoacylation of tRNAs encoded by *Chlorella* virus CVK2. *Virology* **263**: 220-229.

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.

Noble, R.T., and Fuhrman, J.A. (1999) Breakdown and microbial uptake of marine viruses and other lysis products. *Aquatic Microbial Ecology* **20**: 1-11.

Not, F., Simon, N., Biegala, I.C., and Vaulot, D. (2002) Application of fluorescent in situ hybridization coupled with tyramide signal amplification (FISH-TSA) to assess eukaryotic picoplankton composition. *Aquatic Microbial Ecology* **28**: 157-166.

Not, F., Latasa, M., Marie, D., Cariou, T., Vaulot, D., and Simon, N. (2004) A single species, *Micromonas pusilla* (*Prasinophyceae*), dominates the eukaryotic picoplankton in the western English channel. *Applied and Environmental Microbiology* **70**: 4064-4072.

Not, F., Massana, R., Latasa, M., Marie, D., Colson, C., Eikrem, W., Pedros-Alio, C., Vaulot, D., and Simon, N. (2005) Late summer community composition and abundance of photosynthetic picoeukaryotes in Norwegian and Barents Seas. *Limnology and Oceanography* **50**: 1677-1686.

O'Kelly, C.J., Sieracki, M.E., Thier, E.C., and Hobson, I.C. (2003) A transient bloom of *Ostreococcus* (Chlorophyta, *Prasinophyceae*) in the West Neck Bay, Long Island, New York. *Journal of Phycology* **39**: 850-854.

Ogata, H., Abergel, C., Raoult, D., and Claverie, J.M. (2005) Comment on "The 1.2-megabase genome sequence of Mimivirus". *Science* **308**.

Onimatsu, H., Suganuma, K., Uenoyama, S., and Yamada, T. (2006) C-terminal repetitive motifs in Vp130 present at the unique vertex of the Chlorovirus capsid are essential for binding to the host *Chlorella* cell wall. *Virology* **353**: 433-442.

Onji, M., Nakano, S., and Suzuki, S. (2003) Virus-like particles suppress growth of the red-tide-forming marine dinoflagellate *Gymnodinium mikimotoi*. *Marine Biotechnology* **5**: 435-442.

Ottenhof, H.H., Ashurst, J.L., Whitney, H.M., Saldanha, S.A., Schmitzberger, F., Gweon, H.S., Gweon, H.S., Blundell, T.L., Abell, C., and Smith, A.G. (2004)

Organisation of the pantothenate (vitamin B-5) biosynthesis pathway in higher plants. *Plant Journal* **37**: 61-72.

Ozols, J. (1989) Structure of cytochrome b_5 and its topology in the microsomal membrane. *Biochimica Et Biophysica Acta-Gene Structure and Expression* **997**: 121-130.

Pace, M.L., and Cole, J.J. (1996) Regulation of bacteria by resources and predation tested in whole- lake experiments. *Limnology and Oceanography* **41**: 1448-1460.

Page, R.D.M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**: 357-358.

Paget, M., and Helmann, J. (2003) The sigma 70 family of sigma factors. *Genome Biology* **4**: 203.

Palenik, B., Grimwood, J., Aerts, A., Rouze, P., Salamov, A., Putnam, N. *et al.,* (2007) The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proceedings of the National Academy of Sciences of the United States of America* **104**: 7705-7710.

Parodi, E.R., and Muller, D.G. (1994) Field and culture studies on virus infections in *Hincksia hincksiae* and *Ectocarpus fasciculatus* (*Ectocarpales*, Phaeophyceae). *European Journal of Phycology* **29**: 113-117.

Partensky, F., Blanchot, J., and Vaulot, D. (1999) Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: a review. In *Marine Cyanobacteria*. Charpy, L., and Larkum, A.W.D. (eds), pp. 457-475.

Paul, J.H., Jiang, S.C., and Rose, J.B. (1991) Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. *Applied and Environmental Microbiology* **57**: 2197-2204.

Perler, F.B., Olsen, G.J., and Adam, E. (1997) Compilation and analysis of intein sequences. In, pp. 1087-1093.

Perler, F.B. (2002) InBase: the Intein Database. *Nucleic Acid Research* **30:**. 383-384.

Peters, A.F., and Ramirez, M.E. (2001) Molecular phylogeny of small brown algae, with special reference to the systematic position of *Caepidium antarcticum* (Adenocystaceae, *Ectocarpales*). *Cryptogamie Algologie* **22**: 187-200.

Peters, A.F., Marie, D., Scornet, D., Kloareg, B., and Cock, J.M. (2004) Proposal of *Ectocarpus siliculosus* (*Ectocarpales*, Phaeophyceae) as a model organism for brown algal genetics and genomics. *Journal of Phycology* **40**: 1079-1088.

Peterson, J., Prough, R. (1986) Cytochrome P-450 reductase and cytochrome b5 in cytochrome P-450 catalysis. In *Cytochrome P-450: structure, mechanism and biochemistry*. Ortiz de Montellano, P. (ed). New York: Plenum Press, pp. 89-117.

Pfeiffer, P., and Hohn, T. (1983) Involvement of reverse transcription in the replication of cauliflower mosaic virus: a detailed model and test of some aspects. *Cell* **33**: 781-789.

Pietrokovski, S. (1998a) Modular organization of inteins and C-terminal autocatalytic domains. *Protein Science* **7:** 64-71.

Pietrokovski, S. (1998b) Identification of a virus intein and a possible variation in the protein-splicing reaction. *Current Biology* **8**: R634-R638.

Pietrokovski, S. (2001) Intein spread and extinction in evolution. *Trends in Genetics* **17**: 465-472.

Piganeau, G., and Moreau, H. (2007) Screening the Sargasso Sea metagenome for data to investigate genome evolution in *Ostreococcus* (*Prasinophyceae*, Chlorophyta). *Gene* **406**: 184-190.

Pingree, R.D., Holligan, P.M., and Mardell, G.T. (1978) The effects of vertical stability on phytoplankton distributions in the summer on the northwest European shelf. *Deep-Sea Research Part* **25**: 1011-1028.

Pitta, P., Giannakourou, A., Divanach, P., and Kentouri, M. (1998) Planktonic food web in marine mesocosms in the eastern Mediterranean: bottom-up or top-down regulation? *Hydrobiologia* **363**: 97-105.

Pomeroy, L.R. (1974) The ocean food web, a changing paradigm. *Bioscience* **24**: 499-504.

Prescott, A.G. (1993) A dilemma of dioxygenases: or where molecular biology and biochemistry fail to meet. *Journal of Experimental Botany* **44**: 849-861.

Prockop, J.D. (1995) Collagens: Molecular biology, diseases, and potentials for therapy. *Annual Review of Biochemistry* **64**: 403-434.

Rao, S., and Rossmann, M.G. (1973) Comparison of super-secondary structures in proteins. *Journal of Molecular Biology* **76**: 241-256.

Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., and Claverie, J.M. (2004) The 1.2-megabase genome sequence of Mimivirus. *Science* **306**: 1344-1350.

Reed, W., Carroll, J., Agramonte, A., Lazear, J. (1901) Senate documents. **66**: 156.

Reisser, W., and Kapaun, E. (1991) Entry of a *Chlorella* virus into its host cell. *Journal of Phycology* **27**: 609-613.

Reisser, W. (1993) Viruses and virus-like particles of freshwater and marine eukaryotic algae - a review. *Archiv Fur Protistenkunde* **143**: 257-265.

Rines, J.E.B., and Hargraves, P.E. (1987) The seasonal distribution of the marine diatom genus *Chaetoceros ehr* in Narragansett Bay, Rhode Island (1981-1982). *Journal of Plankton Research* **9**: 917-933.

Ring, B., Yarnell, W., and Roberts, J. (1996) Function of *E. coli* RNA polymerase sigma factor sigma 70 in promoter-proximal pausing. *Cell* **86**: 485 - 493.

Roca, J. (1995) The mechanisms of DNA topoisomerases. *Trends in Biochemical Sciences* **20**: 156-160.

Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A. Arellano, A., Coleman, M., Hauser, L., Hess, W. R., Johnson, Z. I., Land, M., Lindell, D., Post, A. F., Regala, W., Shah, M., Shaw, S. L., Steglich, C., Sullivan, M. B., Ting, C. S., Tolonen, A., Webb, E. A., Zinser, E. R., Chisholm, S. W. (2003) Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042-1047.

Rodriguez, F., Fernandez, E., Head, R.N., Harbour, D.S., Bratbak, G., Heldal, M., and Harris, R.P. (2000) Temporal variability of viruses, bacteria, phytoplankton and zooplankton in the western English Channel off Plymouth. *Journal of the Marine Biological Association of the United Kingdom* **80**: 575-586.

Rodriguez, F., Derelle, E., Guillou, L., Le Gall, F., Vaulot, D., and Moreau, H. (2005) Ecotype diversity in the marine picoeukaryote *Ostreococcus* (Chlorophyta, *Prasinophyceae*). *Environmental Microbiology* **7**: 853-859.

Rohozinski, J., Girton, L.E., and Van Etten, J.L. (1989) *Chlorella* viruses contain linear nonpermuted double-Stranded DNA genomes with covalently closed hairpin ends. *Virology* **168**: 363-369.

Rohwer, F., and Edwards, R. (2002) The phage proteomic tree: a genome-based taxonomy for phage. *Journal of Bacteriology* **184**: 4529-4535.

Romari, K., and Vaulot, D. (2004) Composition and temporal variability of picoeukaryote communities at a coastal site of the English Channel from 18S rDNA sequences. *Limnology and Oceanography* **49**: 784-798.

Rondon, M.R., August, P.R., Bettermann, A.D., Brady, S.F., Grossman, T.H., Liles, M.R., Loiacono, K.A., Lynch, B.A., MacNeil, I.A., Minor, C., Tiong, C.L., Gilman, M., Osburne, M.S., Clardy, J., Handelsman, J. And Goodman, R.M. (2000) Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. *Applied and Environmental Microbiology* **66**: 2541-2547.

Rose, M.G., Farrell, M.P., and Schmitz, J.C. (2002) Thymidylate synthase: a critical target for cancer chemotherapy. *Clinical Colorectal Cancer* **I**: 220-229.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and Barrell, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics* **16**: 944-945.

Safferman, R.S., and Morris, M.E. (1963) Algal virus: Isolation. *Science* **140**: 679-680.

Sahlsten, E. (1998) Seasonal abundance in Skagerrak-Kattegat coastal waters and host specificity of viruses infecting the marine photosynthetic flagellate *Micromonas pusilla*. *Aquatic Microbial Ecology* **16**: 103-108.

Sahlsten, E., and Karlson, B. (1998) Vertical distribution of virus-like particles (VLP) and viruses infecting *Micromonas pusilla* during late summer in the southeastern Skagerrak, North Atlantic. *Journal of Plankton Research* **20**: 2207-2212.

Sandaa, R.A., Heldal, M., Castberg, T., Thyrhaug, R., and Bratbak, G. (2001) Isolation and characterization of two viruses with large genome size infecting *Chrysochromulina ericina* (*Prymnesiophyceae*) and *Pyramimonas orientalis* (*Prasinophyceae*). *Virology* **290**: 272-280.

Sandaa, R.A. (2008) Burden or benefit? Virus-host interaction in the marine environment. *Research in Microbiology* **159:** 374-381.

Scanlan, D.J., Mann, N.H., and Carr, N.G. (1993) The response of the picoplanktonic marine cyanobacterium *Synechococcus* species WH7803 to phosphate starvation involves a protein homologous to the periplasmic phosphate-binding protein of *Escherichia coli. Molecular Microbiology* **10**: 181-191.

Scanlan, D.J., Silman, N.J., Donald, K.M., Wilson, W.H., Carr, N.G., Joint, I., and Mann, N.H. (1997) An immunological approach to detect phosphate stress in populations and single cells of photosynthetic picoplankton. *Applied and Environmental Microbiology* **63**: 2411-2420.

Scanlan, D.J. (2003) Physiological diversity and niche adaptation in marine *Synechococcus. Advances in Microbial Physiology, Vol 47* **47**: 1-64.

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

Schroeder, D.C., Oke, J., Malin, G., and Wilson, W.H. (2002) Coccolithovirus (*Phycodnaviridae*): Characterisation of a new large dsDNA algal virus that infects *Emiliania huxleyi*. *Archives of Virology* **147**: 1685-1698.

Schroeder, D.C., Oke, J., Hall, M., Malin, G., and Wilson, W.H. (2003) Virus succession observed during an *Emiliania huxleyi* bloom. *Applied and Environmental Microbiology* **69**: 2484-2490.

Schroeder, D.C., Park, Y., Yoon, H.-M., Lee, Y.S., Kang, S.W., Meints, R.H., Ivey, R.G., Choi, T.J. (2009) Genomic analysis of the smallest giant virus - *Feldmannia* sp. virus 158. *Virology* **384**: 223-232.

Schuster, A.M., Burbank, D.E., Meister, B., Skrdla, M.P., Meints, R.H., Hattman, S. Swinton, D., Van Etten J.L. (1986) Characterization of viruses infecting a eukaryotic *Chlorella*- like green alga. *Virology* **150**: 170-177.

Seaton, G.G.R., Hurry, V.M., and Rohozinski, J. (1996) Novel amplification of non-photochemical chlorophyll fluorescence quenching following viral infection in *Chlorella*. *FEBS Letters* **389**: 319-323.

Sengco, M.R., Brautigam, M., Kapp, M., and Muller, D.G. (1996) Detection of virus DNA in *Ectocarpus siliculosus* and *E. fasciculatus* (Phaeophyceae) from various geographic areas. *European Journal of Phycology* **31**: 73-78.

Shackleton, L.A., and Holmes, E.C. (2004) The evolution of large DNA viruses: combining genomic information of viruses and their hosts. *Trends in Microbiology* **12**: 458-465.

Sherman, L.A., and Brown, R.M. (1978) Cyanophages and viruses of eukaryotic algae. In *Comprehensive Virology*. Fraenkel-Conrat, H., and Wagner, R.R. (eds). New York: Plenum Press, pp. 145-234.

Sherr, E.B., and Sherr, B.F. (1991) Planktonic microbes - tiny cells at the base of the oceans food webs. *Trends in Ecology & Evolution* **6**: 50-54.

Shinada, A., Ikeda, T., Ban, S., and Tsuda, A. (2001) Seasonal dynamics of planktonic food chain in the Oyashio region, western subarctic Pacific. *Journal of Plankton Research* **23**: 1237-1247.

Shirai, Y., Tomaru, Y., Takao, Y., Suzuki, H., Nagumo, T., and Nagasaki, K. (2008) Isolation and characterization of a single-stranded RNA virus infecting the marine planktonic diatom *Chaetoceros tenuissimus* Meunier. *Applied and Environmental Microbiology* **74**: 4022-4027.

Short, C.M., and Suttle, C.A. (2005) Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. *Applied and Environmental Microbiology* **71**: 480-486.

Short, S.M., and Suttle, C.A. (2002) Sequence analysis of marine virus communities reveals that groups of related algal viruses are widely distributed in nature. *Applied and Environmental Microbiology* **68**: 1290-1296.

Short, S.M., and Suttle, C.A. (2003) Temporal dynamics of natural communities of marine algal viruses and eukaryotes. *Aquatic Microbial Ecology* **32**: 107-119.

Silva, P., Thompson, E., Bailey, S., Kruse, O., Mullineaux, C.W., Robinson, C., Mann, N.H., Nixon, P.J. (2003) FtsH is involved in the early stages of repair of photosystem II in *Synechocystis* sp PCC 6803. *Plant Cell* **15**: 2152-2164.

Sime-Ngando, T., and Colombet, J. (2009) Viruses and prophages in aquatic ecosystems. *Canadian Journal of Microbiology* **55**: 95-109.

Simpson, A.A., Nandhagopal, N., Van Etten, J.L., and Rossmann, M.G. (2003) Structural analyses of *Phycodnaviridae* and *Iridoviridae. Acta Crystallographica Section D-Biological Crystallography* **59**: 2053-2059.

Sindelar, L. (1986) The content of ATP, ADP, AMP, PI, the activity of enzymes involved in the glycolytic pathway and some problems of its regulation and energy-balnace in tobacco plants infected wtih potato virus Y. *Biologia Plantarum* **28**: 449-459.

Sindelar, L., Sindelarova, M., and Burketova, L. (1999) Changes in activity of glucose-6-phosphate and 6-phosphogluconate dehydrogenase isozymes upon potato virus Y infection in tobacco leaf tissues and protoplasts. *Plant Physiology and Biochemistry* **37**: 195-201.

Six, C., Finkel, Z.V., Rodriguez, F., Marie, D., Partensky, F., and Campbell, D.A. (2008) Contrasting photoacclimation costs in ecotypes of the marine eukaryotic picoplankter *Ostreococcus*. *Limnology and Oceanography* **53**: 255-265.

Smith, M.A., Cross, A.R., Jones, O.T.G., Griffiths, W.T., Stymne, S., and Stobart, K. (1990) Electron transport components of the 1-acyl-2-oleoyl-sn-glycero-3-phosphocholine delta-12-desaturase (delta-12-desaturase) in microsomal preparations

from developing safflower (*Carthamus tinctorius* L) cotyledons. *Biochemical Journal* **272**: 23-29.

Smith, M.A., Jonsson, L., Stymne, S., and Stobart, K. (1992) Evidence for cytochrome-b_5 as an electron donor in ricinoleic acid biosynthesis in microsomal preparations from developing castor bean (*Ricinus communis* L). *Biochemical Journal* **287**: 141-144.

Smith, M.A., Stobart, A.K., Shewry, P.R., Napier, J.A. (1998) Cytochrome b_5 and polyunsaturated fatty acid biosynthesis. In *Engineering Crop Plants for Industrial End Uses*. P.R. Shewry, J.A.Napier, P.J. Davis (ed). London: Portland Press, pp. 181-188.

Smith, W.O., Codispoti, L.A., Nelson, D.M., Manley, T., Buskey, E.J., Niebauer, H.J., and Cota, G.F. (1991) Importance of *Phaeocystis* blooms in the high latitude ocean carbon cycle. *Nature* **352**: 514-516.

Sode, K., Oozeki, M., Asakawa, K., Burgess, J.G., and Matsunaga, T. (1994) Isolation of a marine cyanophage infecting the marine unicellular cyanobacterium, *Synechococcus* sp. NKBG 042902. *Journal of Marine Biotechnology* **1**: 189-192.

Steward, G.F., Fandino, L.B., Hollibaugh, J.T., Whitledge, T.E., and Azam, F. (2007) Microbial biomass and viral infections of heterotrophic prokaryotes in the sub-surface layer of the central Arctic Ocean. *Deep-Sea Research Part I-Oceanographic Research Papers* **54**: 1744-1757.

Stoeck, T., and Epstein, S. (2003) Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments. *Applied and Environmental Microbiology* **69**: 2657-2663.

Strasser, P., Zhang, Y.P., Rohozinski, J., and Van Etten, J.L. (1991) The termini of the *Chlorella* virus PBCV-1 genome are identical 2.2 kbp inverted repeats. *Virology* **180**: 763-769.

Sullivan, M.B., Coleman, M.L., Weigele, P., Rohwer, F., and Chisholm, S.W. (2005) Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* **3**: e144.

Sullivan, M.B., Lindell, D., Lee, J.A., Thompson, L.R., Bielawski, J.P., and Chisholm, S.W. (2006) Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *Plos Biology* **4**: 1344-1357.

Suttle, C.A., Chan, A.M., and Cottrell, M.T. (1990) Infection of phytoplankton by viruses and reduction of primary productivity. *Nature* **347**: 467-469.

Suttle, C.A., Chan, A.M., and Cottrell, M.T. (1991) Use of ultrafiltration to isolate viruses from seawater which are pathogens of marine phytoplankton. *Applied and Environmental Microbiology* **57**: 721-726.

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. (1995) Viruses infecting the marine Prymnesiophyte *Chrysochromulina* sp. - Isolation, preliminary characterization and natural abundance. *Marine Ecology-Progress Series* **118**: 275-282.

Suttle, C.A. (2000) Ecological, evolutionary and geochemical consequences of viral infection of cyanobacteria and eukaryotic algae. In *Viral Ecology*. Hurst, C.J. (ed): Academic Press, pp. 247-296.

Suttle, C.A. (2007) Marine viruses - major players in the marine global ecosystem. *Nature Reviews Microbiology* **5**: 801-812.

Sym, S., and Pienaar, R. (1993) The class *Prasinophyceae*. In *Progress in phycological research*. Round, F., and Chapman, D. (eds). Bristol, UK: Biopress Ltd.

Tai, V., Lawrence, J.E., Lang, A.S., Chan, A.M., Culley, A.I., and Suttle, C.A. (2003) Characterization of HaRNAV, a single-stranded RNA virus causing lysis of *Heterosigma akashiwo* (Raphidophyceae). *Journal of Phycology* **39**: 343-352.

Tanaka, T., and Rassoulzadegan, F. (2004) Vertical and seasonal variations of bacterial abundance and production in the mesopelagic layer of the NW Mediterranean Sea: bottom-up and top-down controls. *Deep-Sea Research Part I-Oceanographic Research Papers* **51**: 531-544.

Tarutani, K., Nagasaki, K., and Yamaguchi, M. (2000) Viral impacts on total abundance and clonal composition of the harmful bloom-forming phytoplankton *Heterosigma akashiwo*. *Applied and Environmental Microbiology* **66**: 4916-4920.

Tarutani, K., Nagasaki, K., Itakura, S., and Yamaguchi, M. (2001) Isolation of a virus infecting the novel shellfish-killing dinoflagellate *Heterocapsa circularisquama. Aquatic Microbial Ecology* **23**: 103-111.

Tarutani, K., Nagasaki, K., and Yamaguchi, M. (2006) Virus adsorption process determines virus susceptibility in *Heterosigma akashiwo* (Raphidophyceae). *Aquatic Microbial Ecology* **42**: 209-213.

Teller, J.H., Powers, S.G., Snell, E.E. (1976) Ketopantoate hydroxymethyltransferase. I. Purification and role in pantothenate biosynthesis. *Journal of Biological Chemistry* **251**: 3780-3785.

Teodoro, J.G., and Branton, P.E. (1997) Regulation of apoptosis by viral gene products. *Journal of Virology* **71**: 1739-1746.

Thingstad, T.F., Heldal, M., Bratbak, G., and Dundas, I. (1993) Are Viruses Important Partners in Pelagic Food Webs. *Trends in Ecology & Evolution* **8**: 209-213.

Thingstad, T.F., Hagstrom, A., and Rassoulzadegan, F. (1997) Accumulation of degradable DOC in surface waters: Is it caused by a malfunctioning microbial loop? *Limnology and Oceanography* **42**: 398-404.

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, J.D., Higgins, D.G., and Gibson, T.J. (1994) Clustal-W - Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position- specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.

Throndsen, J. (1976) Occurrence and productivity of small marine flagellates. *Norwegian Journal of Botany* **23**: 269-293.

Throndsen, J., and Kristiansen, S. (1991) *Micromonas pusilla* (*Prasinophyceae*) as part of picoplankton and nanoplankton communities of the Barents Sea. *Polar Research* **10**: 201-207.

Thyrhaug, R., Larsen, A., Thingstad, T.F., and Bratbak, G. (2003) Stable coexistence in marine algal host-virus systems. *Marine Ecology Progress Series* **254**: 27-35.

Tomaru, Y., Tarutani, K., Yamaguchi, M., and Nagasaki, K. (2004) Quantitative and qualitative impacts of viral infection on a *Heterosigma akashiwo* (Raphidophyceae) bloom in Hiroshima Bay, Japan. *Aquatic Microbial Ecology* **34**: 227-238.

Tonetti, M., Zanardi, D., Gurnon, J.R., Fruscione, F., Armirotti, A., Damonte, G., Sturla, L., De Flora, A., and Van Etten, J.L. (2003) *Paramecium bursaria Chlorella* virus-1 encodes two enzymes involved in the biosynthesis of GDP-L-fucose and GDP-D-rhamnose. *Journal of Biological Chemistry* **278**: 21559-21565.

Torriani, A.M. (1990) From cell membrane to nucleotides - the phosphate regulon in *Escherichia coli*. *Bioessays* **12**: 371-376.

Towler, D.G., Glaser, L. (1986) Protein fatty acid acylation: enzymatic synthesis of an N -myristoylglycyl peptide. *Proceedings of the National Academy of Sciences of the United States of America* **83**: 2812-2816.

Travers, A.A. (1969) Bacteriophage sigma factor for RNA polymerase. *Nature* **223**: 1107-1110.

Tremblay, G., Belzile, C., Gosselin, M., Poulin, M., Roy, S., and Tremblay, J.E. (2009) Late summer phytoplankton distribution along a 3500 km transect in Canadian Arctic waters: strong numerical dominance by picoeukaryotes. *Aquatic Microbial Ecology* **54**: 55-70.

Tzaras, A., Pick, F.R., Mazumder, A., and Lean, D.R.S. (1999) Effects of nutrients, planktivorous fish and water column depth on components of the microbial food web. *Aquatic Microbial Ecology* **19**: 67-80.

Tzlil, S., Kindt, J.T., Gelbart, W.M., and Ben-Shaul, A. (2003) Forces and pressures in DNA packaging and release from viral capsids. *Biophysical Journal* **84**: 1616-1627.

Van den Hoek, C., Mann, D.G., and Jahns, H.M. (1995) *Algae: an introduction to phycology*. Cambridge: Cambridge University Press.

Van Etten, J.L. (1995) Giant *Chlorella* Viruses. *Molecules and Cells* **5**: 99-106.

Van Etten, J.L. (2003) Unusual life style of giant *Chlorella* viruses. *Annual Review of Genetics* **37**: 153-195.

Van Etten, J.L., and Ghabrial, S.A. (1991) *Phycodnaviridae*. In *Classification and Nomenclature of Viruses Arch Virol Suppl 2*. FRrancki, R.I.B., Fauget, C.M., Knudson, D.L., and Brown, F. (eds). Vienna: Springer-Verlag, pp. 137-139.

Van Etten, J.L., and Meints, R.H. (1999) Giant viruses infecting algae. *Annual Review of Microbiology* **53**: 447-494.

Van Etten, J.L., Meints, R.H., Burbank, D.E., Kuczmarski, D., Cuppels, D.A., and Lane, L.C. (1981) Isolation and characterization of a virus from the intracellular green alga symbiotic with *Hydra viridis*. *Virology* **113**: 704-711.

Van Etten, J.L., Meints, R.H., Kuczmarski, D., Burbank, D.E., and Lee, K. (1982) Viruses of symbiotic *Chlorella*-like algae isolated from *Paramecium bursaria* and *Hydra viridis*. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **79**: 3867-3871.

Van Etten, J.L., Burbank, D.E., Xia, Y., and Meints, R.H. (1983a) Growth cycle of a virus, PBCV-1, that infects *Chlorella*-like algae. *Virology* **126**: 117-125.

Van Etten, J.L., Burbank, D.E., Kuczmarski, D., and Meints, R.H. (1983b) Virus infection of culturable *Chlorella*-like algae and development of a plaque assay. *Science* **219**: 994-996.

Van Etten, J.L., Burbank, D.E., Joshi, J., and Meints, R.H. (1984) DNA synthesis in a *Chlorella*-like alga following infection with the virus PBCV-1. *Virology* **134**: 443-449.
Van Etten, J.L., Burbank, D.E., Schuster, A.M., and Meints, R.H. (1985) Lytic viruses infecting a *Chlorella*-Like alga. *Virology* **140**: 135-143.

Van Etten, J.L., Lane, L.C., and Meints, R.H. (1991) Viruses and virus-like particles of eukaryotic algae. *Microbiological Reviews* **55**: 586-620.

Van Etten, J.L., Graves, M.V., Muller, D.G., Boland, W., and Delaroque, N. (2002) *Phycodnaviridae* - large DNA algal viruses. *Archives of Virology* **147**: 1479-1516.

Vardi, A., Berman-Frank, I., Rozenberg, T., Hadas, O., Kaplan, A., and Levine, A. (1999) Programmed cell death of the dinoflagellate *Peridinium gatunense* is mediated by CO_2 limitation and oxidative stress. *Current Biology* **9**: 1061-1064.

Vaulot, D., Eikrem, W., Viprey, M., and Moreau, H. (2008) The diversity of small eukaryotic phytoplankton (<= 3 υm) in marine ecosystems. *FEMS Microbiology Reviews* **32**: 795-820.

Vergères, G., and Waskell, L. (1995) Cytochrome b_5, its functions, structure and membrane topology. *Biochimie* **77**: 604-620.

Verity, P.G., and Smetacek, V. (1996) Organism life cycles, predation, and the structure of marine pelagic ecosystems. *Marine Ecology-Progress Series* **130**: 277-293.

Verity, P.G., Redalje, D.G., Lohrenz, S.R., Flagg, C., and Hristov, R. (2002) Coupling between primary production and pelagic consumption in temperate ocean margin pelagic ecosystems. *Deep-Sea Research Part Ii-Topical Studies in Oceanography* **49**: 4553-4569.

Villarreal, L.P., and DeFilippis, V.R. (2000) A hypothesis for DNA viruses as the origin of eukaryotic replication proteins. *Journal of Virology* **74**: 7079-7084.

Wang, I.N., Li, Y., Que, Q.D., Bhattacharya, M., Lane, L.C., Chaney, W.G., and Van Etten, J.L. (1993) Evidence for virus encoded glycosylation specificity. *Proceedings of the National Academy of Sciences of the United States of America* **90**: 3840-3844.

Wanner, B.L. (1993) Gene regulation by phosphate in enteric bacteria. *Journal of Cell Biochemistry* **51**: 47-54.

Wanner, J.L. (1996) Phosphorus assimilation and control of the phosphate regulon. . In *Escherichia coli and Salmonella: Cellular and Molecular Biology*. Neidhardt, R.C.I., Ingrahan, J.L., Lin, E.C.C., Low, K.B., Magasanik, B., Reznikoff, W.S., Riley, M., Schaechter, M., Umbrager, H.E. (ed). Washington, D.C.: American Society for Microbiology, pp. 1357-1381.

Warbrick, E. (2000) The puzzle of PCNA's many partners. *Bioessays* **22**: 997-1006.

Waterbury, J.B., and Valois, F.W. (1993) Resistance to cooccurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Applied and Environmental Microbiology* **59**: 3393-3399.

Waters, E., Hohn, M.J., Ahel, I., Graham, D.E., Adams, M.D., Barnstead, M.., Beeson, K.Y., Bibbs, L., Bolanos, R., Kelle, M., Kretz, K., Lin, X., Mathur, E., Ni, J., Podar, M., Richardson, T., Sutton, G.G., Simon, M., Soll, D., Stetter, K.O., Short, J.M., and Noordewier, M. (2003) The genome of *Nanoarchaeum equitans*: Insights into early archaeal evolution and derived parasitism. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 12984-12988.

Waters, R.E., and Chan, A.T. (1982) *Micromonas pusilla* Virus - the virus growth-cycle and associated physiological events within the host cells - host range mutation. *Journal of General Virology* **63**: 199-206.

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. (1995) Significance of viruses versus heterotrophic nanoflagellates for controlling bacterial abundance in the Northern Adriatic Sea. *Journal of Plankton Research* **17**: 1851-1856.

Weinbauer, M.G., and Rassoulzadegan, F. (2004) Are viruses driving microbial diversification and diversity? *Environmental Microbiology* **6**: 1-11.

Weinbauer, M.G., Brettar, I., and Hofle, M.G. (2003) Lysogeny and virus-induced mortality of bacterioplankton in surface, deep, and anoxic marine waters. *Limnology and Oceanography* **48**: 1457-1465.

Weinbauer, M.G., Wilhelm, S.W., Suttle, C.A., and Garza, D.R. (1997) Photoreactivation compensates for UV damage and restores infectivity to natural marine virus communities. *Applied and Environmental Microbiology* **63**: 2200-2205.

Westley, J., Adler, H., Westley, L., and Nishida, C. (1983) The sulfurtransferases. *Fundamental and Applied Toxicology* **3**: 377-382.

Weynberg, K.D., Allen, M.J., Ashelford, K., Scanlan, D.J., and Wilson, W.H. (2009) From small hosts come big viruses: the complete genome of a second *Ostreococcus tauri* virus, OtV-1. *Environmental Microbiology*.

Wilcox, C., Hu, J.S., and Olson, E.N. (1987) Acylation of proteins with myristic acid occurs cotranslationally. *Science* **238**: 1275-1278.

Wilhelm, S.W., and Suttle, C.A. (1999) Viruses and nutrient cycles in the sea - Viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* **49**: 781-788.

Wilhelm, S.W., Weinbauer, M.G., Suttle, C.A., Pledger, R.J., and Mitchell, D.L. (1998) Measurements of DNA damage and photoreactivation imply that most viruses in marine surface waters are infective. *Aquatic Microbial Ecology* **14**: 215-222.

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., Tarran, G., and Zubkov, M.V. (2002a) Virus dynamics in a coccolithophore-dominated bloom in the North Sea. *Deep-Sea Research Part Ii-Topical Studies in Oceanography* **49**: 2951-2963.

Wilson, W.H., Tarran, G.A., Schroeder, D., Cox, M., Oke, J., and Malin, G. (2002b) Isolation of viruses responsible for the demise of an *Emiliania huxleyi* bloom in the English Channel. *Journal of the Marine Biological Association of the UK* **82**: 369-377.

Wilson, W.H., Van Etten, J.L., Schroeder, D.S., Nagasaki, K., Brussaard, C., Delaroque, N. *et al.,* (2005a) Family: *Phycodnaviridae*. In *Virus Taxonomy, VIIIth ICTV Report*. Fauquet, C.M., Mayo, M.A., Maniloff, J., Dusselberger, U., and Ball, L.A. (eds). London: Elsevier/Academic Press, pp. 163-175.

Wilson, W.H., Schroeder, D.C., Allen, M.J., Holden, M.T.G., Parkhill, J., Barrell, B.G., Churcher, C., Hamlin, N., Mungall, K., Norbertczak, H., Quail, M.A., Price, C. Rabbinowitsch, E., Walker, D., Craigon, M., Roy, D. and Ghazal, P. (2005b) Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science* **309**: 1090-1092.

Wilson, W.H., Van Etten, J.L., and Allen, M.J. (2009) The *Phycodnaviridae:* The story of how tiny giants rule the world. Current Topics in Microbiology and Immunology 328. In *Lesser Known Large dsDNA Viruses*. Van Etten, J.L. (ed). Berlin Heidelberg: Springer-Verlag.

Wolf, S., Muller, D.G., and Maier, I. (2000) Assembly of a large icosahedral DNA virus, MclaV-1, in the marine alga *Myriotrichia clavaeformis* (Dictyosiphonales, Phaeophyceae). *European Journal of Phycology* **35**: 163-171.

Wommack, K.E., and Colwell, R.R. (2000) Virioplankton: Viruses in aquatic ecosystems. *Microbiology and Molecular Biology Reviews* **64**: 69-114.

Wommack, K.E., Hill, R.T., and Colwell, R.R. (1995) A simple method for the concentration of viruses from natural water samples. *Journal of Microbiological Methods* **22**: 57-67.

Wood, T. (1986) Distribution of the pentose phosphate pathway in living organisms. *Cell Biochemical Functioning* **4**: 235-240.

Worden, A.Z. (2006) Picoeukaryote diversity in coastal waters of the Pacific Ocean. *Aquatic Microbial Ecology* **43**: 165-175.

Worden, A.Z., Nolan, J.K., and Palenik, B. (2004) Assessing the dynamics and ecology of marine picophytoplankton: the importance of the eukaryotic component. *Limnology and Oceanography* **49**: 168-179.

Worden, A.Z., Lee, J.H., Mock, T., Rouze, P., Simmons, M.P., Aerts, A.L.., (2009) Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* **324**: 268-272.

Xia, Y.N., Morgan, R., Schildkraut, I., and Van Etten, J.L. (1988) A site-specific single-strand endonuclease activity induced by NYS-1 virus infection of a *Chlorella*-like green alga. *Nucleic Acids Research* **16**: 9477-9487.

Xiao, C.A., Chipman, P.R., Battisti, A.J., Bowman, V.D., Renesto, P., Raoult, D., and Rossmann, M.G. (2005) Cryo-electron microscopy of the giant mimivirus. *Journal of Molecular Biology* **353**: 493-496.

Yamada, T., Onimatsu, H., and Van Etten, J.L. (2006) *Chlorella* viruses. *Advances in Virus Research, Vol 66* **66**: 293-+.

Yan, X.D., Chipman, P.R., Castberg, T., Bratbak, G., and Baker, T.S. (2005) The marine algal virus PpV01 has an icosahedral capsid with T=219 quasisymmetry. *Journal of Virology* **79**: 9236-9243.

Yan, X.D., Olson, N.H., Van Etten, J.L., Bergoin, M., Rossmann, M.G., and Baker, T.S. (2000) Structure and assembly of large lipid-containing dsDNA viruses. *Nature Structural Biology* **7**: 101-103.

Zauberman, N., Mutsafi, Y., Ben Halevy, D., Shimoni, E., Klein, E., Xiao, C., Sun, S., and Minsky, A. (2008) Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga Mimivirus. *Plos Biology* **6**: 1104-1114.

Zeidner, G., Bielawski, J.P., Shmoish, M., Scanlan, D.J., Sabehi, G., and Beja, O. (2005) Potential photosynthesis gene recombination between *Prochlorococcus* and *Synechococcus* via viral intermediates. *Environmental Microbiology* **7**: 1505-1513.

Zhang, Y., Xiang, Y., Van Etten, J.L., and Rossmann, M.G. (2007) Structure and Function of a *Chlorella* virus encoded glycosyltransferase. *Structure* **15**: 1031-1039.

Zhang, Y.P., Nelson, M., Nietfeldt, J., Xia, Y.N., Burbank, D., Ropp, S., and Van Etten, J.L. (1998) *Chlorella* virus NY-2A encodes at least 12 DNA endonuclease/methyltransferase genes. *Virology* **240**: 366-375.

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.

Zhu, F., Massana, R., Not, F., Marie, D., and Vaulot, D. (2005) Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiology Ecology* **52**: 79-92.

Zingone, A., Sarno, D., and Forlani, G. (1999) Seasonal dynamics in the abundance of *Micromonas pusilla* (*Prasinophyceae*) and its viruses in the Gulf of Naples (Mediterranean Sea). *Journal of Plankton Research* **21**: 2143-2159.

Zingone, A., Natale, F., Biffali, E., Borra, M., Forlani, G., and Sarno, D. (2006) Diversity in morphology, infectivity, molecular characteristics and induced host resistance between two viruses infecting *Micromonas pusilla. Aquatic Microbial Ecology* **45**: 1-14.

Appendices

Appendix I. Table of primers used in this study to close contigs in the sequence finishing of the OtV-1 and OtV-2 genomes and to confirm 454 sequence data of the OtV-1 CDS OtV1_117, a putative virus inclusion body gene.

OtV-1 primers	
OtV105677_F	5'-tatgggactctttgcgggta-3'
OtV105677_R	5'-tatttgacccctcgatcgtc-3'
OtV105669_F	5'-gtcagatcgctcaaggcaat-3'
OtV105669_R	5'-gacattgccttgagcgatct-3'
OtV105510_F	5'-cattggtggaagaggcaact-3'
OtV105510_R	5'-cagaggtgccttgagaggtc-3'
OtV105632_F	5'-tttgccatcccaaaaacct-3'
OtV105632_R	5'-tgggtgataccaatatatttgacttt-3'
OtV105629_F	5'-ggcggatctcaaggctaac-3'
OtV105629_R	5'-attcttgtcccgaggtggac-3'
OtV105350_F	5'-gggttgtgagtcctcaaagg-3'
OtV105350_R	5'-gcgatgccttaaaactccaa-3'
OtV105617_F	5'-tccgcaatttggagtggt-3'
OtV105617_R	5'-ctccaaattgcggagagaaa-3'
OtV10009_F	5'-ccaggttcttgatccgtgtt-3'
OtV10009_R	5'-cgcgagactttcaaggtga-3'
OtV105353_F	5'-atgccaataccttcggagtg-3'
OtV105353_R	5'-ttgacggtgcgtcatatcat-3'
OtV105539_F	5'-ggcgcattactctgtgcttt-3'
OtV105539_R	5'-gcacagagtaatgcgccaat-3'
OtV10007_F	5'-ggggtgagggacatttttg-3'
OtV10007_R	5'-ttgcccatggatgatgaata-3'
OtV105633_F	5'-cgacatcgagctttgtgaaa-3'
OtV105633_R	5'-tgatggtggaaacactaatataaag-3'
OtV105668_F	5'-cctctcaaggcaccaacg-3'
OtV105668_R	5'-gcggagatgccttgattgta-3'
OtV105674_F	5'-cgaagatgccgtgaaaagat-3'
OtV105674_R	5'-ccaggacttgcatgggtact-3'
OtV105675_F	5'-ccctgtggtcccatacattc-3'
OtV105675_F	5'-gaggatttcgcgatggtaaa-3'
OtV105676_F	5'-gacccagaagacctgacacc-3'
OtV105676_R	5'-acttatcggggattgggaag-3'
OtV105670_F	5'-ctcaaccctctatgggacga-3'
OtV105670_R	5'-tccaaataaacttaatccaatacc-3'
OtV10008_R	5'TACCAAGCCCCTCCAGATAG-3'
OtV10008_F	5' TGGCTTCCAGTGTCATACAC-3'

OtV-2 primers	
OV2-180_R	5'-CGGGTTGGTTAGGGTGATTA-3'
OV2-0324_F	5'-TTCGTCCAATCCAAATCCAT-3'
OV2-0324_R	5'-CAAAAAGTAGGGTTGGCAAGG-3'
OV2-0332_F	5'-AGTGTGTTGAAGACCGACCA-3'
OV2-0332_R	5'-TGAATGGTCGGTCTTCAACA-3'
OV2-0416_F	5'-CATTTCCATGAATTGCGTTG-3'
OV2-0416_R	5'-CTGCATTAACGACGCAAAAA-3'
OV2-0417_F	5'-AAAACAGGGCTGCAAACAAC-3'
OV2-0417_R	5'-CAGCCCTGTTTTTGGTTGTT-3'
OV2-0438_F	5'-GAGCATGTCGCCCACTTTAT-3'
OV2-0438_R	5'-TACCCTACGAAGACGGTTGG-3'
OV2-0732_F	5'-TCGCCAAAGGTTCTAAGCAG-3'
OV2-0732_R	5'-GCTTAGAACCTTTGGCGAGA-3'
OV2-0960_F	5'-CAATTCGTATGGGCAACAGA-3'
OV2-0960_R	5'-CACCTGCATCGAGATTTGAG-3'
OV2-03621_F	5'-CTTGCTTGAGCCATCGTGTA-3'
OV2-03621_R	5'-AACACACCTGGACGCCTTAC-3'
OV2_03622_F	5'-TCGCATTTTTCCATTCTGT

	G-3'
OV2_03622_R	5'-TGGAAAAATGCGATCAACAC-3'
OV2-03624_R	5'-CCTTCCACTGTTCGGATTTC-3'
OV2-03624_F	5'-CTCGCACAAAAAGATGACGA-3'
OV2-3625_F	5'-GAGATGAGTCAGCGGTCGTT-3'
OV2-3625_R	5'-AACGACCGCTGACTCATCTC-3'
OV2-3626_F	5'-GTAACAGGCGGTTGTGGTTT-3'
OV2-3626_R	5'-AAACCACAACCGCCTGTTAC-3'
OV2-3627_F	5'-CCCAGTTTGACGACTTACGG-3'
OV2-3627_R	5'-ATCACGAAACCCGTAAGTCG-3'
OV2-3630_F	5'-TGAGCATCTGGCTTTTCCTT-3'
OV2-3630_R	5'-AGGATCACTGGTTGCATGGT-3'
OV2-3681_F	5'-CAGGCCTTTACCGTGTCAAT-3'
OV2-3681_R	5'-CGGAGGGGATAAAAAGGAAG-3'
OV2-3709_F	5'-TCACCATCGTAAGTGCCATC-3'
OV2-3709_R	5'-AACCACCAATGTAGCGGAAG-3'
OV2-3744_F	5'-CCCTTTGTAAGCCAAAACCA-3'
OV2-3744_R	5'-TTGGAAGAAACAGCCAAAGC-3'

OV2-3784_F	5'-ATTCGTTTGCCAAGTGATGG-3'
OV2-3784_R	5'-TTTCACCATCCTACCGTCAA-3'
OV2-4056_F	5'-CACTGAACTCCACCCCTCTC-3'
OV2-4056_R	5'-TGATGAGAGAGGGGTGGAGT-3'
OV2-4063_F	5'-TAAAGGTACGGTTGCCATCA-3'
OV2-4063_R	5'-ATTGATGGCAACCGTACCTT-3'
OV2-4091_F	5'-AGTGTGTTGAAGACCGACCA-3'
OV2-4091_R	5'-TGAATGGTCGGTCTTCAACA-3'
OV2-4092_F	5'-ATCCGGCACATTTTGAAGAG-3'
OV2-4092_R	5'-CATGAGATGTTCGGCACAAG-3'
OV2-4212_F	5'-GTGAGGAGGGCAACCTTGTA-3'
OV2-4212_R	5'-GAGGGGCAAATTGTAGCAGA-3'
OV2-4215_F	5'-CGGTAGGAGTTCGCACAAAT-3'
OV2-4215_R	5'-GGAGGGACATGAGTTGGAAG-3'
OV2-4216_F	5'-ACGCTATCCAGGGTTGTGTC-3'
OV2-4216_R	5'-ATTTGTGCGAACTCCTACCG-3'
OV2-4487_F	5'-CCCGCTTGTTTGGTAACATC-3'
OV2-4487_R	5'-AAGAGTTCCCCCGTCTGTTT-3'
OV2-4508_F	5'-

	GCAAGTTCTGCGGTCTTC AT-3'
Putative Virus inclusion body gene (OtV1_117) primers	
VIB_F	5'-GAAAACGTCGCGAAGAA GAG-3'
VIB_R	5'-GCTTGGGATCTGTCCTCT TG-3'
VIB.1_F	5'-CGACCAACTCTTTGACGC AA-3'
VIB.1_R	5'-CTCCATTCTTCACAGCAC CA-3'
VIB.2_F	5'-GCGATGAGCTACTCGGAA -3'
VIB.2_R	5'-AGCAACCTTCACTGCCTT G-3'
VIB.3_F	5'-GATGATCTCGCAGCTCTC AA-3'
VIB.3_R	5'-GTCGGGTTGGTTGAGGTA GA-3'
VIB.4_F	5'-CGTTCCTGAACCGTATCA CA-3'
VIB.4_R	5'-CTTGGCTTCCTTCTCCCTT T-3'
VIB.5_F	5'-GATAACGGTATTCGGGCT CA-3'
VIB.5_R	5'-GAGCACGGTTCCTCACAT TT-3'
VIB.6_F	5'-GGCTGTCAATGCGAAGA AG-3'
VIB.6_R	5'-TTGGTGTCGCGTTCACAT AG-3'
VIB.7_F	5'-GCCAAAGATGACACGGA AGT-3'
VIB.7_R	5'-CGCTGCTTCTGTTGAGCT T-3'