ABSTRACT

Title of Document: BIOLOGY AND ECOLOGY OF *SYNECHOCOCCUS* AND THEIR VIRUSES IN THE CHESAPEAKE BAY

Kui Wang, Doctor of Philosophy, 2007

Directed By: Dr. Feng Chen, Associate Professor
University of Maryland Biotechnology Institute

The biology and ecology of *Synechococcus* and their viruses in open oceans have been studied extensively in the past decade. However, little is known about these virus-host systems in the estuarine environment. In this study, *Synechococcus* spp. isolated from the Chesapeake Bay were characterized in terms of their RuBisCO gene and ITS sequences. Chesapeake Bay harbors unique and diverse *Synechococcus* spp., which belong to a poorly defined cluster, named marine cluster B (MC-B) *Synechococcus*. This finding revived the phylogenetic position of MC-B cluster of marine *Synechococcus*. The estuarine *Synechococcus* strains can tolerate a much broader range of salinity compared to oceanic *Synechococcus* spp., suggesting the adaptation of *Synechococcus* to the dynamic estuarine ecosystem.

Seven cyanophages isolated from four MC-B *Synechococcus* strains were characterized in terms of their phenotypic and genetic traits. Among the seven MC-B *Synechococcus* phages, three are podoviruses, three are siphoviruses and one is a

myovirus. Six of seven phage isolates did not cross infect any other closely related MC-B *Synechococcus* strains, indicating the prevalence of highly specific cyanophages for MC-B strains. The podoviruses have significantly shorter latent periods compared to the myo- and siphoviruses. For the first time, photosynthetic gene (*psbA*) was found in the podoviruses infecting marine *Synechococcus*. DNA polymerase gene (*pol*) sequences were obtained from three MC-B *Synechococcus* podoviruses, and they cluster with all the known podoviruses of marine picocyanobacteria. Viral capsid assembly gene (g20) was found to be conserved among cyanomyoviruses for marine picocyanobacteria.

Synechococcus abundance often exceeded 10^6 cells ml^{-1} in summer, and sometimes contributed more than 50% of total phytoplankton biomass and primary production in the Chesapeake Bay. Cyanophage titer ranged from undetectable to over 10^5 MPN ml^{-1} in the Bay. Both *Synechococcus* abundance and their phage titers varied dramatically in different seasons, and the two co-varied on temporal and spatial scales. No synchronized seasonal succession was seen for population compositions of *Synechococcus* and cyanomyovirus, suggesting that "kill the winner" module may not apply to polyvalent cyanomyoviruses. *Synechococcus* and their viruses living in the Chesapeake Bay may develop an ecological strategy different from their oceanic counterparts.

BIOLOGY AND ECOLOGY OF *SYNECHOCOCCUS* AND THEIR VIRUSES IN THE CHESAPEAKE BAY

By

Kui Wang

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2007

Advisory Committee:
Associate Professor Feng Chen, Chair
Professor M. Robert Belas
Associate Professor Russell T. Hill
Associate Professor K. Eric Wommack
Dr. Daniel E. Terlizzi

Acknowledgements

It is difficult to overstate my deep and foremost gratitude to my Ph.D. supervisor, Dr. Feng Chen, who encouraged me to explore the world of viruses at the very beginning. His enthusiasm, inspiration and critical thinking ensured the smooth progress of my research project. He is always available when I need to discuss with him. He has allowed me a great deal of freedom to explore the topics of my interest. My dissertation has incorporated numerous remarks and comments from him. I sincerely thank my Shifu for his solid training, friendship and strong faith in me!

I gratefully thank my committee advisors, Drs. Belas, Hill, Terlizzi and Wommack for their continued scientific guidance and valuable advice during my academic pursuit. Particularly, I would like to thank Dr. Wommack for his tremendous effort for planning and coordinating numerous research cruises, as well as many insightful discussions during the cruises.

Big thanks are to my colleagues, Jinjun Kan, Ju Shen, Kate O'Mara and Erla Ornolfsottir for help with the bench work and thoughtful discussions. I am grateful to my friends at the Center of Marine Biotechnology during my six-year study. Among them, Matthew Anderson, Olivier Peraud, Naglaa Mohamed, Sonja K. Fagervold and Naomi Montalvo deserve special thanks. I thank them for sharing, caring, as well as moral support.

I thank all the "MOVErs" who participated in the "Microbial Observatory for Virioplankton Ecology" project. Special thanks are due to Wayne Coats, Kurt E. Williamson, Shannon Williamson, Bench Shellie, Rebekah R. Helton, Danielle M. Winget and Yoanna Eissler for sharing the good and tough days during the cruises. I

want to thank the crew of the R/V *Cape Henlopen* and R/V *Hugh R. Sharp* for their expertise and collecting water samples that were critical to my research project.

I thank Dr. Hill for his generosity and continued support on PFGE work. I also thank K. Czymmek and D. Powell at Delaware Biotechnology Institute (University of Delaware) for their technical supports on TEM work.

I would like to thank the many people at Center of Marine Biotechnology for helping the center to run smoothly and helping me in countless ways.

Finally, and most importantly, I am forever indebted to my family and particularly to my mother (Dingwei Wang) for their understanding, continuous support, endless patience and encouragement to me. This dissertation is dedicated to my beloved family.

Statement of Contribution

Dr. Jinjun Kan and Dr. Marcelino T. Suzuki (Chesapeake Biological Laboratory, University of Maryland) provided *Synechococcus* ITS sequence data from Chesapeake Bay bacterioplankton rRNA operon clone libraries for Chapter 3.

Dr. D. Wayne Coats (Smithsonian Environmental Research Center) provided Chl *a* data of Chesapeake Bay water samples for Chapter 6.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction: Marine *Synechococcus* and cyanophages

Unicellular cyanobacteria of the genus *Synechococcus* are ubiquitous and abundant in the oceans. They are photosynthetic microorganisms and contribute significantly to the global carbon fixation. Recently, viruses or phages that infect *Synechococcus* have been found to be abundant and concurrent with their hosts. A great deal of effort was made in the past decade to understand the ecological role of *Synechococcus* phages. This chapter reviews the historic studies and recent progress in studies of marine *Synechococcus*-phage system, followed by the research motivation and scope of this dissertation.

1.1 Ecological importance of marine *Synechococcus*

Cyanobacteria are probably the most ancient organisms on the planet Earth. They were extant ca. 3.5 billion years ago, and became dominant oxygenic phototrophs 2.5 billion years ago (Schopf 2000). In addition to their diverse morphology and nearly universal distribution in various aquatic and terrestrial environments, cyanobacteria contribute greatly to global biogeochemical cycles (Whitton and Potts 2000). The discovery of the tiny but widespread planktonic cyanobacteria *Synechococcus* (Johnson and Sieburth 1979; Waterbury et al. 1979) and *Prochlorococcus* (Chisholm et al. 1988) in marine environments, has fundamentally changed our view of the composition and function of marine ecosystems. In general, *Synechococcus* cells (0.6-2.1 μm) are slightly larger than *Prochlorococcus* cells (0.5-1.6 μm) (Herdman et al. 2001). *Synechococcus* contain

phycobilisomes as their main light-harvesting apparatus, while *Prochlorococcus* use divinyl chlorophylls a and b for light harvesting (Ting et al. 2002). *Synechococcus* can be found in a wide range of habitats including lakes, rivers, estuaries, coastal and oceanic waters, and are more diverse than *Prochlorococcus* in marine environments (Fuller et al. 2003; Ahlgren and Rocap 2006). Both *Synechococcus* and *Prochlorococcus* are key photoautotrophic organisms in the open oceans. They contribute substantially to global primary production (Platt et al. 1983; Waterbury et al. 1986; Partensky et al. 1999a). About half of the global primary production occurs in the oceans (Whitman et al. 1998), and marine picocyanobacteria contribute over 50% of the oceanic primary production (Goericke and Welschmeyer 1993; Li 1995; Liu et al. 1997a).

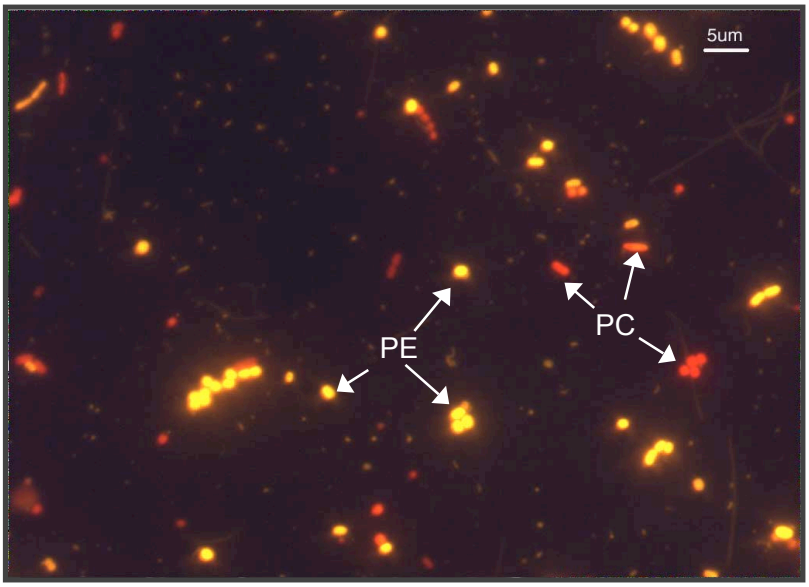


Fig. 1-1. Epifluorescence micrograph of an unfiltered Chesapeake Bay water sample (magnification, ×1,000). The water sample was collected from Stat. 804 in July 2004 (Chapter 6). Picocyanobacteria of the genus *Synechococcus* are abundant during summer months in the Chesapeake Bay. Phycocyanin-only (PC) *Synechococcus* show in red, while phycoerythrin-containing (PE) *Synechococcus* show in orange under green excitation (excitation BP 510-560, emission LP 590).

1.1.1 Taxonomy of marine *Synechococcus*

The genus *Synechococcus* is a provisional taxon that was loosely defined as unicellular coccoid to rod-shaped cyanobacteria (< 3 μm in diameter), whose cells divide by binary fission in a single plane (Waterbury and Rippka 1989). Traditionally, marine *Synechococcus* have been classified into three clusters (i.e. marine clusters A, B and C) based on their morphological, physiological and biochemical characteristics (Waterbury and Rippka 1989). Members of the marine cluster A (MC-A) group (mol % G+C = 55-62) were isolated from euphotic zone from both open oceans and coastal waters and they have elevated salt requirement for growth. All the *Synechococcus* contain phycocyanin (PC) as a light-harvesting pigment, while the MC-A strains also contain phycoerythrin (PE) pigments. MC-A strains can be further differentiated based on the presence or absence of phycourobilin (PUB), while the ratio of PUB to phycoerythrobilin (PEB) can vary in different strains (Wood et al. 1985). All strains in marine clusters B (MC-B, with mol % G+C = 63-69.5) and C (MC-C, with mol% G+C =47. 5-49.5) appear to be halotolerant and contain only phycocyanin (PC) as their major light harvesting pigment. The MC-B and MC-C *Synechococcus* were mostly isolated from coastal or estuarine waters, and much fewer strains were isolated for these two clusters compared to the MC-A cluster (Waterbury and Rippka 1989).

Classification of marine *Synechococcus* has recently been examined by molecular approaches. The MC-A cluster is well supported by phylogenetic analyses of 16S rRNA gene and 16S-23S ribosomal DNA internal transcribed spacer (ITS) sequences (Honda et al. 1999; Robertson et al. 2001; Laloui et al. 2002; Rocap et al. 2002). The MC-C cluster currently contains only four *Synechococcus* strains

according to 16S rRNA gene phylogeny (Robertson et al. 2001). In contrast, the MC-B cluster contains diverse marine *Synechococcus* strains whose relationship is poorly supported by molecular phylogeny. The MC-B cluster originally contained four *Synechococcus* strains, PCC7001, WH5701, WH8007, WH8101, all isolated from coastal waters (Waterbury and Rippka 1989). Strain PCC7001 has now been reclassified as a member of the *Cyanobium* cluster based on 16S rRNA gene phylogeny (Herdman et al. 2001). Strain WH8101 was found to be more closely related to MC-A (or *Synechococcus* subcluster 5.1) than to strain WH5701, while WH5701 was phylogenetically distinct from all other marine *Synechococcus* isolates (Rocap et al. 2002; Fuller et al. 2003; Chen et al. 2004). Strain WH8007 has rarely been included in previous phylogenetic analyses of *Synechococcus*. All evidence points to an unstable taxonomy for the MC-B members. Therefore, much more effort is needed to elucidate a clear phylogenetic position of MC-B strains. Characterization of more *Synechococcus* strains isolated from different estuarine and coastal waters may provide insights into the classification of the MC-B group.

1.1.2 Distribution and abundance of marine *Synechococcus*

Marine *Synechococcus* are widely distributed in the oceans (Waterbury et al. 1979; Li et al. 1983; Li 1998). *Synechococcus* are present in both warm and cold marine regions, as well as eutrophic and oligotrophic waters. Although present in low abundance, they can be found in the polar seas where the water temperature as low as 2°C (Shapiro and Haugen 1988). In contrast, picocyanobacteria of *Prochlorococcus* are warm species whose distribution is essentially confined between 40°N and 40°S in oligotrophic oceans (Partensky et al. 1999a). Typically, *Synechococcus* abundance

ranges from 10^3-10^4 cells ml^{-1} in euphotic zones of oceans (Waterbury et al. 1979; Partensky et al. 1999b). Geographically, their abundance was found to be low in the gyre (up to 4×10^3 cells ml^{-1}), higher in subtropical area (up to 2×10^4 cells ml^{-1}) and highest in the upwelling region (up to 10^5 cells ml^{-1}) in the open ocean (Zwirglmaier et al. 2007). The abundance of *Synechococcus* appeares to increase from oligotrophic ocean (ca. 10^3 cells ml^{-1}) to eutrophic coastal waters (up to 10^5 cells ml^{-1}) (Sullivan et al. 2003). Annual cycles of *Synechococcus* abundance in the ocean appear to be regulated by water temperature. The abundance of *Synechococcus* is high in warm seasons and low in cold seasons (Li 1998). Generally, *Synechococcus* are less abundant than *Prochlorococcus* in oligotrophic oceans but much higher in coastal waters (Partensky et al. 1999b). *Synechococcus* and *Prochlorococcus* appear to complement each other, and occupy different niches in the ocean (Chisholm 1992). Altogether, they are estimated to account for up to 80% of the primary production in some oceans (Li 1994; Liu et al. 1997a; Partensky et al. 1999a).

1.1.3 Physiological and genetic diversity of marine *Synechococcus*

Cultured isolates of *Synechococcus* exhibit a remarkable physiological difference in terms of pigment composition, response to irradiance, nutrient utilization, motility and salt requirement. (Ferris and Palenik 1998; Scanlan and West 2002). For example, many MC-A strains are capable of chromatic adaptation, by altering the ratio of phycourobilin (PUB) to phycoerythrobilin (PEB) to maximize the photosynthetic efficiency under varied light spectral availability (Wood et al. 1985; Palenik 2001). Some MC-A strains possess a unique type of motility (Waterbury et al. 1985; Toledo et al. 1999) different from the typical flagella- or pilus-driven

motility. The swimming ability in these *Synechococcus* strains appears to be driven by a unique cell-surface polypeptide encoded by the *swmA* gene (Brahamsha 1996a). Nutrient utilization also varies among *Synechococcus*. Marine *Synechococcus* strains lack N_2 fixation ability but are capable of using a wide range of inorganic (ammonium, nitrate and nitrite) as well as organic (urea and some amino acids) N sources (Waterbury et al. 1986; Collier et al. 1999; Moore et al. 2002; Palenik et al. 2003). The physiological flexibility of *Synechococcus* reflects their great potential in niche partitioning.

Genetic diversity of cultured *Synechococcus* and their natural assemblages has been investigated using several different gene markers such as 16S rRNA gene (Fuller et al. 2003), ITS region (Rocap et al. 2002), DNA-dependent RNA polymerase gene (*rpo*C1)(Toledo and Palenik 1997), and ribulose 1,5-bisphosphate carboxylase-oxygenase gene (*rbcL*) (Pichard et al. 1997). Despite the different focus of individual studies, they all points to the presence of diverse *Synechococcus* in marine environments. Seven distinguishable clades in MC-A *Synechococcus* were resolved based on the *rpo*C1 gene (Toledo and Palenik 1997; Toledo et al. 1999). Fuller et al. (2003) showed that at least 10 discrete clades were found in marine *Synechococcus* based on 16S rRNA gene phylogeny (Fuller et al. 2003). Rocap et al. (2002) identified at least 16 clusters of MC-A *Synechococcus* based on ITS sequences (Rocap et al. 2002; Ahlgren and Rocap 2006). Novel *Synechococcus* genotypes are emerging when more ecosystems are investigated, reflecting the broad niche adaptation of marine *Synechococcus* (Moore et al. 1998; Urbach et al. 1998; Honda et al. 1999; Robertson et al. 2001; Rocap et al. 2002; Scanlan and West 2002; Crosbie et

al. 2003; Ernst et al. 2003; Ferris et al. 2003; Fuller et al. 2003; Becker et al. 2004; Chen et al. 2004; Ahlgren and Rocap 2006). The high genetic diversity of *Synechococcus* has challenged traditional classification systems which are primarily based on their phenotypic features. Overall, little effort has been made to understand the diversity of *Synechococcus* living in more productive coastal estuaries, where the MC-B and MC-C members are possibly the dominant types.

1.1.4 *Synechococcus* genomes

Currently, 20 genomes of *Synechococcus* spp. have been sequenced or are in progress (Table 1-1). Sixteen of them were isolated from various marine environments; two were from freshwaters and two from Yellowstone hot springs. Among the 16 marine *Synechococcus* spp., 14 are MC-A strains, with the remaining two strains from MC-C (*Synechococcus* PCC7002), and a *Cyanobium* relative (*Synechococcus* WH5701). Strain WH5701 was thought to be a representative of the MC-B *Synechococcus*; however, our recent phylogenetic analysis placed this strain close to the *Cyanobium* cluster (Chen et al. 2006a). Although many genome sequences from marine *Synechococcus* are now available, the overwhelming bias has been towards MC-A strains.

Table 1-1. Completed and on-going *Synechococcus* genome sequencing projects (modified from NCBI).

***Synechococcus* strain**	**Source of Isolation**	**Group**	**Genome size (Mb)**	**Sequencing facility**
Completed genomes				
Synechococcus sp. PCC 6301	Freshwater	Non-MC	2.70	Nagoya University
Synechococcus sp. PCC7942	Freshwater	Non-MC	2.70	DOE-JGI
Synechococcus sp. JA-2-3B'a	Hot spring	Non-MC	3.05	TIGR
Synechococcus sp. JA-3-3Ab	Hot spring	Non-MC	2.93	TIGR
Synechococcus sp. CC9605	California current	MC-A	2.51	DOE-JGI
Synechococcus sp. CC9902	California current	MC-A	2.23	DOE-JGI
Synechococcus sp. WH8102	Open ocean	MC-A	2.43	DOE-JGI
Synechococcus sp. CC9311	California current	MC-A	2.76	TIGR
Synechococcus sp. RCC307	Mediterranean Sea	MC-A	2.22	Genoscope
Synechococcus sp. WH7803	North Atlantic ocean	MC-A	2.36	Genoscope
In progress (or in draft assembly)				
Synechococcus sp. BL107	Mediterranean Sea	MC-A	2.5	J. Craig Venter Institute
Synechococcus sp. Eum14	Tropical Atlantic	MC-A	2.5	J. Craig Venter Institute
Synechococcus sp. PCC7002	Fish pens, Puerto Rico	MC-C	3.3	Penn State University
Synechococcus sp. RS9916	Red Sea	MC-A	2.5	J. Craig Venter Institute
Synechococcus sp. RS9917	Red Sea	MC-A	2.5	J. Craig Venter Institute
Synechococcus sp. WH5701	Long Island Sound	Cyanobium	3.0	J. Craig Venter Institute
Synechococcus sp. WH7805	Atlantic ocean	MC-A	2.6	J. Craig Venter Institute
Synechococcus sp. M11.1	Gulf of Mexico	MC-A	2.5	J. Craig Venter Institute
Synechococcus sp. M16.17	Gulf of Mexico	MC-A	2.5	J. Craig Venter Institute
Synechococcus sp. MITS9220	Equatorial Pacific	MC-A	2.5	J. Craig Venter Institute

Genome sequences of marine picocyanobacteria have shed light on how these microorganisms adapt to unique ecological niches in the ocean. For example, genome sequences from two *Synechococcus* strains (Palenik et al. 2003; Palenik et al. 2006) and three *Prochlorococcus* strains (Dufresne et al. 2003; Rocap et al. 2003) have shown unique genetic features reflecting their ecological adaptation to distinct marine niches. Genome streamlining can be seen among the picocyanobacterial members. For example, high-light adapted (HL) *Prochlorococcus* appear to economize their genome size (as small as 1.6 Mbp in MED4) to in adaptation to the relatively stable conditions and low nutrient waters. The relatively large genomes of *Synechococcus* (2.43 Mbp for WH8102 and 2.76 Mbp for CC9311) tend to provide more versatility, which enable them to exploit more diverse resources. Comparative genomic analysis of an oceanic *Synechococcus* strain WH8102 and a coastal *Synechococcus* strain CC9311 also illustrates the adaptation of picocyanobacteria to particular niches. The coastal strain *Synechococcus* CC9311 contained more genes encoding the proteins for sensing and responding to variable environments compared to the oceanic strain *Synechococcus* WH8102 (Palenik et al. 2006). Genome sequencing has become a powerful tool for understanding the potential ecophysiology and evolution of picocyanobacteria. I recommend that the genomes from a few MC-B *Synechococcus* strains should be sequenced in the near future.

1.2 Cyanophages infecting marine *Synechococcus*

Bacteriophages (viruses infecting bacteria) are now known to be the most abundant biological entities on earth with approximately 4×10^{30} present in the oceans (reviewed in Suttle 2005b; Suttle 2007). The discovery of highly abundant

viral particles (ca. 10^7 per ml) in natural waters (Bergh et al. 1989; Proctor and Fuhrman 1990) re-initiated research into to the ecological roles of marine viruses.

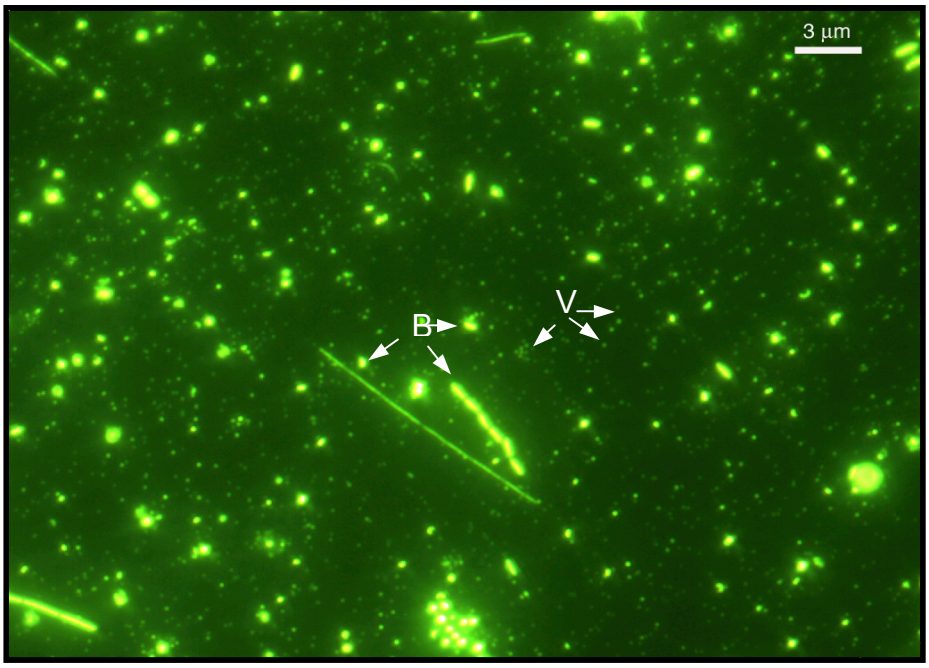


Fig. 1-2. Epifluorescence micrograph of an unfiltered Chesapeake Bay water sample stained with SYBR Gold. Bacterial cells (B) and virus-like particles (V) are visualized under blue excitation (excitation BP 485, emission LP 520) with a Zeiss Axioplan epifluorescence microscope at ×1,000 magnification (Chapter 5).

Viruses were estimated to account for 10 to 40 % of bacterial mortality in marine and freshwater environments (Fuhrman and Noble 1995; Hennes and Simon 1995), equal or even greater than average grazing-induced mortality (Weinbauer and Peduzzi 1995; Steward et al. 1996; Fuhrman 1999; Wommack and Colwell 2000). Viral lysis could affect carbon and nutrient flow through different compartments of the foodweb (Wommack and Colwell 2000), remove about 6 to 25% of primary production into dissolved organic matter (DOM) pool (Wilhelm and Suttle 1999), and consequently promote bacterial production at the community level (Fuhrman 1999).

Unlike protozoan grazing, viral infection does not have an equal effect on all bacterial species in a microbial community. The host-specific nature of viral infection can change the species richness and evenness in a bacterial community and therefore control the bacterial community composition (Fuhrman and Suttle 1993; Wommack and Colwell 2000; Weinbauer and Rassoulzadegan 2004). The selective pressure of viral infection applied on bacterial community has been incorporated into a "killing the winner" hypothesis (Thingstad and Lignell 1997). In this hypothesis, diversity levels in bacterial community is partially maintained by viruses killing the winners of inter-species resource competition (Thingstad and Lignell 1997; Thinstad 2000). Wommack and Cowell (2000) further extended this hypothesis into a conceptual model depicting the changes of bacterial and phage abundance in a microbial community. In this model, in a microbial consortium, the succession of the most abundant bacterial species is followed by the co-variation of their specific viral abundance along a time scale. From a genetic and evolutionary point of view, virus-mediated gene transfer is commonly seen in the prokaryotes and thought to be an important mechanism to sustain both host and virus diversity and fitness (Saye and Miller 1989; Chiura 1997; Paul 1999; Fuhrman 1999). It is now widely recognized that viruses are active and important component in aquatic microbial loop, with respect to their impacts on microbial mortality, production, community structure, driving microbial diversification and biogeochemical cycling (Suttle et al. 1990; Fuhrman 1999; Wommack and Colwell 2000; Weinbauer and Rassoulzadegan 2004).

1.2.1 Discovery of highly abundant marine *Synechococcus* phages

The first cyanophage was isolated from freshwater filamentous cyanobacteria in 1963 (Safferman and Morris 1963). Since then, numerous cyanophage systems have been isolated from fresh waters and characterized in order to understand the biological and ecological interactions between the these viruses and hosts (Martin and Benson 1988). Viral infection of marine picocyanobacteria became evident in early 1990s (Proctor and Fuhrman 1990; Suttle et al. 1990). Many cyanophages infecting marine *Synechococcus* were subsequently isolated and characterized (Suttle and Chan 1993; Waterbury and Valois 1993; Wilson et al. 1993; Lu et al. 2001). *Synechococcus* phages are widely distributed in fresh waters (Dorigo et al. 2004; Wilhelm et al. 2006), rivers (Lu et al. 2001), coastal waters (Suttle and Chan 1993; Waterbury and Valois 1993; Marston and Sallee 2003; Sandaa and Larsen 2006) and open oceans (Suttle and Chan 1993; Wilson et al. 1993; Suttle and Chan 1994; Sullivan et al. 2003). Reported abundance of infectious *Synechococcus* phages often exceeds 10^4-10^5 infectious units ml^{-1} in seawater (Suttle 2000), and are correlated with *Synechococcus* abundance (Waterbury and Valois 1993; Marston and Sallee 2003; Mühling et al. 2005; Sandaa and Larsen 2006).

1.2.2 Taxonomy and diversity of *Synechococcus* phages

Synechococcus phages exhibit considerably variation in morphology, genetic diversity, and host range (for reviews, see Suttle 2000 and Mann 2003). Among more than 5500 characterized bacteriophages, over 96% of them are tailed phages belonging to three viral families of double-stranded DNA viruses: Myoviridae (contractile tails); Siphoviridae (long non-contractile tails); and Podoviridae (short

tails) (Ackermann 2007). Currently, all known cyanophage isolates belong to these three well-recognized families (Suttle 2000; Mann 2003). The vast majority of characterized cyanophages have been isolated using MC-A *Synechococcus* as host cells (Suttle and Chan 1993; Waterbury and Valois 1993; Wilson et al. 1993; Lu et al. 2001). Among those MC-A *Synechococcus* phages, cyanomyoviruses are the most common phage type. Cyanomyoviruses have a broad host range, and are able to cross infect different oceanic *Synechococcus* strains (Waterbury and Valois 1993; Lu et al. 2001), or even marine *Prochlorococcus* (Sullivan et al. 2003). Only a few podoviruses and siphoviruses infecting oceanic *Synechococcus* have been reported (Suttle and Chan 1993; Waterbury and Valois 1993; Chen and Lu 2002). Interestingly, podoviruses were the only phage type isolated from high-light-adapted marine *Prochlorococcus* (Sullivan et al. 2003). In contrast to cyanomyovirus, podoviruses that infect oceanic *Synechococcus* and *Prochlorococcus* are lytic and host specific (Suttle and Chan 1993; Waterbury and Valois 1993; Chen and Lu 2002; Sullivan et al. 2003). Only a few siphoviruses infecting marine picocyanobacteria have been characterized and they tend to have a narrow host range (Suttle and Chan 1993; Sullivan et al. 2003). Distinguishable morphologies (i.e. head shape and size, length of tail and tail fibers) can be seen from different cyanophages infecting a same host, indicating the presence of diverse cyanophage in the natural environment.

In addition to the morphological diversity, genetic diversity of *Synechococcus* phages was studied by analyzing their restriction digestion patterns, protein profiles, and structural genes (Suttle 2000; Mann 2003). Restriction digestion analysis of phage DNA showed the genetic variation among cyanophage isolates (Wilson et al.

1993; Lu et al. 2001). Cross-hybridization of cyanomyovirus restriction digested DNA revealed a conserved region that contains a gene homologous to the g20 gene of coliphage T4 (Fuller et al. 1998). The g20 gene encodes portal vertex protein involved in capsid assembly and viral genome packaging in coliphage T4 (Rao et al. 1992). Characterization of the g20 gene from three different *Synechococcus* myoviruses permitted the design of specific PCR primers for the cyanomyovirus group (Fuller et al. 1998). Several sets of g20 gene primers have been developed and used to investigate the genetic diversity of cyanophage isolates and their natural assemblages (Table 1-2).

Table 1-2. Comparison of different PCR primer sets used for targeting the g20 gene in natural aquatic environments.

Primer set	Degeneracy	Amplicon Size (bp)	References	Applications	Ecosystem tested
CPS1/CPS2	4/8	165	Fuller et al. 1998	PCR and Competitive PCR	Coastal waters
CPS4/CPS5	0/0	165	Wilson et al. 1999	DGGE	Atlantic Ocean
			Wilson et al. 2000	DGGE	Atlantic Ocean
			Frederickson et al. 2003	DGGE	River estuaries
			Mühling et al. 2005	DGGE	Gulf of Aqaba
CPS1/CPS8	4/12	592	Zhong et al. 2002	Clone library	Estuarine, coastal and oceanic waters
			Marston and Sallee 2003	RFLP (isolates)	Coastal waters (Rhode Island)
			Wang and Chen 2004	T-RFLP	Estuary (Chesapeake Bay)
			Dorigo et al. 2004	Clone library	Freshwater (Lake Bourget)
			Wilhelm et al, 2006	Clone library	Freshwater (Lake Erie)
CPS4/G20-2	0/1024	592	Short and Suttle 2005	DGGE	Gulf of Mexico, Pacific Ocean, Arctic, lakes, catfish pond, Chuckchi Sea (3000m deep sea)

High genetic diversity of g20 gene sequences has been found in various aquatic environments (Wilson et al. 1999; Zhong et al. 2002; Frederickson et al. 2003; Dorigo et al. 2004; Wang and Chen 2004; Mühling et al. 2005; Short and Suttle 2005; Sandaa and Larsen 2006; Wilhelm et al. 2006). Zhong et al. (2002) compared the g20 gene diversity between coastal and oceanic waters. Phylogenetic analysis of 114 gp20 homologs recovered from six coastal and oceanic water samples revealed nine distinct phylogenetic clades. In one water sample, six different phylogenetic clades and 29 different genotypes could be identified based on g20 gene sequences (Zhong et al. 2002). At least three main phylogenetic groups were seen among cultured cyanomyoviruses (Zhong et al. 2002; Marston and Sallee 2003). Gene g20 sequence diversity also varied from coastal to oceanic waters and from surface water to the deep chlorophyll maximum depth, but there was no correlation of g20 phylogenetic groups with geographic location (Zhong et al. 2002). Such observations have been further extended and supported by later studies (Short and Suttle 2005; Wilhelm et al., 2006). Notably, it was found that many of g20 sequences recovered from natural environments were not closely related to known cyanomyoviruses infecting oceanic *Synechococcus* (Zhong et al. 2002; Short and Suttle 2005; Wilhelm et al. 2006). It was speculated that these unidentified g20 sequences could either arise from phages infecting *Prochlorococcus* (a close relative of *Synechococcus*), or from phages infecting other groups of bacteria (Zhong et al. 2002; Short and Suttle 2005; Wilhelm et al. 2006). The origin of those unmatched environmental g20 sequences is still mysterious.

Rapid molecular fingerprinting approaches such as denaturing gradient gel electrophoresis (DGGE) and terminal restriction fragment length polymorphism (T-RFLP) have been used to investigate the composition of cyanophage in nature. DGGE analysis of 165-bp DNA fragments of g20 genes amplified by original cyanomyovirus specific PCR primers (CPS4 and CPS5) was first used to examine the structure of natural cyanophage populations in the Atlantic Ocean (Wilson et al. 1999; Wilson et al. 2000). High genetic diversity of cyanophage was found in these studies and significant changes in the population structure was observed from surface to deep water. Maximum diversity was correlated to maximum *Synechococcus* abundance in a south-north transect of Atlantic Ocean (Wilson et al. 1999; Wilson et al. 2000). T-RFLP analysis of the g20 amplicons from Chesapeake Bay revealed strong temporal and spatial variations with the cyanomyoviruses (Wang and Chen 2004). In general, population dynamics of cyanomyoviruses appears to be influenced by *Synechococcus* abundance and population composition (Wilson et al. 1999; Frederickson et al. 2003; Wang and Chen 2004; Mühling et al. 2005).

The DNA polymerase (*pol*) gene was demonstrated to be a suitable genetic marker for examining the diversity and evolutionary relationship between algal viruses and other large double-stranded DNA viruses (Chen and Suttle 1995, 1996; Chen et al. 1996). For cyanophage, the T7-like *pol* gene was first found in a podovirus (P60) infecting marine *Synechococcus* (Chen and Lu 2002). Despite the great genomic divergence and mosaic nature of viruses (Hendrix et al. 1999; Hendrix et al. 2000), a small set of key genes (e.g. replication genes) appear to be conserved among certain groups of viruses. Podoviruses are mostly lytic and can be very

virulent. For example, a few cyanopodovirus P60 can lyse their host bacterial lawn within 10 days (Fig. 1-3). Podoviruses possess the DNA *pol* gene in their genome and this could be important for rapid replication of their DNA (Chen and Lu 2002). Recently, the DNA *pol* gene was also found in two podoviruses infecting *Prochlorococcus* (Sullivan et al. 2005) and another oceanic *Synechococcus* strain (Pope et al. 2007). These findings further indicate that the DNA *pol* gene is conserved among T7-like podovirus of marine picocyanobacteria. It is possible that the genetic diversity of cyanopodoviruses could be explored using molecular approaches based on T7-like *pol* gene in the near future.

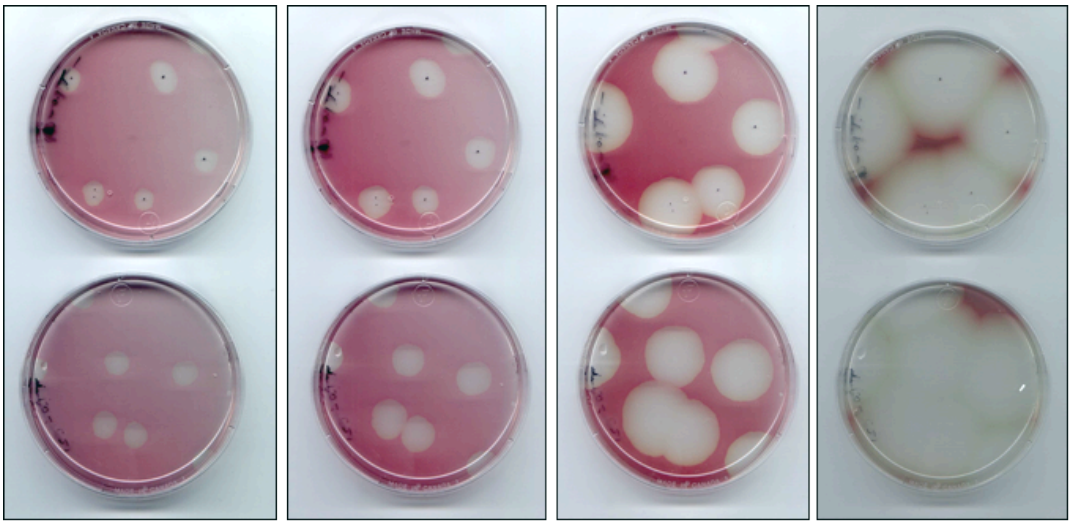


Fig. 1-3. Progression of cyanopodovirus P60 growing on the host *Synechococcus* WH7805 lawn. Plaque assay followed the method described in Chapter 4.

Only a limited number of cyanosiphoviruses has been reported to date. Our knowledge about genes and genetic characteristics for this group of cyanophage is still lacking. Further studies are warranted to understand the biology and ecology of cyanosiphoviruses.

1.2.3 Cyanophage genomes

The first complete cyanophage genome (podovirus P60) was reported in 2002 (Chen and Lu 2002). The DNA replication module of P60 was similar to that of coliphage T7 and a podovirus (SIO1) infecting a marine roseobacter (Rohwer et al. 2000; Chen and Lu 2002), suggesting DNA replication genes could be essential for this group of lytic phages. A high sequence homology in the regions encoding exonuclease, ribonucleoside triphosphate reductase and thymidylate synthase was found between marine cyanopodovirus P60 and marine *Synechococcus* WH8102, suggesting a horizontal gene transfer between the cyanophage and host cyanobacteria (Chen and Lu 2002). Currently, the full genome sequences from four *Synechococcus* phages and three *Prochlorococcus* phages have been documented (Table 1-3). Genome sequencing of cyanophages has provided many new insights into ecological, biological and evolutionary interaction between viruses and their hosts. Perhaps the most intriguing finding is the possession of key photosynthetic genes by cyanophages. Many marine cyanophages contain the photosynthetic genes *psbA* and *psbD*, which functionally encode photosystem II core protein D1 and D2 of their hosts (Mann et al. 2003; Lindell et al. 2004; Millard et al. 2004; Sullivan et al. 2006). The cyanophage-encoded *psbA* gene is expressed upon infection (Lindell et al. 2005; Clokie et al. 2006), implying that possession of this gene might benefit the viral

fitness (Clokie and Mann 2006). A recent survey also revealed that diverse virus-like *psbA* gene is present in marine (Sharon et al. 2007) and estuarine environments (Bench et al. 2007), and the expression of viral *psbA* can be detected in the sea (Sharon et al. 2007). More recently, genome-wide expression of cyanophage-encoded host metabolism genes (*psbA, hli, talC and nrd*) upon infection revealed the co-evolution features between marine virus and host (Lindell et al. 2007). It is believed that the growing number of phage genomes will provide new insights into the phage-host interaction and co-evolution.

Table 1-3. Sequenced marine cyanophage genomes.

Phage	Genome Size (bp)	Family	Host strain	Accession #	References
P60	47,872	Podoviridae	*Synechococcus* WH7805	AF338467	Chen and Lu 2002
S-PM2	196,280	Myoviridae	*Synechococcus* WH7803	AJ630128	Mann et al. 2005
P-SSP7	44,970	Podoviridae	*Prochlorococcus* MED4	AY939843	Sullivan et al. 2005
P-SSM2	252,401	Myoviridae	*Prochlorococcus* NATL1A	AY939844	Sullivan et al. 2005
P-SSM4	178,249	Myoviridae	*Prochlorococcus* NATL2A	AY940168	Sullivan et al. 2005
Syn5	46,214	Podoviridae	*Synechococcus* WH8109	EF372997	Pope et al. 2007
Syn9	176,847	Myoviridae	*Synechococcus* WH8012	DQ149023	Weigele et al. 2007

1.2.4 Impacts of cyanophages on *Synechococcus* in natural environments

Cyanophage titers are often found to be correlated with their host *Synechococcus* abundance in seawater. *Synechococcus* phage abundance is usually highest when host cell density reach the seasonal maximum (Waterbury and Valois 1993; Marston and Sallee 2003; Mühling et al. 2005). Cyanophage titers decrease

from near shore waters to open oceans, along with host cell concentration (Suttle and Chan 1994; Sullivan et al. 2003).

Viral lysis of specific host leads to microbial mortality, and therefore could affect microbial population composition. About 3% of *Synechococcus* cells in natural communities have been documented to be visibly infected (Proctor and Fuhrman 1990), while up to 6.6% of *Synechococcus* cells were lysed daily in coastal and offshore environments as estimated based on virus-host contact rates and viral decay rates (Waterbury and Valois 1993; Suttle 1994; Suttle and Chan 1994). Recent metagenomic survey also estimated that up to 12% of natural picocyanobacteria of *Synechococcus* and *Prochlorococcus* were infected by cyanophages (DeLong et al. 2006). Cell lysis mediated by viral infection could ultimately reduce the transfer of primary production to higher trophic levels (Fuhrman 1992; Wilhelm and Suttle 1999). In addition to the biomass control of host populations, there is evidence that cyanophage can influence the population succession of *Synechococcus* in the oligotrophic ocean (Mühling et al. 2005). Because cyanophage are abundant, diverse and active, they have the potential to alter their host *Synechococcus* population structure and influence the genomic evolution of *Synechococcus* via genetic recombination (Hambly et al. 2001; McDaniel et al. 2002; Ortmann et al. 2002).

Prophage-like genes have been found in the *Synechococcus* WH8102 genome (Palenik et al. 2003). However, no intact prophage genome has been identified on the genome sequences of marine *Synechococcus* spp. (Palenik et al. 2003; Palenik et al. 2006). Attempts to induce lysogeny in many MC-A *Synechococcus* strains were not successful (Waterbury and Valois 1993). On the other hand, lysogeny of

Synechococcus was detected in the natural waters (McDaniel et al. 2002; McDaniel and Paul 2005; Hewson and Fuhrman 2007). The lysogenic potential of marine *Synechococcus* is still controversial at this point.

1.3 Bacteria and viruses in the Chesapeake Bay

Chesapeake Bay is the largest estuary in the United States. It is an extremely complex temperate ecosystem in which temperature; salinity and nutrients vary greatly, forming strong environmental gradients along the Bay. Salinity varies from 0 to 30 ppt from the upper to lower Bay, while water temperature reaches the freezing point in winter, and 30 °C in summer. The Chesapeake Bay estuary is subject to significant influence of freshwater from rivers. Susquehanna River on the northern end of the Bay contributes over 50% of total freshwater input (USGS, United States Geological Survey) and accounts for ca. 70% of total nitrogen and 60% of total phosphorous loads into the Bay (Malone 1992). The excess nutrient input subsequently triggers algal blooms in the spring, and decomposition of massive algal biomass by bacterial activities often leads to severe anoxia condition in the bottom water of the Bay (Officer et al. 1984).

The Chesapeake Bay is a well-characterized estuarine ecosystem, early studies have been focused on the bacterial biomass, growth rate, respiration, production and relationship between bacteria and nutrients (Jonas *et al.*, 1988; Jonas and Tuttle 1990; Horrigan *et al.*,1990; Shiah and Ducklow 1994). Bacteria are abundant (average 3×10^6 cells ml^{-1}) and dynamic microbial components in the Chesapeake Bay (Wommack et al. 1992). High bacterial biomass and production rates are often observed in middle Cheapeake Bay region (Ducklow and Shiah 1993; Shiah and

Ducklow 1994; Ducklow et al. 1999), and are likely controlled by grazing or virus-related mortality (Smith and Kemp 2003). Bacterial growth rate appears to be limited by organic carbon in northern Bay, but by inorganic nutrients in southern Bay (Smith and Kemp 2003).

Early studies showed that bacterial community composition varies both seasonally and spatially in the Chesapeake Bay (Bidle and Fletcher 1995; Noble et al. 1997). For instance, bacteria of *Cytophaga* group were found predominant at the turbidity maximum in the middle Bay (Bouvier and del Giorgio 2002). The distribution and abundance of *Gamma*-proteobacteria and *Vibrio* spp. varied seasonally along the Choptank River tributary (Heidelberg et al. 2002). In northern Chesapeake Bay, where the salinity is low, *Vibrio cholerae* was more frequently detected during the warmer months (Louis et al. 2003). Most recent analysis based on bacterial rRNA operons (16S rRNA-ITS-23S rRNA) clone libraries revealed that the Chesapeake Bay contains diverse bacterial groups and the bacterial community composition varies dramatically between seasons (Kan et al. 2007). *Alpha*-proteobacteria, *Gamma*-proteobacteria, *Bacteroidetes*, *Cyanobacteria* and *Actinobacteria* were the dominant major groups found in warm season (September), while *Alpha*-proteobacteria, *Beta*-proteobacteria, and *Actinobacteria* were most abundant in wintertime (March) (Kan et al. 2007). *Roseobacter* and *Rhodobacter*-like sequences are predominant in the March clone library. In contrast, *Synechococcus*-like cyanobacterial sequences are more abundant in September but not detected in March. Chesapeake Bay has a long residence time with average of seven months (Nixon et al. 1996), which allows for the development of an indigenous bacterial

community (Kan et al. 2006). For example, novel groups of SAR11, *Roseobacters*, SAR86, and *Actinobacteria* were found to be endemic to the Bay (Kan et al. 2007)., Drastic seasonal shifts in bacterial species composition were observed for SAR11, *Roseobacters, Cyanobacteria* (*Synechococcus*) and *Actinobacteria*, suggesting that a strong seasonal variation exits within these diverse bacterial groups (Kan et al. 2006).

Picocyanobacteria are important to the ecological function of the Chesapeake Bay. Picophytoplankton (dominated by *Synechococcus* spp.) can contribute 10 to 14% of the total primary production in early summer (Ray et al. 1989). In late summer, *Synechococcus* can reach nearly10^6 cells ml^{-1} and account for 55.6% of primary production in the southern Bay (Affronti and Marshall 1994). Picocyanobacterial primary production often exceeds heterotrophic bacterial productivity (Malone et al. 1991) during summer months in the Bay. The composition of *Synechococcus* phenotypes (phycoerythrin-rich vs. phycocyanin-rich) appears to correlate with the salinity gradient (Affronti and Marshall 1993). The earlier studies on *Synechococcus* surveyed only the sub-estuary regions of Chesapeake Bay (i.e. Chop-Pax transect and York River). Moreover, the species composition of *Synechococcus* in the Chesapeake Bay remains largely unexplored.

Past studies demonstrated that viruses are also abundant (ranging from 2.6×10^6 to 1.4×10^8 ml) in the Chesapeake Bay (Wommack et al. 1992). A strong correlation between bacterial and viral abundance has been observed in the Chesapeake Bay (Wommack et al. 1992). Viral abundance is usually 3-25 fold higher than that of bacteria in the Bay (Wommack et al. 1992). PFGE analysis of the Chesapeake Bay virioplankton revealed that annual variation in viral community

structure was correlated with time, geographical location and extent of water column stratification (Wommack et al. 1999b). Further hybridization analysis of spatial and temporal dynamics of specific viruses supported the idea that virus may control the bacterial community composition (Wommack et al. 1999a).

Virus and bacteria communities in the Chesapeake Bay are abundant, complex and dynamic, but little is known about how viruses interact bacteria interact with their viruses. There have been no detailed investigations on the effects of virus on a particular group of bacteria in the Chesapeake Bay. Given the importance of picocyanobacteria in the Chesapeake Bay ecosystem, it is necessary to investigate how viruses affect this particular group of bacteria in the Chesapeake Bay.

1.4 Objectives of the dissertation

Current knowledge on the interaction between *Synechococcus* and cyanophages is largely built on what has been learned from the open ocean ecosystem. Little effort has been made to understand the biology and ecology of *Synechococcus* and their viruses in the estuarine environment. Currently, only a few *Synechococcus* strains have been isolated from estuaries, and virtually no cyanophages have been isolated from estuarine *Synechococcus*. After reviewing the literature, several questions appear to deserve further investigation: 1) Does the Chesapeake Bay harbor its own *Synechococcus* species? 2) If so, are they more closely related to freshwater or marine *Synechococcus*? 3) Can we isolate viruses infectious for estuarine *Synechococcus*? 4) How diverse are *Synechococcus* and their viruses in the Bay? 5) What are the seasonal and spatial patterns of *Synechococcus* and their phage populations in the

Bay, and what factors influence their distributions? 6) Are viruses important to host population dynamics in the Bay?

The Microbial Observatory for Virioplankton Ecology in the Chesapeake Bay (MOVE project, from September 2002 to February 2007) provided me the opportunity to address the above questions.

Chapter 2: Phylogenetic diversity of *Synechococcus* in the Chesapeake Bay revealed by ribulose-1,5-bisphophate carboxylase-oxygenase (RuBisCO) large subunit gene (*rbcL*) sequences

Abstract

In order to understand how *Synechococcus* in the estuarine environment of the Chesapeake Bay are phylogenetically related to other known marine *Synechococcus*, partial *rbcL* gene sequences from 25 strains of *Synechococcus* spp. isolated from estuarine, coastal and oceanic waters were sequenced. The *rbcL* gene phylogeny showed that Chesapeake Bay *Synechococcus* isolates, together with other marine *Synechococcus* spp. formed a monophyletic group and belongs to the form IA RuBisCO. All the Chesapeake Bay *Synechococcus* were able to grow in a wide range of salinity (0-30 ‰), and most of them belong to the Marine Cluster B (MC-B). Interestingly, several phycoerythrin (PE)-containing *Synechococcus* isolated from the Bay were clustered in the MC-B group, which previously contained only non-PE *Synechococcus*. A set of PCR primers was developed to specifically amplify the *rbcL* gene from natural *Synechococcus* populations. After screening 232 clones randomly selected from five clone libraries (built on five estuarine samples respectively), at least seven different *rbcL* genotypes or the operational taxonomic units (OTUs) were identified. Despite the great genetic diversity among the OTU sequences, they were all clustered with 13 Chesapeake Bay isolates. The distribution frequency of these

OTUs varied dramatically from the upper to lower Bay. Our results suggest that the Chesapeake estuary provides an ideal environment for the MC-B type *Synechococcus* populations to thrive. Marine *Synechococcus* appear to adapt well to various ecological niches, and a clear boundary solely based on phenotypic features may not exist when more and more novel ecotypes and genotypes are unveiled with molecular tools.

Introduction

Phycobilisome-containing unicellular cyanobacteria of the genus *Synechococcus* are among the most abundant members of the picophytoplankton in the open ocean (Johnson and Sieburth 1979; Waterbury et al. 1979). Marine *Synechococcus* species contribute 5 to 30% of primary production in the surface waters of the world's oceans (Waterbury et al. 1979). The *Synechococcus* group (Chroococcales) is a provisional assemblage that was loosely defined as unicellular coccoid to rod-shaped cyanobacteria (< 3 μm in diameter) whose cells divided by binary fission in a single plane (Waterbury and Rippka 1989). Based on morphological, physiological and chemical properties, and DNA base ratios, *Synechococcus* spp. have been divided into six clusters: Cyanobacterium, *Synechococcus*, Cyanobium, Marine Clusters (MC) A, B and C (Waterbury and Rippka 1989). It is now known that MC-A, MC-B, the Cyanobium cluster, and the *Prochlorococcus* group formed a monophyletic group, distinct from freshwater *Synechococcus* and the MC-C cluster (Fuller et al. 2003). All strains in the MC-A cluster contain phycoerythrin (PE) as their major light-harvesting pigment. PE-containing *Synechococcus* strains can be distinguished by the presence or absence of

phycourobilin (PUB) and the ratio of PUB to phycoerythrobilin (PEB) contained in the individual cells. All strains in MC-B and C contain only phycocyanin as their major light harvesting pigment, but do not contain PE (Waterbury and Rippka 1989). The PE type *Synechococcus* is a major primary producer in pelagic oceans (Campbell et al. 1983; Campbell and Iturriaga 1988; Olson et al. 1990), and the phylogenetic relationship among PE isolates has been examined extensively (Wood and Townsend 1990; Urbach et al. 1992; Toledo and Palenik 1997; Urbach et al. 1998; Toledo et al. 1999; Rocap et al. 2002; Fuller et al. 2003). The divergence of PE strains based on 16S rDNA sequences appeared to be narrow (< 4%). The term "microdiversity" has been used to describe a narrow genetic niche for unicellular cyanobacteria including *Synechococcus* and *Prochlorococcus* (Fuhrman and Campbell 1998; Moore et al. 1998). The RNA polymerase gene (*rpoC1*) has been used to study the phylogenetic linkage among the PE type *Synechococcus* (Toledo and Palenik 1997; Toledo et al. 1999). The *rpoC1* based phylogeny appeared to correlate well with the motility of PE *Synechococcus*. Most recently, the phylogenetic study based on the 16S-23S ribosomal DNA internal transcribed spacer (ITS) sequences divided the MC-A *Synechococcus* into six clades (Rocap et al. 2002).

In order to further resolve the phylogeny of the MC-B strains, more *Synechococcus* strains isolated from coastal estuaries are needed. Compared with other *Synechococcus* clusters, MC-B is not a well-characterized group because it contains only four strains (WH8007, WH8101, WH5701, and PCC7001) (Waterbury and Rippka 1989). Recently, strains WH8101 and PCC7001 have been re-classified into MC-A and the *Cyanobium* cluster, respectively (Herdman et al. 2001). Although

marine *Synechococcus* and *Prochlorococcus* frequently co-exist in oceanic waters, the former dominates in coastal waters (Partensky et al. 1999b). Chromatic adaptation by *Synechococcus* enables this group to adapt to broad ecological niches; a number of different genotypes or ecotypes of *Synechococcus* have been identified in natural environments (Palenik 2001; Ferris et al. 2003). It is now known that *Synechococcus* are also abundant and dynamic members of the plankton in coastal estuaries. For example, the abundance of *Synechococcus* in the Bedford Basin varied from a few cells to 4.5×10^5 cells ml^{-1}, and was greatly influenced by temperature (Li 1998). The seasonal survey in the Chesapeake Bay showed that *Synechococcus* density reached the maximum of 3.3×10^6 cells ml^{-1} in the summer and dropped to few hundred cells ml^{-1} in the winter (Chapter 6). In a recent review, Scanlan and West (2002) indicated that *Synechococcus* from estuarine or coastal waters have not been well studied so far and are likely under-represented in *Synechococcus* culture collections and within phylogenetic trees as well. Is it possible that *Synechococcus* in coastal estuaries exhibit greater genetic diversity compared to the PE type due to their complex habitats?

Ribulose 1,5-bisphosphate carboxylase-oxygenase (RuBisCO) is a key enzyme in the primary pathway of photosynthetic carbon fixation of all primary producers. RuBisCO catalyzes the assimilation of CO_2 to organic carbon via the Calvin-Benson cycle. The enzyme is made of large and small subunits. The large subunit is responsible for carbon fixation (Miziorko and Lorimer. 1983). This enzyme has been found to have two natural forms, Forms I and II. Form I contains eight large and eight small subunits (L_8S_8), encoded by *rbcL* and *rbcS*, respectively. Form II is comprised

of large subunits only, usually as a dimer (L_2) (Watson and Tabita. 1997). The form I RuBisCOs are believed to have evolved in response to the decline of CO_2 and the emergence of oxygen while the global atmosphere changed (McFadden and Tabita. 1974). According to the phylogeny of the *rbcL* gene, the form I RuBisCOs can be divided into four different forms (A to D). The form IA consists of two marine cyanobacteria (marine *Synechococcus* WH7803 and *Prochlorococcus marinus*) and chemolithoautotrophs such as *Chromatium vinosum*, *Thiobacillus denitrificans*, and *Thiobacillus ferrooxidans* (Watson and Tabita. 1997). Form IB contains chloroplasts of terrestrial plants, green algae and cyanobacteria (e.g. *Anabaen*a, freshwater *Synechococcus* and *Synechocystis*, *Prochloron*). The form IC has been found in a few alpha- and delta-proteobacteria, and many non-green algae carry the from ID RuBisCO (Watson and Tabita. 1997). Unlike the ribosomal RNA genes which are universal for all living organisms, RuBisCO (the *rbcL* gene) is found primarily in photoautotrophs and chemolithotrophs. Because RuBisCO plays a central role in photosynthesis, the phylogenetic analyses based on the *rbcL* gene sequences have provided many new insights on the evolution of photosynthesis (Clegg 1993). The *rbcL* gene has been used as a phylogenetic marker to investigate the genetic diversity of marine phytoplankton communities (Pichard et al. 1997) and microorganisms living in the deep-sea (Elsaied and Naganuma 2001), and groundwater (Alfreider et al. 2003).

The objective of this study was to understand the phylogenetic diversity of *Synechococcus* living in the Chesapeake Bay, the largest and probably the most complex estuary in the United States. The partial *rbcL* sequences from 13 Chesapeake

Bay *Synechococcus* isolates and 12 Woods Hole *Synechococcus* isolates were sequenced and used for phylogenetic reconstruction. The *rbcL* gene phylogeny revealed that the vast majority of Chesapeake Bay *Synechococcus* isolates belong to the MC-B group. We also developed a new set of PCR primers specific to the *rbcL* gene of marine *Synechococcus* as a tool for exploring the genetic diversity of *Synechococcus* populations in the Bay. The environmental *rbcL* sequences showed that *Synechococcus* living in the Bay are diverse, but still closely related to Chesapeake Bay *Synechococcus* isolates.

Material and Methods

Synechococcus strains

Synechococcus strains used in this study were described in Table 2-1. The Chesapeake Bay *Synechococcus* spp. (with CB prefix) were isolated in our laboratory. To isolate CB strains, water samples collected from the Bay were pre-incubated to increase *Synechococcus* density (Waterbury and Willey 1988). A pour plating method (Brahamsha 1996b) was used to isolate *Synechococcus* strains from the pre-incubated waters. Marine *Synechococcus* strains were grown in the SN medium (Waterbury et al. 1986) at 26°C in constant light (20-30 μE m^{-2} s^{-1}) in an illuminated incubator. Salinity of SN media was adjusted to be close to the salinity of water sample. Cultures of marine *Synechococcus* isolates WH7803, WH7805, WH8101 and WH8103 were provided by B. Binder at the University of Georgia and the other WH strains from L. Campbell's laboratory. Strains WH8007 and WH5701 were purchased from the Center for Culture of Marine Phytoplankton (CCMP) at Bigelow Laboratory.

In vivo pigment absorption spectra

In vivo pigment absorption spectra for the *Synechococcus* strains were measured using a Beckman DU640 spectrophotometer. Cultures (2 ml) under exponential growth were scanned from 400 nm to 800 nm with a 1 nm interval.

Enumeration of *Synechococcus* cells

To determine the *Synechococcus* abundance in a water sample, 10 ml water sample was filtered onto a 0.2 μm pore size 25 mm black polycarbonate membrane filter (Osmonics, CA). A Nikon Eclipse E400 (Nikon, Japan) epifluorescence microscope was used to count the total *Synechococcus* cells, PC and PE type cells, respectively. The total cells were enumerated under the green light excitation (528-553 nm). Under the blue light excitation (465-495 nm), the PE cells showed bright yellow-orange color while the PC cells showed dim red in the same field. The PE cells were counted under the blue excitation. The difference between total counts and PE cell counts yielded the PC cell counts. At least 20 fields with the minimal of 200 cells per water sample were counted.

Oligonucleotide primers

When we started this project in 1998, there was only one complete *rbcL* gene sequence available for marine *Synechococcus* (WH7803) in GenBank (Watson and Tabita. 1996). Although PCR primers became available to amplify a 480-483 bp fragment of *rbcL* gene in 1997 (Pichard et al. 1997), the primers are degenerate and not specific for marine *Synechococcus*. A reverse degenerate PCR primer that is specific for the Form IA *Synechococcus* was published in 2003 (Wawrik et al. 2003). Our goal was to design a set of non-degenerate primers specific for marine

Synechococcus, which can be extended for DGGE or T-RFLP analysis of natural marine *Synechococcus* populations.

In this study, two primer sets were used to amplify the *rbcL* gene from various marine *Synechococcus* isolates, and a third primer set was developed for environmental *Synechococcus*. The first primer set (WHF1 and WHR1) was used to amplify the *rbcL* gene from all the PE strains. This primer set was designed based on aligned nucleotide sequences of the *rbcL* genes of all the freshwater cyanobacteria and WH7803. The WHF1/WHR1 are non-degenerate primers and specific for WH7803 strain. The sequences of WHF1 and WHR1 are 5' ATGAGCAAGAAGTACGA 3' (sequence position of 1-17 as for WH7803) and 5' GGTCTCCTGCTCGGACAG 3' (sequence position of 655-672 as for WH7803), respectively. The *rbcL* gene from all the PE strains in Table 2-1 were amplified with WHF1/WHR1 and sequenced afterwards.

The second set of *rbcL* primers (CF1 and CR1) was designed based on the conserved regions on the *rbcL* sequences of 10 known PE strains and other cyanobacterial strains. This primer set was used to amplify the *rbcL* gene from PC strains that could not be amplified with WHF1 /WHR1. The sequences for CF1 and CR1 are 5' AC(TC)TACT(GA)GACTCC(TC)GATTAC 3' (sequence position of 43-63 as for WH7803) and 5' GAA(GA)GGCTG(GA)GAGTT(GA)ATGTT 3' (sequence position of 589-609 as for WH7803). The *rbcL* gene from all the PC strains was amplified using CF1/CR1.

The third primer set MSF1 and MSR1 was designed to specifically amplify the *rbcL* gene from the Marine *Synechococcus* Clade (in Form IA, Fig. 2-4). The

primer MSF1 contains nucleotide sequence 5' GGTCCACTGTGTGGTCCGAGG 3', which corresponds to the sequence position of 88-108 as for WH7803. The primer MSR1 has nucleotide sequence 5' GTTCTCGTCGTCCTTGGTGAAGTC 3', which corresponds to the sequence position of 568-591 as for WH7803. The non-degenerate MSF1/MSR1 primers were tested against the natural bacterial communities in the Chesapeake Bay, with the goal to develop rapid fingerprinting techniques like DGGE for monitoring *Synechococcus* populations in the natural marine ecosystem.

Field sample collection

Water samples were collected from 2 m depth water of Chesapeake Bay and other estuarine stations (Fig. 2-1) using Niskin bottles on board the R/V *Cape Henlopen* on September 26-30, 2002. Water samples from the Roosevelt Inlet, Lewis, Delaware were collected from surface water at the dock using a plastic bucket. For each station, 250 ml water was filtered through 0.2 μm pore-size polycarbonate filters (47 mm diameter, Millipore,) and microbes retained on the filters were stored at -20˚C until DNA extraction.

Nucleic acid extraction

Total DNA was extracted according to a protocol developed by Schmidt et al. (1991) with minor modifications. Briefly, DNA from bacterial community was extracted by treating with lysozyme and proteinase K concomitant with phenol extraction and isopropanol precipitation. DNA were dissolved in ddH_2O and stored at -20°C for further analysis.

PCR amplification

For *Synechococcus* isolates, DNA released from boiled cultures was used as templates. For preparation of cell lysis, *Synechococcus* cultures were harvested by centrifugation at 10,000 rpm for 5 min. Culture pellets were re-suspended in 50 μl of lysis solution (0.1 M Tris-HCl, pH8.0, 0.01% P-40) and boiled for 5 minutes. PCR was conducted in 0.2 ml tubes with a final volume of 25 μl reaction mixture, which contained 1 unit of Taq DNA polymerase (Promega, Madison, WI), 200 μM of each dNTP, 1.5 mM of $MgCl_2$, 1 μM each oligonucleotide primer and 1 μl of cell lysates. For the primer set WHF1 and WHR1, the PCR amplification profile consisted of an initial denaturation at 94°C for 3 min, followed by 30 cycles of 94°C for 1 min, 50°C for 1 min, and 72°C for 1 min. The PCR condition for the primer set CF1 and CR1 was similar to WHF1/WHR1 except that the annealing temperature for CF1/CR1 was increased to 55°C.

Primer set MSF1/MSR1 was used to amplify the *rbcL* gene from the environmental DNAs. The PCR program for primers MSF1/MSR1 included an initial denaturation at 94°C for 5 min and 10 touchdown cycles of denaturation at 94°C for 1 min, annealing at 65°C (with temperature decreasing 1 °C per cycle) for 1 min, and extension at 72 °C for 1 min, followed by 25 cycles of 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min. The final extension of PCR amplification was at 72°C for 10 min. PCR amplicons was examined by electrophoresis in 2% agarose gel stained with ethidium bromide in 0.5 × TBE buffer. Gel image was captured and analyzed using Kodak EDAS 290 electrophoresis documentation and analysis system (Eastman Kodak Company, New Heaven, CT).

Clone library construction

The five PCR products amplified from environmental samples were cloned. Each PCR product was purified by using Wizard PCR Prep DNA Purification System (Promega). The purified DNA fragments were cloned into the pGEM-T Easy cloning vector (Promega) and then transformed into JM109 competent cells (Promega) according to the manufacturer's instructions. Positive clones (white-colonies) were picked randomly and transferred onto a new agar plate for further use.

Restriction fragment length polymorphism (RFLP) analysis

To avoid redundant sequencing, the *rbcL* clones amplified from five estuarine microbial communities were pre-screened with RFLP. About 50-60 positive clones from each clone library were randomly picked and the plasmid inserts were PCR-amplified with vector-specific primers T7 (5'-TAATACGACTCACTATAGGGC GA-3') and SP6 (5'-ATTTAGGTGACACTATAG-3') along with Taq DNA polymerase (Promega). PCR amplification cycles involved a 3-min initial denaturation at 94°C and followed by 35 cycles of 94°C for 1 min, 56°C for 1 min, 72°C for 2 min, and a 5-min final extension at 72°C. Once a clone was confirmed to contain an insert of the appropriate size (ca. 510 bp), the insert was subjected to RFLP analysis. A subsample (10 μl) from each PCR mixture was digested with 5 U *Hae*III in 1 × Buffer C (Promega) at 37°C overnight in a final reaction volume of 25 μl. Digested DNA was separated by 2% agarose gel electrophoresis as described above. The resulting RFLP patterns were examined and compared by visual check. The distinguishable RFLP patterns were assigned as operational taxonomic units (OTUs).

Sequencing and phylogenetic analysis

Prior to sequencing, PCR products were purified by using the DNA Purification System, Wizard DNA Clean-up (Promega) according the manufacturer's protocol. Purified PCR products were then sequenced using an automated sequencer ABI310 (PE Applied Biosystems) in the Biological and Analytical Laboratory at the Center of Marine Biotechnology, University of Maryland Biotechnology Institute. Both strands of the DNA were sequenced if there were uncertain nucleotides in the first strand of DNA sequence. Sequence alignment and phylogenetic reconstruction were conducted using Mac Vector 7.1 program (GCG, Madison, WI). Jukes-Cantor distance matrix analysis was used to calculate the distances from the aligned DNA sequences, and neighbor-joining method was used to construct a phylogenetic tree. To construct a phylogenetic tree based on the deduced amino acids, the pairwise similarity was calculated by using Blosum30 as matrix with open gap penalty of 10.0 and extend gap penalty of 0.1. The phylogenetic tree was constructed by using neighbor-joining method based on 166 amino acid residues. Bootstrap values for both trees were obtained from analysis of 1000 re-samplings of the data set.

Nucleotide sequence accession numbers

The nucleotide sequences determined in this study have been deposited in GenBank. The assigned accession numbers for the *Synechococcus* strains used in this study are shown in Table 2-1. The assigned accession numbers for OTUs are as follows: OTU1A (AY453140), OTU2B (AY453144), OTU3 (AY453145), OTU4C (AY453148), OTU5 (AY453149), OTU6 (AY453150), OTU7 (AY453152). The reference strains retrieved from GenBank included *Nostoc* sp. (previously *Anabaena*

sp). PCC7120 (P00879), *Chromatium vinosum* (M26396*)*, *Cylindrotheca* sp. (P24673), *Heterosigma akashiwo* (M24288), *Nitrobacter winogradskyi* (AF109915), *Prochlorococcus marinus* CCMP1378 (ZP_00104542), *Prochlorococcus marinus* MIT9313 (ZP_00113526), *Prochlorothrix hollandica* (P27568), *Rhodobacter capsulatus* (L82000), *Synechococcus* PCC 6301 (J01536), *Synechococcus* PCC 7002 (Q44176), *Synechocystis* PCC 6803 (BA000022), *Thiobacillus denitrificans* (L42940), *Thiobacillus ferrooxidans* A (Q07087)*, Thiobacillus ferrooxidans* B (P28895)*, Trichodesmium erythraeum* IMS101 (ZP_00072133).

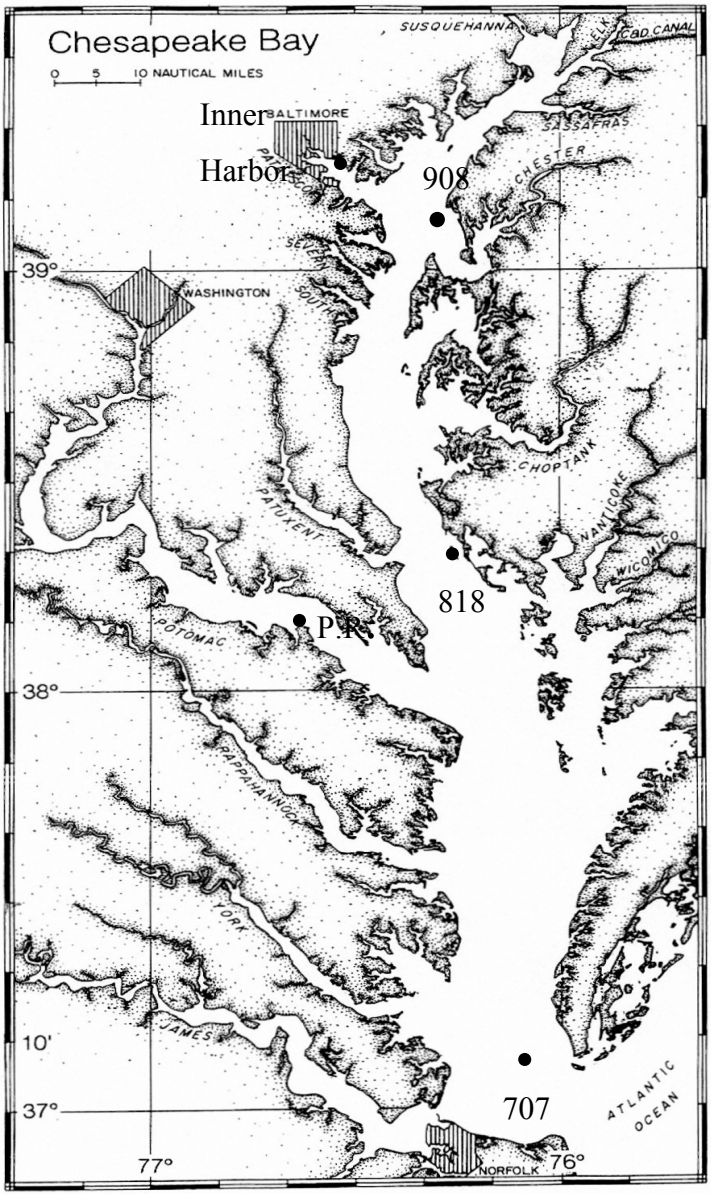


Fig. 2-1. Sampling locations in the Chesapeake Bay. Water samples were taken at 2 meter depth. At the time of sampling, the salinity at the stations 908, 818, 707, and Potomac River (PR) was 15, 17, 26, and 0‰, respectively. The salinity of Lewes Dock (38˚46’ N, 75˚08’ W) in the Delaware Bay (not shown in the map) was about 8%.

Results and Discussion

Characterization of Chesapeake Bay *Synechococcus* isolates

Among 13 CB strains, seven strains are the PC type (or non-PE type) and six strains are the PE type (Table 2-1). Microscopic examination showed that all the Chesapeake Bay *Synechococcus* isolates have the size ranging from 1 to 3 μm. The absorption spectra of Chesapeake Bay PC strains were similar to that of *Synechococcus* WH8101 in the Marine Cluster (MC- B), while the PE strains isolated from the Bay have an absorption spectrum similar to WH7805 (Fig. 2-2).

Table 2-1. *Synechococcus* strains used in this study

Strain	PE	PUB:PEB	Location	Source	Accession No.
CB0101	-	no PUB	Inner Harbor, CB	this study	AY244815
CB0102	-	no PUB	Inner Harbor, CB	this study	AY244816
CB0103	-	no PUB	Inner Harbor, CB	this study	AY244817
CB0104	-	no PUB	Inner Harbor, CB	this study	AY244818
CB0201	-	no PUB	Stn. 818, CB	this study	AY244819
CB0203	-	no PUB	Stn. 818, CB	this study	AY452725
CB0204	-	no PUB	Stn. 707, CB	this study	AY244820
CB0205	+	no PUB	Stn. 818, CB	this study	AY244833
CB0206	+	no PUB	Stn. 707, CB	this study	AY452726
CB0207	+	no PUB	Stn. 707, CB	this study	AY452727
CB0208	+	no PUB	Stn. 707, CB	this study	AY452728
CB0209	+	no PUB	Stn. 707, CB	this study	AY452729
CB0210	+	no PUB	Stn. 707, CB	this study	AY452730
WH5701	-	no PUB	Long Island Sound	Guillard, R	AY244822
WH8007	-	no PUB	Gulf of Mexico	Provasoli, L	AY244825
WH8101	-	no PUB	WHOI Harbor	Valois, F	AF081831
WH8102	+	high	Sargasso Sea	Waterbury, J	AF081832
WH8103	+	high	Sargasso Sea	Waterbury, J	AY244827
WH8108	+	high	Sargasso Sea	Waterbury, J	AY244831
WH7803	+	low	Sargasso Sea (25 m)	Waterbury, J	U46156
WH8104	+	low	Sargasso Sea	Waterbury, J	AY244828
WH8111	+	low	Sargasso Sea	Waterbury, J	AY244832
WH7805	+	no PUB	Sargasso Sea	Brand, L	AF081833
WH8006	+	no PUB	Gulf of Mexico	Waterbury, J	AY244824
WH8008	+	no PUB	Gulf of Mexico	Waterbury, J	AY244826

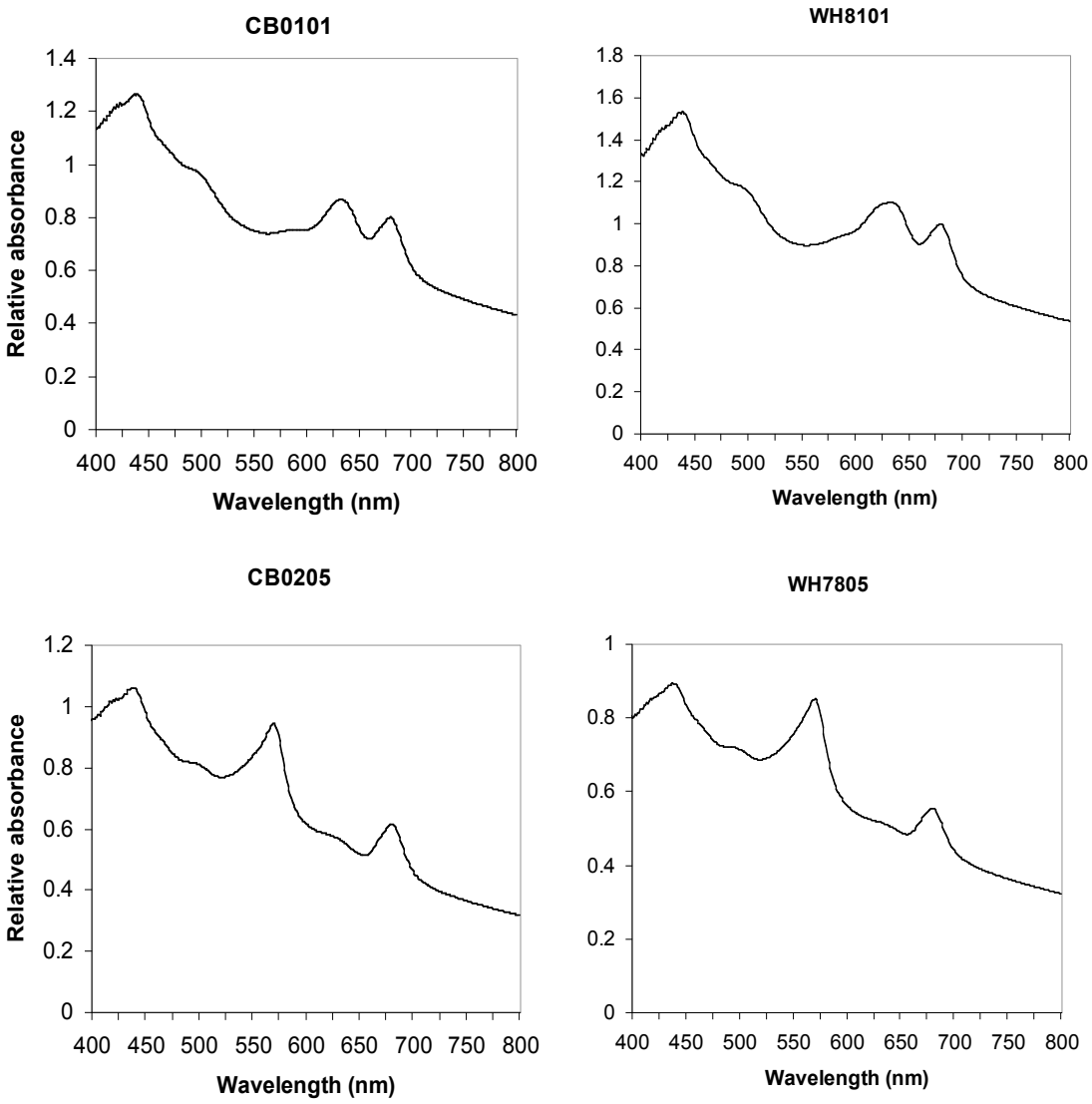


Fig. 2-2. Absorption spectra of Chesapeake Bay *Synechococcus* strains and MC-A stains. Left: Chesapeake Bay strain CB0101 (PC type) and CB0205 (PE type). Right: MC-A strain WH8101 (PC type) and WH7805 (PE type).

Six CB strains (CB0101, CB0104, CB0201, CB0205, CB0208, CB0210) were able to grow in the SN medium with a wide range of salinity (0, 15, and 30‰, respectively). In contrast, WH7803 and WH7805 could not grow in the SN medium with 0‰ salinity (Fig. 2-3). These data suggest that the Chesapeake Bay *Synechococcus* have been adapted to the estuarine ecosystem.

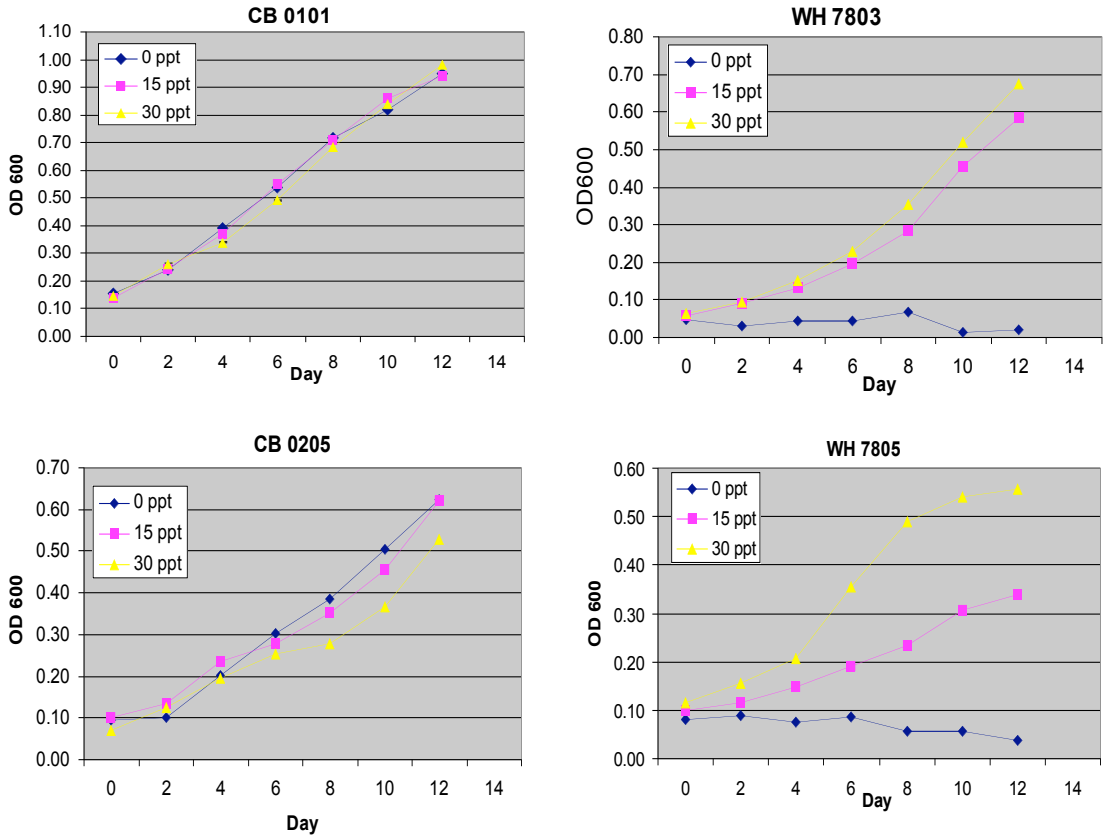


Fig. 2-3. Growth curve of *Synechococcus* strains in different salinity (0, 15 and 30 ppt) SN media. Left: Chesapeake Bay strain CB0101 (PC type) and CB0205 (PE type). Strain CB0104, CB0201, CB0208 and CB0210 have the similar growth curves (not shown). Right: MC-A strain WH7803 (PE type) and WH7805 (PE type).

During the course of isolation of *Synechococcus* from the Chesapeake Bay, we found that most of colonies (80-95% of all colonies) recovered from the Baltimore Inner Harbor and upper Chesapeake Bay (with salinity <15 ‰) were green strains. The percentage of the green colonies grown on the plate decreased to 56-65% at the mouth of the Bay (Station 707) where salinity is typically in the range of 25-30 ‰. The percentage of pigment types estimated based on the colonies recovered from the pre-incubated samples may not reflect the actual distribution of PC and PE *Synechococcus* in the Bay. Further examination of original water samples using epifluorescence microscopy showed that PC type *Synechococcus* were predominant (ca. 87%) in the upper Bay and Baltimore Inner Harbor, while PE type *Synechococcus* could account for approximately 82% of total *Synechococcus* populations in the lower Bay (station 707) (Fig. 2-4). The ratio of PC vs. PE type appeared to decrease from upper to lower Bay (Fig. 2-4).

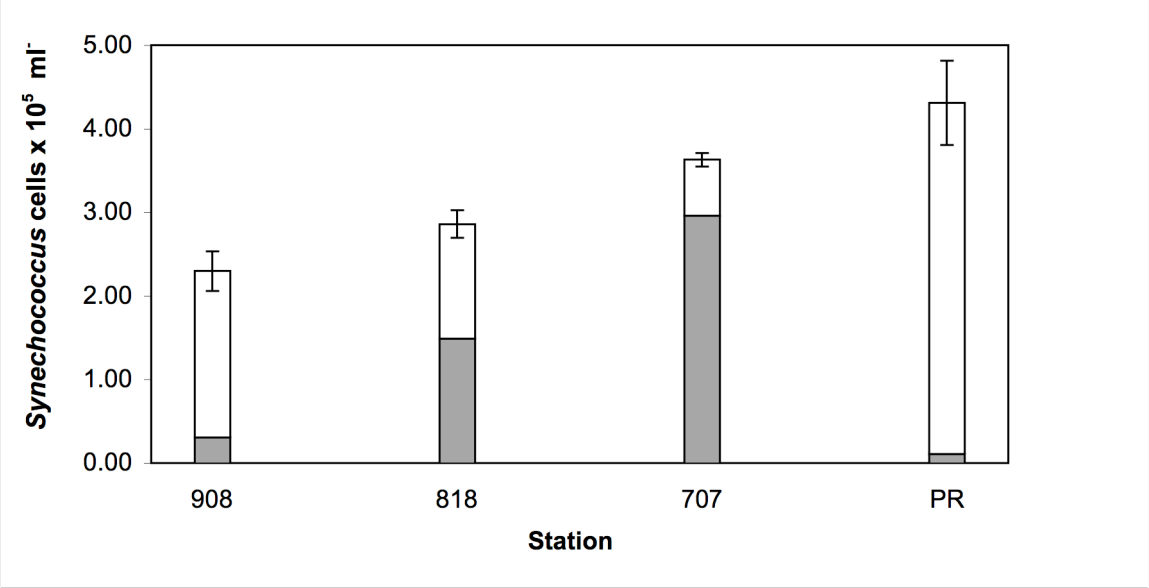


Fig. 2-4. Concentration of PC and PE type *Synechococcus* at the stations 908, 818 and 707, and in the Potomac River. Error bars represent SD of total cell counts. The percentages shown on top of each bar are the percent PE cells among all the *Synechococcus* counts. Closed bar: PE type; open bar: PC type.

Most of Chesapeake *Synechococcus* isolates belong to the MC-B cluster

The phylogenetic tree based on the *rbcL* DNA sequences of CB strains and the Bay samples revealed that *Synechococcus* living in the Chesapeake Bay are very diverse (Fig. 2-5). The BC strains alone exhibited the genetic diversity equivalent to the Woods Hole PC strains WH5701, WH8007 and WH8101 which were isolated from Long Island Sound, the Pier Harbor of Woods Hole Oceanography Institute, and the Gulf of Mexico, respectively. All the PE strains (Woods Hole collection) isolated from Sargasso Sea and the Gulf of Mexico (by Waterbury and Brand, see Table 2-1) were clustered within the MC-A group. Interestingly, five Chesapeake Bay PE strains (CB0206, CB0207, CB0208, CB0209 and CB0210) were closely related to WH8007, a member of MC-B (Fig. 2-5). Our data suggested that MC-B cluster not only contains the PC type *Synechococcus*, but also the PE type *Synechococcus*. Nine out of 13 CB isolates were grouped with WH8007, but none of them were closely related to WH5701. Strain CB0104 appeared to be in between WH5701 and WH8007. At least 10 CB strains can be considered belonging to the MC-B cluster (or subcluster 5.2), which contains only strains WH8007 and WH5701 previously.

According to the *rbcL* phylogeny, strain WH8101 could be clustered within the MC-A group, but with a relatively deep and independent branch. Recently, strain WH8101 has been re-classified into the MC-A or subcluster 5.1 (Herdman et al. 2001), and the close relationship between WH8101 and the MC-A group were also supported by the 16S-23S ribosomal DNA internal transcribed sequences (Rocap et al. 2002) and 16S rDNA gene sequences (Fuller et al. 2003), respectively. With PE strains found in the MC-B cluster and PC strains in the MC-A cluster, it is believed

that pigment types (PE or PC) are no longer a critical feature for separating MC-A and MC-B *Synechococcus*. Moreover, the designation of MC-B should be redefined when more data are available.

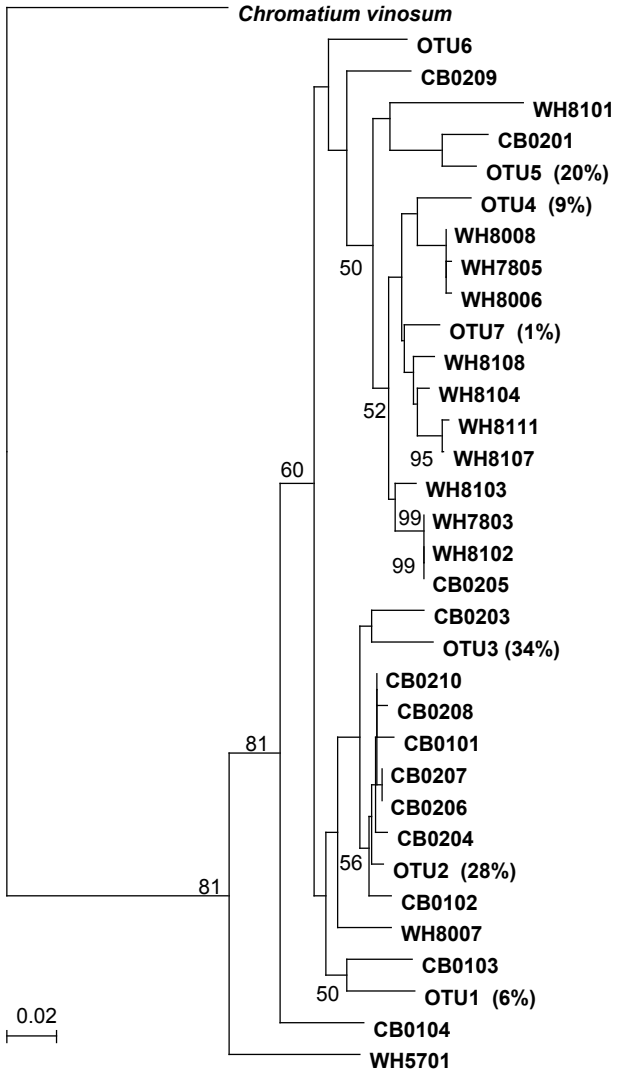


Fig. 2-5. Neighbor-joining tree of marine *Synechococcus* strains constructed using the partial *rbcL* nucleotide sequences (414 nt). OTUs included in this tree represent the seven distinct environmental *rbcL* genotypes (shown in Fig. 2-9). The percentage in the parenthesis is the relative distribution frequency of each OTU among 233 *rbcL* clones analyzed. Numbers at tree branches indicated bootstrap values with 1000 replicates. Bootstrap values less than 50 were not shown. The scale bar indicated 0.02 substitution per site.

***Synechococcus* strains from estuarine, coastal and oceanic waters form a monophyletic group.**

The phylogenetic affiliation at the amino acid sequence level showed that all the estuarine, coastal and oceanic *Synechococcus* used in this study, together with marine *Prochlorococcus marinus* share a common ancestry and belong to the Form IA RuBisCO (Fig. 2-6). *Prochlorococcus marinus* MIT9313 and CCMP1378 were closely related to each other, and also formed a close relationship with the marine *Synechococcus* cluster. The chemoautotrophic bacteria (e.g. *Chromatium vinosum*, *Nitrobacter winogradskyi*, *Rhodobacter capsulatus*, *Thiobacillus denitrificans*, *Thiobacillus ferrooxidans*) were in another branch of Form IA RuBisCO. All the freshwater cyanobacteria (e.g. *Anabaena* PCC7120, *Synechococcus* PCC6301 and PCC7002 and *Synechocystis* PCC6803) and marine filamentous cyanobacteria *Trichodesmium erythraeum* were within the Form IB RuBisCO.

Despite the wide range of salinity distribution in the Chesapeake Bay, all the CB strains (13 isolates) are still more closely related to the WH strains than to the freshwater *Synechococcus* which were grouped within the Form IB RuBisCO (Fig. 2-6). The CB and WH *Synechococcus* strains formed a monophyletic group (termed marine *Synechococcus* clade) that is closely related to marine *Prochlorococcus*. The close relationship between marine *Synechococcus* and marine *Prochlorococcus* revealed by the *rbcL* gene phylogeny is consistent with the previous studies based on the 16S rDNA sequences (Palenik and Haselkorn 1992; Urbach et al. 1992). The *rbcL* phylogeny also showed a close relationship between marine unicellular cyanobacteria and chemolithotrophic bacteria (Fig. 2-7). These results further support previous

studies in which the *Prochlorococcus marinus* and *Synechococcus* strain WH7803 *rbcL* sequences are found most closely related to the γ–purple bacterial sequences (Watson and Tabita. 1996; Gregory 1997; Watson and Tabita. 1997). It is very likely that marine *Synechococcus* and *Prochlorococcus* acquired the *rbcL* gene from a chemolithotroph through lateral gene transfer (Watson and Tabita. 1997).

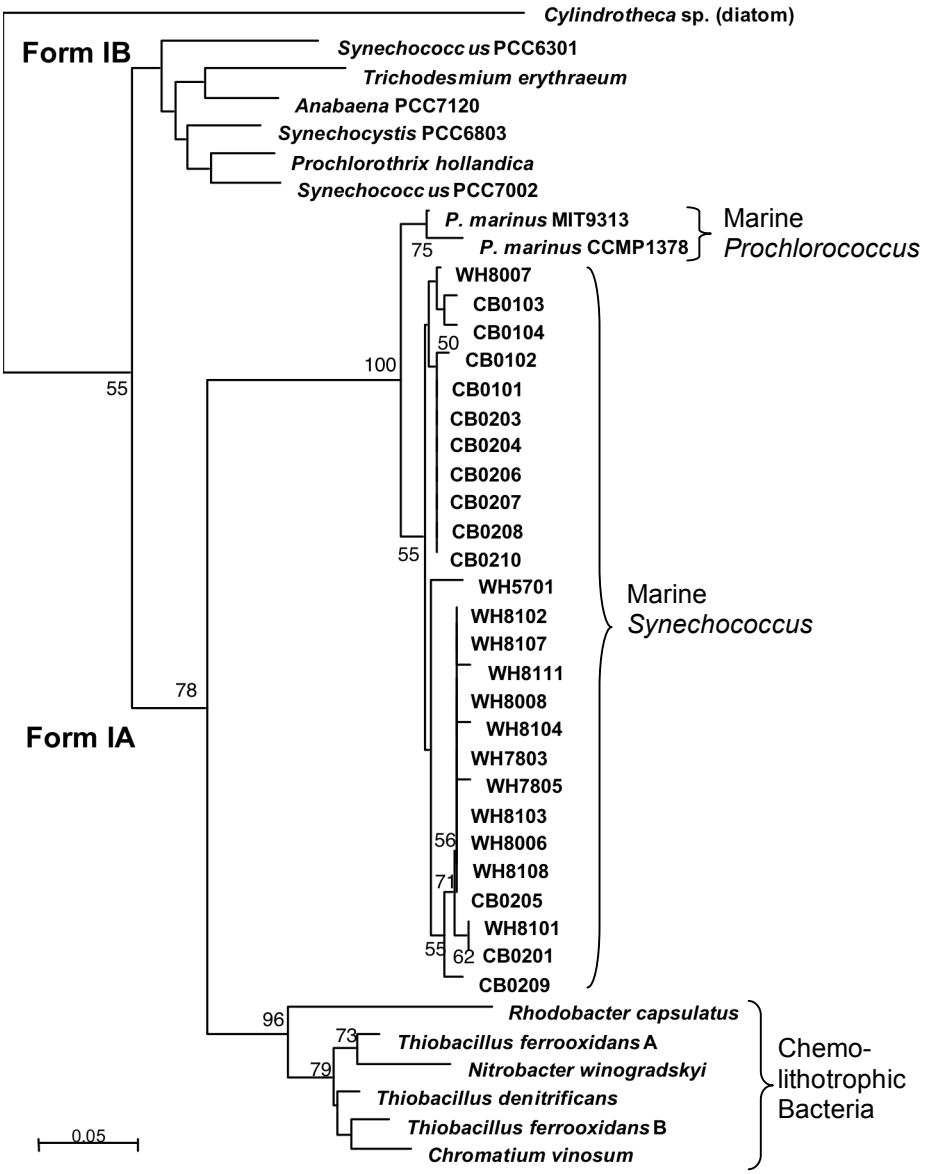


Fig. 2-6. Phylogenetic tree constructed using the neighbor-joining method based on the aligned *rbcL* amino acid sequences (138 residues). Form IA and IB are the two major groups of RuBisCO. Numbers at tree branches indicated bootstrap values with 1000 replicates. Bootstrap values less than 50 were not shown. The scale bar indicated 0.05 substitution per site.

Our result does not support a previous study in which marine *Synechococcus* WH8007 was clustered into the Form IB group based on the *rbcL* gene phylogeny (Pichard et al. 1997). In our study, the *rbcL* gene sequence of WH8007, together with other marine *Synechococcus* were all clustered into the Form IA group. Pichard et al. (1997) also indicated that the two *Prochlorococcus* strains were clustered together with freshwater *Synechococcus* in the From IB group. However, the phylogenetic analysis based on the various molecular markers (e.g. 16S rDNA, 16S-23S ITS, *rpo*C1 and *rbcL*) did not support such a close relationship between *Prochlorococcus* and freshwater *Synechococcus* (Palenik and Haselkorn 1992; Urbach et al. 1992; Shimada et al. 1995; Watson and Tabita. 1997; Badger et al. 2002; Rocap et al. 2002).

Development of marine *Synechococcus*-specific primers based on the *rbcL* gene

Because all marine *Synechococcus* form a unique cluster, a set of non-degenerate PCR primers that are specific to this group of unicellular cyanobacteria was developed. The upstream primer MSF1 is specific for marine *Synechococcus* strains while the downstream primer MSR1 is specific for marine *Synechococcus*, marine *Prochlorococcus*, some chemoautotrophic bacteria and freshwater *Synechococcus*. The primer set (MSF1/MSR1) was tested against the bacterial DNA recovered from several estuarine samples where the salinity ranged from 0 to 30 ‰. In all cases, specific PCR amplification was achieved (Fig. 2-7), suggesting that this primer set is suitable to amplify the *rbcL* gene from *Synechococcus* populations in various marine environments.

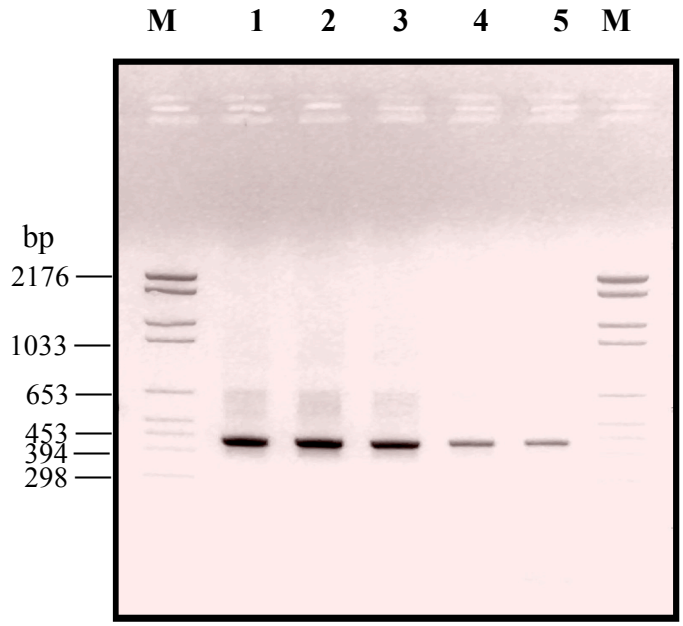


Fig. 2-7. PCR amplification of the *rbcL* gene from natural marine bacterial communities using the primer set MSF1 /MSR1, which is specific for marine *Synechococcus*. Lanes 1 to 5 correspond to stations 908 (Upper Bay), 818 (Mid Bay) and 707 (Lower Bay) of Chesapeake Bay, Potomac River and Lewes Dock (Roosevelt Inlet, Delaware), respectively. Lane M: Molecular Marker VI (Roche).

There are several reasons for developing non-degenerate *rbcL* primers that are specific for marine *Synechococcus*. First, this primer set will allow us to explore the genetic diversity of *Synechococcus* in the natural marine environment. Non-degenerate primers can also be coupled with rapid fingerprinting techniques such as denaturing gradient gel electrophoresis (DGGE) or the terminal restriction fragment length polymorphisms (T-RFLP) to monitor the *Synechococcus* populations over temporal and spatial scales. Second, non-specific PCR amplification is greatly minimized using the *rbcL* primers because the *rbcL* gene mainly exists in phototrophs and some chemolithotrophs. PCR amplification with the *rbcL* primers can improve the specificity and avoid the contamination problem associated with 16S rDNA primers. Third, the *rbcL* gene sequences (also *rpo*C1 and ITS sequences) provide

greater level of genetic resolution relative to the 16S rDNA gene sequences, and are particularly suitable for studying phylogenetic relationship of closely related strains like *Synechococcus*. Finally, this primer set will allow us to study the ecological interaction between *Synechococcus* and cyanophages in the natural marine environment. Viruses that infect marine *Synechococcus* (synechophage) are ubiquitous and abundant in various marine environments (Suttle and Chan 1993, 1994, Waterbury and Valois 1993, Wilson et al. 1993, Lu et al. 2001). The group-specific PCR primers have been applied to investigate the genetic diversity of synechophage in the Chesapeake Bay and other marine environments (Fuller et al. 1998; Zhong et al. 2002; Marston and Sallee 2003; Wang and Chen 2004), so it would be ideal to have a specific primer set for their corresponding hosts. In order to understand the complex interaction between marine viruses and their microbial hosts, it is necessary to look at the genetic diversity of both viral and host communities (Fuhrman 1999).

Highly diverse *Synechococcus* in the Chesapeake Bay

At least seven genotypes (or OTUs) of *Synechococcus* were found in the Chesapeake Bay based on the RFLP analysis of environmental *rbcL* clones (Fig. 2-8). All the estuarine stations except for Stn. 707 contained at least six different OTUs (Fig. 2-9). The seven OTUs recovered from the four stations in the Chesapeake Bay and Lewes Dock in the Delaware Bay were clustered within the Marine *Synechococcus* clade (Fig. 2-5). Among 232 clones tested, 10% (OTUs 4 and 7) were clustered with PE strains in MC-A, 22% (OTUs 5 and 6) were related to WH8101 (PC strain in MC-A), and 68% (OTUs 1, 2 and 3) were affiliated with the MC-B

group (Fig. 2-5). The composition of environmental *rbcL* sequences suggests that Chesapeake Bay is dominated by the MC-B type *Synechococcus*.

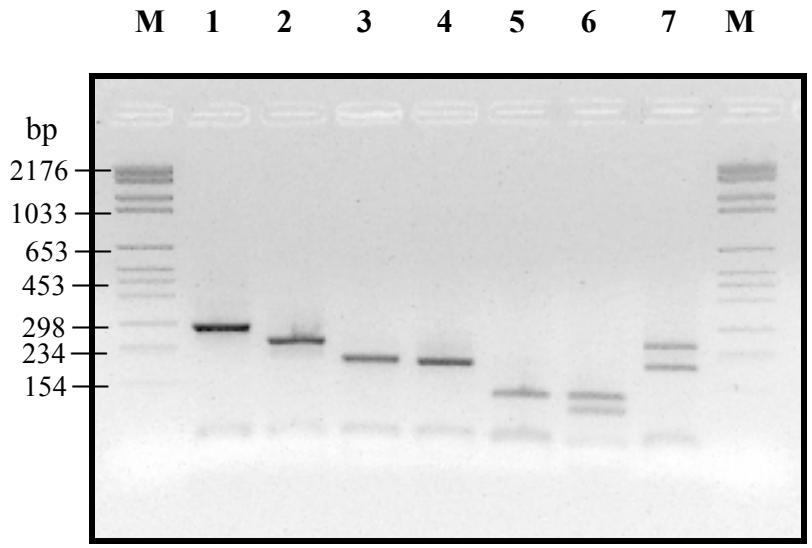


Fig. 2-8. Seven different *rbcL* genotypes or OTUs revealed by RFLP analysis (lanes 1-7) of *rbcL* clone libraries (a total of 232 clones). Lane M: Molecular Marker VI (Roche).

In the Chesapeake Bay, the distribution of these *Synechococcus* genotypes changed from upper to lower Bay (Fig. 2-9). For example, from the upper Bay (Stn. 908) to lower Bay (Stn. 707), OTU2 gradually decreased from 51 to 15%, while OTU3 steadily increased from 16 to 49%, respectively. The dramatic change of *Synechococcus* populations in the Chesapeake Bay is reflected by the complicated ecological habitats in the Bay. The Chesapeake Bay is an very complex ecosystem in which salinity, temperature and nutrients are very dynamic both temporally and spatially. High salinity seawater from the Atlantic Ocean enters the mouth of the Bay, and gradually decreases towards north. Salinity levels in the Bay can range from nearly 0 to 30 ‰ across a north-south transect. In our study, the OTUs were selected

based on the RFLP digestion patterns of one restriction enzyme, and actual genetic diversity of *Synechococcus* in the Bay could be underestimated using this method. In the California Current, seven genetically distinct groups of *Synechococcus* have been identified based on the *rpo*C1 sequence data (Toledo and Palenik 1997; Ferris and Palenik 1998; Toledo et al. 1999; Palenik 2001).

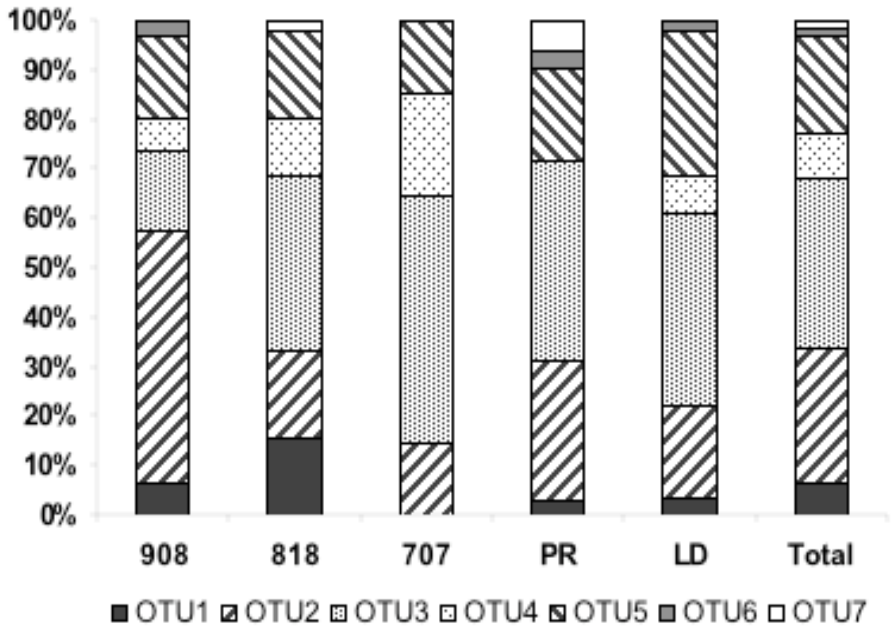


Fig. 2-9. Relative percentage of each OTUs presented in each sampling locations, Chesapeake Bay stations 908, 818 and 707, Potomac River (PR) and Lewes Dock (LD). The relative distribution of each OUT among the total 232 clones is shown as "Total".

Conclusion

Despite the great genetic diversity revealed by the *rbcL* gene sequences of *Synechococcus* isolates and environmental clones, *Synechococcus* populations in the Chesapeake Bay appear to be dominated by the MC-B (or subcluster 5.2) type

Synechococcus. Several novel PE strains within the MC-B cluster were found in the lower Bay. All the CB strains are halotolerant, suggesting they have been adapted to the Chesapeake estuary where salinity could change from almost 0 to 30 ‰. *Synechococcus* in estuarine, coastal and oceanic waters (at least for the MC-A and B) are monophyletic rather than widely dispersed across the *rbcL* phylogenetic tree. With the marine *Synechococcus*-specific primers available, it is now possible to explore the genetic diversity of marine *Synechococcus* populations. At least seven different genotypes of *Synechococcus* from the Bay have been identified, and in most cases six distinct genotypes could be found in one sample. Our long-term goal is to explore the co-variation of both *Synechococcus* and their phage, simultaneously in the natural aquatic environments. The spatial and temporal variations of *Synechococcus* populations, together with synechophage populations in the Chesapeake Bay are currently under investigation.

Chapter 3: Diverse and unique picocyanobacteria found in the Chesapeake Bay, revealed by 16S-23S rRNA internal transcribed spacer sequences

Abstract

Picophytoplankton (< 3μm), predominantly picocyanobacteria, have been known to contribute significantly to total phytoplankton biomass and primary production in the Chesapeake Bay, particularly in summer. Abundance of picocyanobacteria in the Bay exhibits strong seasonal variation. Although diversity of picocyanobacteria in the coastal and open ocean has been studied extensively, little is known about the population structures of picocyanobacteria in estuarine ecosystem, like the Chesapeake Bay. The composition of Chesapeake picocyanobacterial populations based on 16S-23S rRNA internal transcribed spacer (ITS) sequences suggests that the Bay provides a unique ecological niche for the marine cluster B (MC-B) *Synechococcus*, a poorly studied group of picocyanobacteria. ITS phylogeny based on 14 isolates and 22 environmental clones indicates that at least five new subclusters (>95% sequence identity) of *Synechococcus* are novel and unique to the Chesapeake Bay. In a broader context, these data provide a new insight into the classification of picocyanobacteria. Phylogenetic comparison with 82 picocyanobacterial strains or environmental clones from a broad range of aquatic environments found that the vast majority of Chesapeake Bay *Synechococcus* populations are associated with marine rather than freshwater environments. Both

phycocyanin- and phycoerythrin-rich *Synechococcus* were found in the MC-B cluster, suggesting that the accessory pigment is not a stable characteristic for cyanobacterial taxonomy. Picocyanobacteria may acquire or lose the gene that is responsible for producing phycoerythrin through their adaptation to a highly variable environment like the Chesapeake Bay.

Introduction

Phototrophic picoplankton (< 3μm) play an important role in ocean's carbon cycle (Waterbury et al. 1979; Li et al. 1983; Platt et al. 1983). In the Chesapeake Bay, picophytoplankton contribute 10 to 20% of total primary production during summer (Ray et al. 1989; Malone et al. 1991; Affronti and Marshall 1993). Picophytoplankton can reach cell density of about 1 million cells per ml, and are responsible for 56% of total phytoplankton primary production in the lower Bay (Affronti and Marshall 1994). It has been reported that picocyanobacteria account for ca. 98% (cell counts) of the Bay picophytoplankton, while picoeukaryotes (e.g. small diatoms, flagellates and dinoflagellates) make up only 2% of picophytoplankton community (Ray et al. 1989). Recently, picocyanobacterial strains isolated from the Bay were found to be affiliated with a group of marine *Synechococcus* based on the phylogenetic analysis of RuBisCO large subunit gene sequences (Chen et al. 2004). However, little is known about the diversity and distribution of picocyanobacterial populations in different Chesapeake habitats.

The genus *Synechococcus* is loosely defined as small (1-3 μm) unicellular coccoid to rod-shaped cyanobacteria (Waterbury and Rippka 1989). They are abundant and responsible for a significant part of marine primary production (Waterbury et al. 1979b; Waterbury et al. 1986). Genetic heterogeneity among *Synechococcus* is apparent from the wide range of genomic DNA % G+C ratio (39 to 71%), which is almost as broad as that of all prokaryotes (Waterbury and Rippka 1989). *Synechococcus* in aquatic environments are diverse and contain polyphyletic lineages. New ecotypes and genotypes continue to emerge as *Synechococcus* diversity from different ecosystems is explored (Moore et al. 1998; Urbach et al. 1998; Honda et al. 1999; Robertson et al. 2001; Rocap et al. 2002; Crosbie et al. 2003; Ernst et al. 2003; Ferris et al. 2003; Fuller et al. 2003; Becker et al. 2004; Chen et al. 2004). In general, less is known about *Synechococcus* living in coastal and estuarine regions compared to offshore *Synechococcus* strains (Scanlan and West 2002).

Based on morphological, physiological and chemical properties, and DNA base composition, the marine *Synechococcus* lineages has been split into three clusters: Marine Cluster A, B and C (MC-A, MC-B, and MC-C) (Waterbury and Rippka 1989). The MC-A cluster contains diverse *Synechococcus* isolated from coastal waters and open oceans, and its classification is supported by the 16S rRNA and ITS gene phylogenies (Herdman et al. 2001; Rocap et al. 2002; Fuller et al. 2003). The MC-C cluster currently contains four closely related marine *Synechococcus* strains, and is supported by 16S rRNA gene based phylogeny (Herdman et al. 2001; Fuller et al. 2003). In contrast, the MC-B cluster contains diverse marine *Synechococcus* strains whose relationship is poorly supported by the molecular phylogeny. The MC-B cluster originally contained four *Synechococcus*

strains, PCC7001, WH5701, WH8007, WH8101, all isolated from coastal waters (Waterbury and Rippka 1989) and containing only phycocyanin (PC). Strain PCC7001 has now been reclassified as a member of the *Cyanobium* cluster based on 16S rRNA gene phylogeny (Herdman et al. 2001). Strain WH8101 was found to be more closely related to MC-A (or *Synechococcus* subcluster 5.1) than to strain WH5701, while WH5701 was phylogenetically distinct from all other marine *Synechococcus* isolates (Rocap et al. 2002; Fuller et al. 2003; Chen et al. 2004). The remaining MC-B strain WH8007 has seldom been included in previous phylogenetic analyses of *Synechococcus*. Recent phylogenetic analysis based on the *rbcL* gene showed that WH8007 is neither closely related to WH8101, nor to WH5701(Chen et al. 2004). Although current MC-B (or *Synechococcus* subcluster 5.2) still includes WH5701, WH8007 and WH8101 (Herdman et al. 2001), it is necessary to elucidate the phylogenetic relationship among these strains. The observation that many *Synechococcus* isolates from Chesapeake Bay were affiliated with WH8007 (Chen et al. 2004) prompted further investigation of Chesapeake Bay *Synechococcus* populations to provide new insight into the classification of marine *Synechococcus*, particularly for MC-B (or Cluster 5.2). Chesapeake Bay is the largest estuary in U.S., which provides strong hydrological gradients and diverse habitats for picophytoplankton. The study of *Synechococcus* diversity in the Chesapeake Bay will lead to better understanding of the population biology and ecology of this globally significant group of picophytoplankton.

In this study, the phylogenetic relationships among 14 Chesapeake *Synechococcus* strains and 22 cyanobacterial environmental rRNA operon gene clones were determined based on rRNA operon internal transcribed spacer (rRNA-ITS) sequences. The rRNA-

ITS was used in this study because it provides greater sequence heterogeneity than the 16S rRNA gene, and a considerable number of rRNA-ITS from marine and freshwater *Synechococcus* isolates are available (Rocap et al. 2002; Ernst et al. 2003). Using culture and culture-independent approaches could provide a comprehensive picture on the diversity of picocyanobacteria in the Bay.

Materials and Methods

Synechococcus isolates

Phylogenetic relationships were reconstructed based on gene sequences of 59 picocyanobacterial strains from freshwater, brackish water, coastal and oceanic waters (Rocap et al. 2002; Ernst et al. 2003; Chen et al. 2004). Chesapeake isolate CB0202, a PC-rich *Synechococcus* isolated from station 908 in 2002 is the only previously undescribed strain reported here. All the Chesapeake isolates were grown in the SN medium (Waterbury et al. 1986) at 22 °C under 14/10 hr light : dark cycle with 10-20 microeinsteins m^{-2} s^{-1} illumination.

Field sample collection

Water samples (250 ml) from Chesapeake Bay were collected at three stations (Stn. 908, 818 and 707, Table 3-1) at a 2-meter depth using Niskin bottles on board the R/V *Cape Henlopen* on September 26-30, 2002 and March 4-8, 2003. Samples were processed and stored as described in Chapter 2.

Nucleic acid extraction

Nucleic acids from *Synechococcus* isolates and the natural bacterial community were obtained by enzymatic lysis (lysozyme, proteinase K, and sodium dodecyl sulfate) followed by phenol-chloroform extraction and isopropanol precipitation (Schmidt et al.

1991). DNA concentration was measured using a SmartSpec TM 3000 spectrophotometer (Bio-Rad, Hercules, CA).

Analysis of rRNA-ITS region: Isolates

For *Synechococcus* isolates, 16S rRNA-ITS-23S rRNA fragments (ca. 1.2-1.4kb) were amplified using primers 16S-1247F and 23S-241R described by Rocap et al. (Rocap et al. 2002). Twenty-five microliter reactions contained 1 × PCR buffer, 1 pmole of each primer, 0.25 mM dNTPs, 1.5 mM $MgCl_2$, 1 U *Taq* DNA polymerase (Promega, Madison, WI) and 15-50 ng genomic DNA as template. PCR cycles included a 3-min initial denaturing step at 94°C, followed by 30 cycles of 94°C for 30 seconds, 52°C for 30 seconds, 72°C for 1 min, and a 7-min final extension at 72°C. PCR products were purified using the Qiagen PCR Purification Kit (Qiagen, Chatsworth, CA), according the manufacturer's protocol and bidirectionally sequenced on an ABI Prism 3100 Genetic Analyzer using Big Dye V3.1 chemistry (Applied Biosystems, Foster City, CA). Sequencing primers 16S-1247F and 23S-241R, as well as primers targeting tRNA alanine sequences in the ITS (Ala-F: 5'-GAGCGCCTGCTTTGCAAGCAG-3'; Ala-R: 5'-CTGCTTGCAAAGCAGGCGCTC-3') were used in cycle sequencing reactions.

Analysis of rRNA-ITS region: Environmental clones

Clone libraries containing a large portion of the rRNA operon (16S rRNA-ITS-23S rRNA) from bacterioplankton within six surface water samples were constructed as previously described (Suzuki et al. 2001) with the following changes: 1) Platinum HIFI polymerase mix (Invitrogen, Carlsbad, CA) was used to provide hot start amplification; 2) PCR products were A-tailed using the Qiagen A-addition kit (Qiagen, Chatsworth, CA); and 3) products were cloned using the TOPO TA Cloning Kits for Sequencing

(Invitrogen). A minimum of 82 clones from each library were screened by a novel screening method adapted from the ITS-LH-PCR (Internal transcribed spacer-length heterogeneity-PCR) method which measures the length variation of two fragments amplified by PCR with fluorescence-labeled primers (Suzuki et al. 2004). Clones putatively identified as cyanobacteria based on the sequence length of: 1) the entire spacer, and 2) the space between the end of the 16S rRNA gene and the beginning of tRNA-alanine were chosen for further analysis. Plasmids were purified using the FastPlamid (Eppendorf, Westbury, NY) and Montage Miniprep96 (Millipore) kits and then sequenced.

Phylogenetic analysis

ITS Sequences were aligned using Mac Vector 7.2 program (GCG, Madison, WI) using Clustal W with slow pair-wise alignment mode (Open Gap Penalty = 10, Extended gap penalty = 1). Aligned sequences were corrected manually. Poorly aligned regions were masked and excluded from phylogenetic analysis. Jukes-Cantor distance matrix analysis was used to calculate a distance matrix from aligned DNA sequences. Neighbor-joining method was used to construct a phylogenetic tree. The phylogeny was also constructed using the Maximum Parsimony method on the Molecular Evolutionary Genetics Analysis software, MEGA3 (Kumar et al. 2004), with close-neighbor-interchange search method (search level=1, random addition tree +10 replicates). For both distance and maximum parsimony analysis, bootstrap values were obtained from an analysis of 1000 random re-samplings of the data set. Nucleotide sequences determined in this study have been deposited in GenBank and the assigned accession numbers for the

Synechococcus strains and environmental clones described in this study are shown in Fig. 3-1 and Fig. 3-2.

Table 3-1. Three *rrn* operon clone libraries constructed from water samples collected in the upper, middle, and lower Bay in 2002. Twenty-three of 263 clones were identified as cyanobacteria based on prescreening of ITS length. Relevant physical, chemical and biological characteristics at these stations were included.

Clone library	CB01 (upper Bay)	CB11 (mid-Bay)	CB22 (lower Bay)
Station name	Stn. 908	Stn. 818	Stn. 707
Location	39.08 N, 76.20 W	38.18 N, 76.17 W	37.07 N, 76.07 W
Water temperature (˚C)	23.3	23.9	24.2
Salinity (ppt)	15.5	19.4	27
Bacteria counts (10^6 cells ml^{-1})	6.42	2.91	2.57
Synechococcus spp. count (10^6 cells ml^{-1})	0.23	0.29	0.36
Percent *Synechococcus* in total bacteria (%)	3.58	9.97	14.01
Percent PC type in total *Synechococcus* (%)	86.7	47.8	18.4
Total clones	91	84	88
No. of cyanobacterial clones (percentage)	4 (4.4%)	12 (14.3%)	7 (8.0%)
Prescreening size (FAM/HEX):			
943/458			CB22A09
954/464		CB11C11, CB11D02	
964/457		CB11B02, CB11E03, CB11H03	CB22A07
974/464	CB01C11, CB01E02, CB01C12	CB11C04, CB11D06	
996/460		CB11F09, CB11H07	CB22D04, CB22G11
1007/460		CB11D12, CB11G04	CB22H05
1056/486	CB01F08		CB22C09
1123/487		CB11G10	

Results and Discussion

Phylogenetic diversity of Chesapeake Bay picocyanobacteria

Phylogenetic analysis included a total of 82 picocyanobacteria ITS sequences (Fig. 3-1). This collection of sequences included strains and environmental clones from freshwater lakes (14 sequences); Baltic Sea brackish waters (5 sequences); Chesapeake Bay (36 sequences); and coastal and oceanic waters (27 sequences). Thirty-three out of 36 Chesapeake Bay isolates and environmental clones were affiliated with either MC-A or MC-B, suggesting that a vast majority of picocyanobacteria in the Chesapeake Bay are members of marine *Synechococcus* clades. Eleven of 14 Chesapeake isolates clustered with WH8007 (MC-B cluster). One strain CB0201 fell within the MC-A cluster, and two strains CB0104 and CB0203 formed a clade independent of any ITS-defined picocyanobacterial group. Among 22 environmental clones putatively identified as cyanobacteria, 13 clustered within MC-A, 9 within MC-B, and one clone (CB11G10) formed a deep branch within the *Cyanobium* cluster. Possible explanations for the discrepancy between the clustering of isolates versus environmental clones (i.e. 13 environmental clones were members of the MC-A cluster but only one CB isolate) are: 1) The samples used for isolation and clone library construction were not the same; 2) Culture media does not perfectly mimic the actual environment (e.g. different salinity). The salinity of media used for isolation (10-20 ppt) favored the growth of estuarine Chesapeake Bay strains rather than MC-A strains known to have an elevated salt requirement for sustaining growth. Regardless, culture and culture-independent approaches confirmed that freshwater *Synechococcus* are rare in the Chesapeake Bay, even in the upper Bay where salinity is in the range of 5-10 ppt.

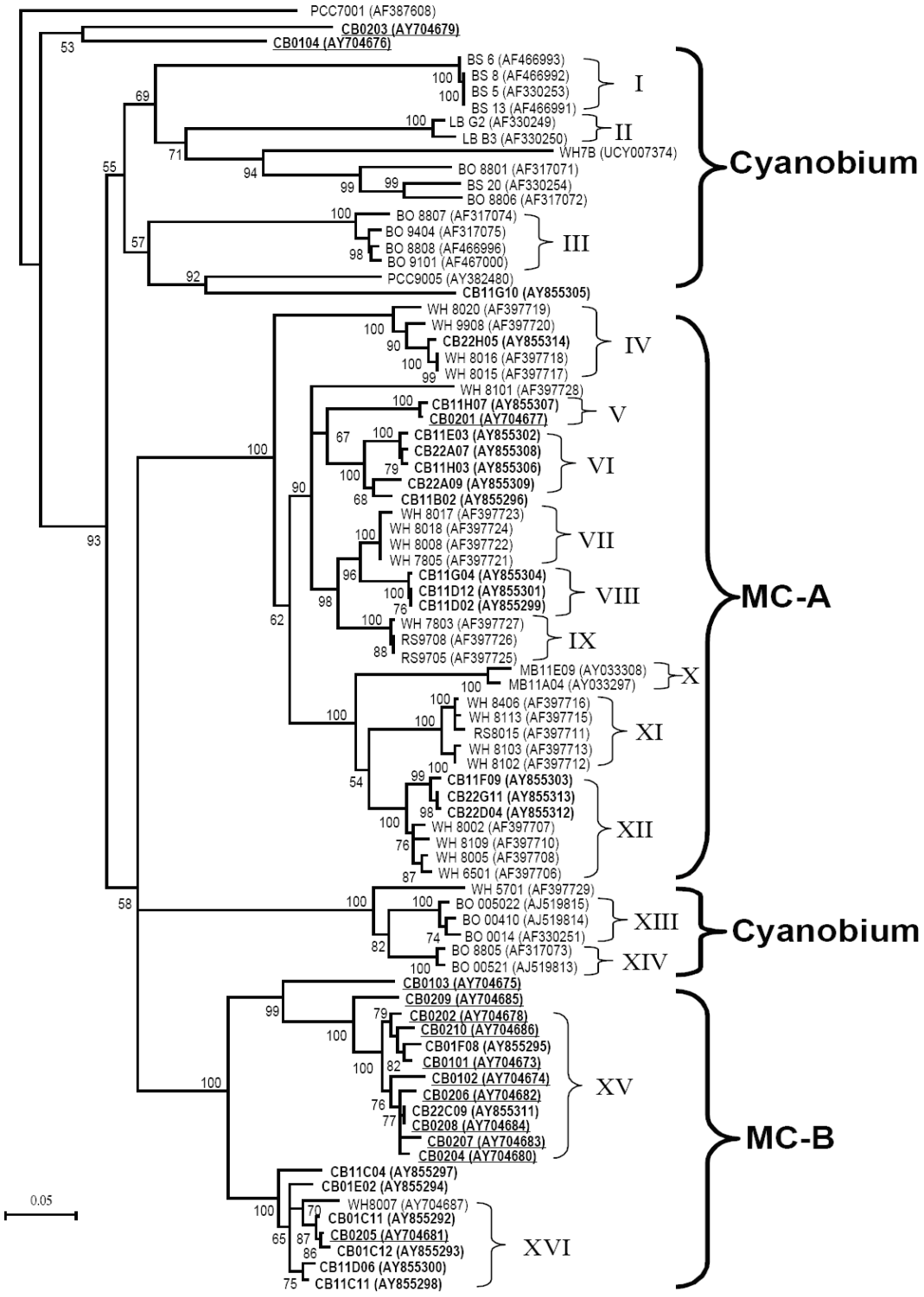


Fig. 3-1. Phylogenetic tree constructed using the neighbor-joining method based on ITS sequence of picocyanobacterial isolates and environmental clones collected from lakes, brackish, estuarine, coastal and oceanic waters. The whole ITS region was aligned, and

final sequence alignment was corrected manually. A total of 786 aligned nucleotide positions were used for constructing the tree. The tree is rooted with PCC7001. Numbers at tree branches indicated the bootstrap values with 1000 replicates. Bootstrap values less than 50 were not shown. The scale bar is equivalent to 0.05 substitution per site. Names in bold type represent the isolates (underlined) or environmental clones (not underlined) from Chesapeake Bay. Prefixes for the *Synechococcus* strains or environmental clones are as follows: (CB) Chesapeake Bay; (WH) Woods Hole; (RS) Red Sea; (MB) Monterey Bay; (LB) Lake Biwa; (BO) Lake Constance (Bodensee); and (BS) Baltic Sea. The strains with PCC prefix are the picocyanobacterial isolates collected by the Pasteur Culture Collection. GenBank accession numbers are included in parenthesis.

At least 16 subclusters (>5% sequence identity) could be identified across all the picocyanobacteria included in this study (Fig. 3-1). All the 16 subclusters were well supported with bootstrap values, 100 with the Neighbor-joining method (Fig. 3-1), and >97 with Maximum Parsimony (data not shown). Five isolates (CB0104, CB0203, WH7B, PCC7001 and PCC9005) and one environmental clone (CB11G10) formed a deep branch without closely related counterparts, suggesting that there are still many novel cyanobacterial genotypes in aquatic environments. Eleven subclusters (I, II, III, IV, VII, IX, X, XI, XII, XIII, and VIX) overlap with previously reported subclusters (Rocap et al. 2002; Ernst et al. 2003), while at least five new subclusters (V, VI, VIII, XV and XVI) were novel and unique to the Chesapeake Bay. Chesapeake isolates and environmental clones were not affiliated with WH5701 and PCC7001 (former MC-B members), or to picocyanobacteria isolated from lakes or Baltic Sea brackish water.

Re-classification of MC-B

Based on ITS phylogeny the four originally described MC-B strains WH5701, WH8007, WH8101, and PCC7001 are distantly related to each other (Fig. 3-1). Strain WH5701, appeared to be related to several *Cyanobium* strains isolated from the pelagic zone of Lake Constance, Germany. WH5701 isolated from Long Island Sound, USA has

a high genomic G+C% content (66%) which falls within the G+C% range of *Cyanobium* (66-71%) (Waterbury and Rippka 1989). This close relationship between WH5701 and a *Cyanobium* subcluster has not been previously reported. Furthermore, our results confirm that WH8101 is more affiliated with the MC-A cluster than MC-B (Rocap et al. 2002; Fuller et al. 2003).

Many *Synechococcus* isolates and environmental clones from the Chesapeake estuary were related to WH8007 (Fig. 3-1). Six environmental clones from the upper and mid Bay, together with CB0205 were clustered closely with WH8007, while another subgroup containing 10 CB isolates and two environmental clones was also affiliated with WH8007. *Synechococcus* strain WH8007 was originally isolated from the Gulf of Mexico and has 63% genomic G+C content (Waterbury and Rippka 1989). This study, combined with earlier work on *rbcL* phylogeny (Chapter 2) suggest that WH8007, rather than WH5701, should be the reference strain for MC-B (or *Synechococcus* Cluster 5.2).

MC-B is a polyphyletic group that contains both PE- and PC-rich *Synechococcus*. At least two subclusters (XV and XVI) can be defined within the MC-B cluster. Five PE-rich *Synechococcus* strains (CB0206, CB0207, CB0208, CB0209, and CB0210) were closely related to four PC-rich *Synechococcus* strains (CB0101, CB0102, CB0202, and CB0204) within subcluster XV. MC-B was originally described as containing only PC-rich *Synechococcus* (Waterbury and Rippka 1989). These results however suggest that both PC-rich and PE-rich strains can be found in MC-B, and that absence of phycoerythrin is not a characteristic of MC-B affiliated *Synechococcus*. The instability of the pigment trait as a taxonomic reference has been previously reported for picocyanobacteria, with different light harvesting pigments are combined in a

monophyletic group. For example, mixture of PE- and PC-rich strains within a well defined cluster like the Bornholm Sea cluster have been reported by Ernst et al. (2003). Moreover, PC-rich *Synechococcus* strain WH8101 is now added to the MC-A which contains only PE-rich strains previously. It has been proposed that some cyanobacterial strains may lose or gain the genes required for producing PE through ecological adaptation (Ernst et al. 2003), and genetic exchange can occur among phyletically related cyanobacterial strains (Rudi et al. 1998). Whether mechanisms like point mutation or down regulation of gene expression are responsible for rapid chromatic adaptation is still unknown. Adaptive radiation allows MC-B *Synechococcus* to become dominant in the upper Bay, and MC-A *Synechococcus* to become prevalent in the lower Bay.

ITS phylogeny vs. *rbcL* phylogeny

For the collection of WH and CB strains ITS phylogeny was in good agreement with the *rbcL* phylogeny (Fig. 3-2). Both ITS and *rbcL* gene phylogeny show a close relationship among CB isolates CB0101, CB0102, CB0204, CB0206, CB0207, CB0208, and CB0210. Close relationships between strains WH7805 and WH8008; WH8107 and WH8108; and WH8102 and WH8103 can be seen in both phylogenetic reconstructions (Fig. 3-2). Placement of some branches (e.g. CB0203 and CB0209) was not the same between the two trees, particularly for deep branches poorly supported by the bootstrap analysis. Based on *rbcL* phylogeny, CB0205 originally clustered with two MC-A strains (WH7803 and WH8102) (Chapter 2); however, based on the ITS phylogeny, CB0205 clustered with WH8007 in MC-B (Fig. 3-1). To confirm this incongruence the *rbcL* gene and ITS region of CB0205 were re-sequenced. Surprisingly, the newly obtained *rbcL* gene sequence of CB0205 was different from the original one (the *rbcL* sequence of

CB0205 has now been corrected with the same GenBank accession No. AY244833). In agreement with ITS phylogeny, the corrected *rbcL* sequence of CB0205 is indeed related to that of WH8007 (Fig. 3-2). It was very likely the original CB0205 culture used for *rbcL* gene analysis was contaminated with MC-A strain.

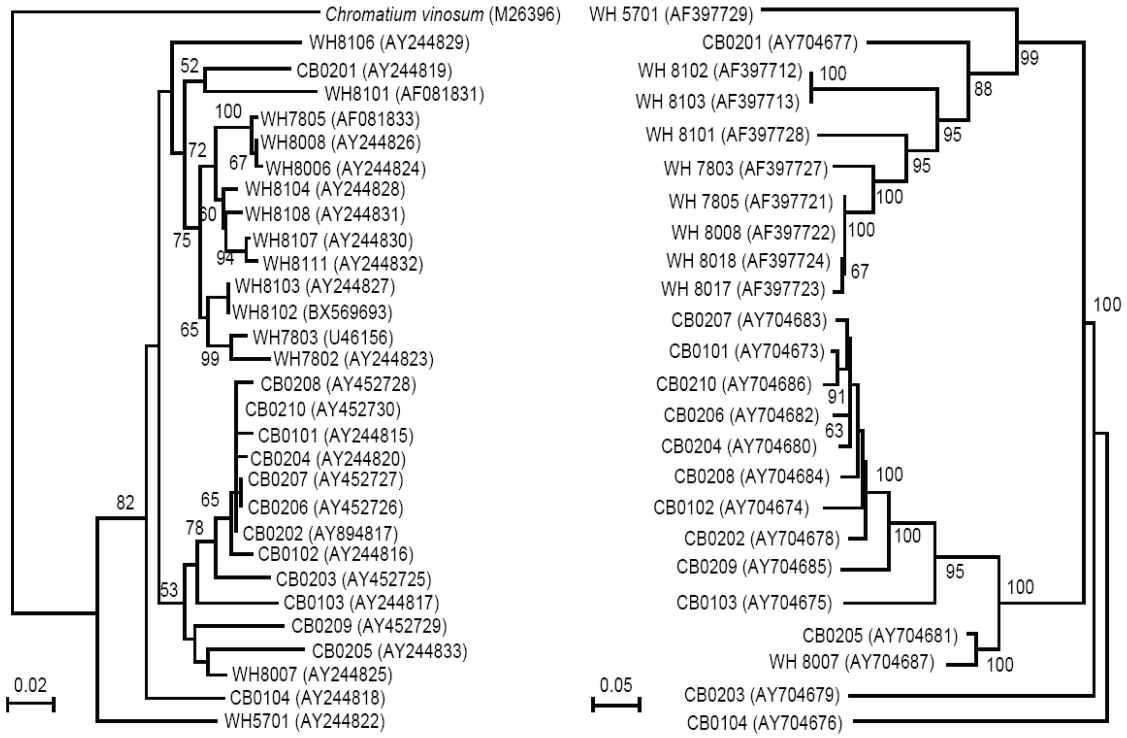


Fig. 3-2. Comparison of phylogenetic trees constructed from *rbcL* gene (left panel) and ITS (right panel) sequences within CB and WH isolates. Bootstrap values were calculated based on 1000 re-sampling. Those values less than 50 are not shown. The scale bar is equivalent to 0.02 substitution per site for the *rbcL* phylogenetic tree, and 0.05 substitution per site for the ITS phylogeny.

Sequence divergence of ITS regions within marine *Synechococcus* is higher than that of 16S rRNA, and other phylogenetic marker genes like *rbcL* and the RNA polymerase gene, *rpo*C1 (Table 3-2). The *rpo*C1 gene has been used to investigate the genetic diversity of *Synechococcus* in marine environments (Toledo et al. 1999; Palenik 2001). In general, the MC-A and MC-B clusters were better separated and supported with higher bootstrap values based on the ITS phylogeny as compared to the *rbcL* phylogeny.

Table 3-2. Sequence divergence of WH *Synechococcus* strains based on several commonly used genetic markers. N.A. : no data available.

Genes	WH7803 vs.WH7805 % identity (No. of nt)	WH7805 vs. WH8101 % identity (No. of nt)	WH5701 vs. WH8103 % identity (No. of nt)
16S rRNA	97% (1235)	96% (1234)	N.D.
rbcL	89% (630)	91% (627)	89% (530)
*rpo*C1	91% (559)	N.A.	86% (559)
ITS	91% (1006)	85% (1020)	71% (1101)

Distribution of picocyanobacterial genotypes in the Bay

Among six rRNA operon clone libraries, only three, constructed from the September samples, contained cyanobacterial genes. Absence of cyanobacteria in the March clone libraries is likely due to low picocyanobacteria cell density. In the Chesapeake Bay, picocyanobacteria in March are usually below 10^3 cells ml^{-1} and contribute less than 0.1% of the total prokaryotic community in terms of cell number (Chapter 6). For September samples, cyanobacterial genes represented 4.4, 14.3 and 8.0% of all clones analyzed in the upper, middle and lower Bay, respectively (Table 3-1). On average, 8.5% of clones from the Chesapeake Bay libraries contained picocyanobacterial rRNA operons. This frequency corresponded well with the 8.1% average fraction of picocyanobacterial cells within the total bacterioplankton community.

Twenty-five environmental clones (from a total of 263 clones) putatively identified as cyanobacteria by prescreening were subsequently identified as marine *Synechococcus* by DNA sequencing. This result indicates the library pre-screening method was accurate enough for the identification of cyanobacteria clones within rRNA operon clone libraries of marine bacterioplankton. All four clones from the upper Bay (Stn. 908) were members of the MC-B cluster. Among 12 clones from the mid Bay (Stn. 818), four were members of the MC-B cluster, seven members of the MC-A cluster and one representative in the *Cyanobium*. Among seven clones in the lower Bay, only one was affiliated with the MC-B cluster, and five were affiliated with the MC-A cluster (Fig. 3-1 and Table 3-1). However, most CB environmental clones in cluster MC-A were not closely related to WH strains isolated from coastal and oceanic waters. Overall, the ITS-based clusters of cyanobacterial strains appear to reflect the ecosystem from which the strain was first isolated.

None of the Chesapeake strains or clones clustered with *Cyanobium* strains isolated from Baltic Sea brackish waters (Ernst et al. 2003). Baltic Sea strains (BS5, BS6, BS8, and BS13) were isolated from the Bornholm Sea (salinity 9 ppt) and the Gotland Sea (salinity 7 ppt). Compared to the Baltic Sea samples, the water samples used for isolating Chesapeake strains had a much wider salinity range (ca. 5 to 25 ppt). Despite this wide range of salinity along the Bay, marine *Synechococcus* (MC-A and MC-B), not *Cyanobium* dominate the Chesapeake picocyanobacterial community.

ITS length heterogeneity and variation of G+C content

The length of the ITS among Chesapeake picocyanobacterial isolates and environmental clones varied widely from 753 to 875 nucleotides (nt) and 606 to 913 nt, respectively

(Table 3-3). The high variability in ITS length among *Synechococcus* allowed for accurate separation of closely related strains based on the library prescreening. The unique position of strains CB0104 and CB0203 in the ITS and *rbcL* phylogenies was also reflected in their longer 16S rRNA to tRNA alanine (ITS-A) sequences (Table 3-3). The ITS-A length for CB0104 and CB0203 is 291 and 271 nt, respectively, while the ITS-A length for the remainder CB and WH strains ranges from 173 to 198 nt. The length of tRNA isoleucine (74 nt) and tRNA alanine (73 nt) were highly conserved among the CB isolates. With the exception of strain CB0209, all the Chesapeake strains and environmental clones clustered in MC-B contained longer sequences between tRNA isoleucine and tRNA alanine (ITS-S spacer, 25 to 26 nt) than those in MC-A (9 nt).

The average G+C % of the ITS sequence for MC-B isolates and clones is 48.8±2.5% (n=19), which is lower than for *Cyanobium gracile* PCC6307 (54%); and higher than those of *Prochlorococcus* (38.6±2.0%) (Rocap et al. 2002) and MC-A isolates and clones (44.1±1.3%, n=21). Mean genomic DNA G+C base composition in picocyanobacteria varies dramatically (31-71%) and is an important character used for the taxonomy of cyanobacteria (Herdman et al. 2001). All the G+C% values in Table 3-3 are based on the nucleotide composition of ITS, and these values are lower than that those based on the genomic DNA. For example, the genomic G+C % content of MC-A (Cluster 5.1) and MC-B (Cluster 5.2) strains ranges between 55-62% and 63-66%, respectively. MC-C contains the lowest genomic G+C% at ~49% (Herdman et al. 2001).

Table 3-3. Comparison of different segments of ITS from Chesapeake *Synechococcus* isolates and clones, and other reference strains. ITS-A: spacer between 16S rRNA and $tRNA^{Ile}$; ITS-S: spacer between $tRNA^{Ile}$ and $tRNA^{Ala}$; ITS-B: spacer between $tRNA^{Ala}$ and 23S rRNA. The G+C content is calculated based on the full ITS sequence. N.M.: not Marine Cluster.

Strains or clones	ITS-A (nt)	tRNA-Ile (nt)	ITS-S (nt)	tRNA-Ala (nt)	ITS-B (nt)	Total (nt)	G+C (%)	Marine Cluster	Sources
Chesapeake Bay isolates (14 total)									
CB0101	192	74	26	73	477	842	50.0	B	This study
CB0102	192	74	26	73	510	875	48.6	B	This study
CB0103	191	74	26	73	408	772	50.3	B	This study
CB0104	291	74	9	73	411	858	52.9	N.M.	This study
CB0201	185	74	9	73	432	773	46.9	A	This study
CB0202	192	74	26	73	429	794	49.0	B	This study
CB0203	276	74	9	73	389	821	53.0	N.M.	This study
CB0204	192	74	26	73	497	862	49.8	B	This study
CB0205	173	74	25	73	408	753	48.4	B	This study
CB0206	192	74	26	73	498	863	50.1	B	This study
CB0207	192	74	26	73	498	863	49.7	B	This study
CB0208	192	74	26	73	471	836	49.1	B	This study
CB0209	192	74	55	73	443	837	47.6	B	This study
CB0210	192	74	26	73	476	841	50.2	B	This study
Chesapeake Bay environmental clones (22 total)									
CB01C11	174	74	25	73	409	757	47.9	B	This study
CB01C12	175	74	25	73	408	756	47.2	B	This study
CB01E02	174	74	26	76	406	758	47.2	B	This study
CB01F08	192	74	26	72	408	757	47.3	B	This study
CB11B02	181	74	9	73	399	738	46.2	A	This study
CB11C04	174	76	25	72	408	757	47.3	B	This study
CB11C11	172	74	25	73	389	735	48.2	B	This study
CB11D02	181	74	9	73	446	785	43.3	A	This study
CB11D06	172	74	25	73	405	751	47.8	B	This study
CB11D12	180	74	9	73	446	784	43.4	A	This study
CB11E03	181	74	9	73	401	740	45.3	A	This study
CB11F09	183	75	9	72	432	774	44.9	A	This study
CB11G04	180	74	9	73	446	784	43.4	A	This study
CB11G10	185	74	35	73	544	913	48.7	N.M.	This study
CB11H03	181	74	9	73	400	739	45.1	A	This study
CB11H07	183	74	9	73	431	772	46.4	A	This study
CB22A07	183	74	9	72	401	741	44.5	A	This study
CB22A09	181	74	9	73	383	722	47.0	A	This study
CB22C09	191	74	26	73	471	837	49.0	B	This study
CB22D04	185	75	9	73	432	776	44.2	A	This study
CB22G11	182	74	9	73	432	772	44.4	A	This study
CB22H05	182	74	9	73	441	781	42.0	A	This study
Other marine reference strains									
WH5701	198	74	9	73	504	858	48	N.M.	Rocap et al. 2002
WH8007	173	74	25	73	438	783	46.3	B	This study
WH8101	185	74	10	73	446	788	44.0	A	Rocap et al. 2002

The two Chesapeake isolates (CB0104 and CB0203) with high G+C% (53%, based on ITS) did not cluster within either MC-A or MC-B (Fig. 3-1). Whether these two strains should be included in genus *Cyanobium* in terms of their high G+C content is uncertain. In concordance with Ernst et al. (2003), our results also suggest that the current *Cyanobium* group contains several deeply branched subclusters equivalent to genera.

Although we did not use cyanobacteria-specific PCR primers to examine the genetic diversity of picocyanobacteria in the Bay, the bacteria-specific PCR primers allowed for a quick snapshot on the distribution of different cyanobacterial genotypes within the whole bacterial community. With the increasing availability of ITS sequences from picocyanobacteria, it is now possible to develop cyanobacteria-specific PCR primers targeted to the ITS region. Such primers will be an important tool for high-resolution analysis of the diversity and distribution of specific subgroups or ecotypes of cyanobacteria (Becker et al. 2002; Laloui et al. 2002; Ferris et al. 2003; Janse et al. 2003; Becker et al. 2004). Conserved and variable regions of cyanobacterial ITS are potential targets for the development of PCR primers or oligonucleotide probes at different taxonomic levels (Iteman et al. 2000). High sequence divergence of ITS may allow us to identify different *Synechococcus* hosts that are infected by different groups of cyanophages, and use this information for studying ecological interactions (i.e. killing the winner population) between *Synechococcus* and their phages.

Conclusion

The composition of picocyanobacterial population structure in the Chesapeake Bay is unique in many respects: 1) it contains several novel clusters of *Synechococcus* that are distinct from freshwater and oceanic picocyanobacteria; 2) it likely reflects

ecological adaptation to a complex estuarine ecosystem; 3) many strains and environmental clones are related to Marine Cluster B *Synechococcus*; and 4) the vast majority of Chesapeake Bay picocyanobacteria appeared to be of marine rather than freshwater origin. By adding sequences of many estuarine organisms to the phylogenetic reconstruction of picocyanobacteria, our study provides a new insight into the classification of picocyanobacteria, particularly of marine *Synechococcus*.

Chapter 4: Isolation and characterization of cyanophages infecting Chesapeake Bay *Synechococcus*

Abstract

Cyanophages that infect coastal and oceanic *Synechococcus* have been studied extensively. However, no cyanophages infecting estuarine *Synechococcus* have been reported. In this study, seven cyanophages (three podoviruses, three siphoviruses and one myovirus) isolated from four estuarine *Synechococcus* strains were characterized in terms of their morphology, host range, growth and genetic features. All the podoviruses and siphoviruses were highly host specific. For the first time, the photosynthesis gene (*psbA)* was found in two podoviruses infecting estuarine *Synechococcus*. However, the *psbA* gene was not detected in the three siphoviruses. The *psbA* sequences from the two *Synechococcus* podoviruses clustered with some environmental *psbA* sequences, forming a unique cluster distantly related to previous known *psbA* clusters. Our results suggest that the *psbA* among *Synechococcus* podoviruses may evolve independently from the *psbA* of *Synechococcus* myoviruses. All three estuarine *Synechococcus* podoviruses contained the DNA polymerase (*pol*) gene, and clustered with other podoviruses that infect oceanic *Synechococcus* and *Prochlorococcus*, suggesting that the DNA *pol* is conserved among marine picocyanobacterial podoviruses. Prevalence of host-specific cyanophages in the estuary suggests that *Synechococcus* and their phages in the estuarine ecosystem may

develop a host-phage relationship different from what has been found in the open ocean.

Introduction

Unicellular cyanobacteria of the genus *Synechococcus* are ubiquitous and abundant photosynthetic picoplankton in a wide range of marine environments (Johnson and Sieburth 1979; Waterbury et al. 1979). Many *Synechococcus* spp. have been isolated from coastal and oceanic waters, and identified as marine cluster A (MC-A) *Synechococcus* (Waterbury and Rippka 1989), or cluster 5.1 (Herdman et al. 2001). In contrast, *Synechococcus* living in the estuarine environments have been poorly characterized until recently. A unique group of *Synechococcus* distantly related to oceanic *Synechococcus* has been isolated from the Chesapeake Bay estuary (Chapter 2). These estuarine *Synechococcus* belong to marine cluster B *Synechococcus* or cluster 5.2, and dominate *Synechococcus* populations from the northern to middle parts of the Chesapeake Bay (Chapter 3). The estuarine MC-B *Synechococcus* (referred to estuarine *Synechococcus* hereafter) appear to be adapted to the distinctive ecological niche in the estuary. The coastal and oceanic MC-A *Synechococcus* (referred to oceanic *Synechococcus* hereafter) were not commonly found in the upper Chesapeake Bay using both culture and molecular methods (Chen et al. 2006a).

Synechophages (viruses that are isolated using *Synechococcus*) are known to be active and abundant pathogens that directly impact the distribution and species composition of *Synechococcus* in the aquatic environment (Proctor and Fuhrman 1990; Suttle et al. 1990; Suttle 2000; Mann 2003). Synechophage titers have been

found to be tightly coupled with *Synechococcus* abundance in marine environments (Waterbury and Valois 1993; Suttle 1994; Suttle and Chan 1994; Sullivan et al. 2003). Our current knowledge on synechophages is mainly built on those infecting oceanic *Synechococcus*. Synechophages belonging to the three major tailed phage families (*Myoviridae*, *Podoviridae* and *Siphoviridae*) have been isolated from oceanic *Synechococcus*, and myoviruses were the dominant type (up to 80%) among phage isolates (Suttle and Chan 1993; Waterbury and Valois 1993; Wilson et al. 1993; Lu et al. 2001; Marston and Sallee 2003). Typically, cyanomyoviruses have a broad host range (polyvalent), and are able to cross infect different oceanic *Synechococcus* strains (Waterbury and Valois 1993; Lu et al. 2001), or even marine *Prochlorococcus* (Sullivan et al. 2003). Only a few podoviruses and siphoviruses infecting oceanic *Synechococcus* have been reported (Suttle and Chan 1993; Waterbury and Valois 1993; Chen and Lu 2002). Interestingly, podoviruses were the dominant phage type isolated from high-light-adapted marine *Prochlorococcus* (Sullivan et al. 2003). All known podoviruses that infect oceanic *Synechococcus* and *Prochlorococcus* are lytic and host specific (Suttle and Chan 1993; Waterbury and Valois 1993; Chen and Lu 2002; Sullivan et al. 2003). In general, siphoviruses infecting marine picocyanobacteria are underrepresented in the current cyanophage collections.

Currently, the genome sequences from seven marine cyanophages (include four *Synechococcus* phages and three *Prochlorococcus* phages) have been reported (Chen and Lu 2002; Mann et al. 2005; Sullivan et al. 2005; Pope et al. 2007; Weigele et al. 2007). The genome sequencing revealed that many marine cyanophages contain the photosynthetic genes *psbA* and *psbD*, which encode photosystem II core protein D1

and D2, respectively (Mann et al. 2003; Lindell et al. 2004; Millard et al. 2004; Sullivan et al. 2006). The phage-encoded *psbA* gene is functional upon infecting host (Lindell et al. 2005; Clokie et al. 2006), and the expression of viral *psbA* can be detected in the sea (Sharon et al. 2007). Diverse *psbA* gene sequences have been found in natural marine environments and a large portion of *psbA* sequences was related to known cyanophages (Zeidner et al. 2005; Sullivan et al. 2006; Bench et al. 2007; Sharon et al. 2007). Recently, it was speculated that a unique *psbA* clone BAC9D04 recovered from the Red Sea could be related to a *Synechococcus* podovirus, but it lacks a culture identity (Zeidner et al. 2005).

Our inter-annual (2002-2007) survey in the Chesapeake Bay (USA) revealed that both *Synechococcus* and cyanophage could be extremely abundant in summer (up to 3×10^6 *Synechococcus* cells ml^{-1}, and 6.2×10^5 cyanophage MPN ml^{-1} for *Synechococcus* WH7803 (Chapter 6). Interestingly, the most abundant phage clones obtained by MPN assay (i.e. end-point phage lysates) were usually either podoviruses or siphoviruses (Chapter 6). This observation is consistent with the high frequency (14%) of cyanopodovirus-like homologs found in the Chesapeake Bay viral metagenomic database (Bench et al., 2007). *Synechococcus* adapted in the Chesapeake Bay estuary are a unique group of marine picocyanobacteria (Chen et al., 2004; 2006a), but no cyanophages have been isolated from these abundant and ecologically important *Synechococcus* prior to our work.

Here, we report the isolation and characterization of seven cyanophages that infect different strains of estuarine MC-B *Synechococcus*. Our goal is to gain a better

understanding of biological and ecological interactions between estuarine *Synechococcus* and their phages.

Materials and Methods

Water samples

Samples were collected from the surface waters of Baltimore Inner Harbor (Pier 5) and Chesapeake Bay on board the R/V *Cape Henlopen* during the research cruises for the MOVE (Microbial Observatory of Virioplankton Ecology) project from September 2002 to July 2004 (Table 4-1). Viral concentrates (VCs) were prepared on board by ultrafiltration and stored as described previously (Wang and Chen 2004). The *Synechococcus* cells were counted as previously described (Chapter 2).

Synechococcus strains

Eleven estuarine *Synechococcus* strains affiliated with MC-B cluster (CB0101, CB0102, CB0103, CB0202, CB0204, CB0205, CB0206, CB0207, CB0208, CB0209 and CB0210) were isolated from the Chesapeake Bay (Chen et al. 2004). They were grown in SN medium (Waterbury and Willey 1988) with 15‰ salinity, at 25 °C under 14/10 hr of light: dark cycle with 10-20 μE (microeinsteins) m^{-2} s^{-1} illumination throughout this study.

Isolation of cyanophage

Two VCs prepared from Baltimore Inner Harbor on June 4, 2001 and July 16, 2002, respectively) and eight VCs obtained from middle Bay station 804 (38˚04' N, 76˚13' W) during the MOVE cruises (September 26-30, 2002; March 1-6; April 4-10, June 1-4, August 21-25 in 2003; and March 20-24, May 6-11 and July11-15 in

2004) were screened against 11 estuarine *Synechococcus* strains for isolating cyanophages. A top agar overlay method (Wilson et al. 1993) was used for isolation and purification of cyanophages and phage titering. Cyanophage titers were measured from triplicate plates using the original host strains.

Cyanophage burst size and latent period determination

A one-step growth curve was measured following the method described by Jiang et al. (1998) with some modifications. In our modified protocol, direct counting of released viral particles instead of conventional plaque assay method was used to avoid tedious plating work and shorten the observation time. The phage lysates (ca. 10^9 to 10^{11} PFU ml^{-1}) were inoculated into the exponentially growing host cultures (ca. 2 x 10^7 cells ml^{-1}, with a doubling time between 16 h to 24 h, as measured by cell counts) with a multiplicity of infection (MOI) of 1.2 to 1.5 for 1 h at 25°C. The control received the same amount of microwave-killed phage lysates. The mixture was then diluted in fresh SN medium (Waterbury et al. 1986) by 100 fold to minimize the further adsorption of phage to host cells. After inoculation (taken as T0), a subsample of suspension was withdrawn from each culture periodically for up to six days. The released virus like particles (VLPs) were stained by SYBR Gold and counted using the protocol described by Chen et al. (2001). The latent period of each phage isolate was estimated by the time interval of first wave of significant increase in VLPs (3 fold increase as threshold) observed in the samples. The burst size was estimated by the following formula:

$$B=dV/dH$$

where *B* represents the burst size, *dV* is the maximum increase in VLPs counts in the sample while *dH* is the reduction in host cell abundance. V and H strand for viral and host cell counts, respectively.

It was shown that viral latent time and burst size are sensitive to host physiology (Wilson et al. 1996; You et al. 2002) and host density (Abedon et al. 2003). In our experiments, the exponentially growing *Synechococcus* cells were obtained by continuous transferring host cultures to the fresh SN media at least two passages before they were used. The VLPs in the controls were less than 1.7×10^4 ml^{-1} throughout the experiments, and were much less than those in the treatments (ranged from 1.3×10^5 to 4.1×10^7 VLPs ml^{-1}).

Host range

Eleven estuarine *Synechococcus* strains, two *Cyanobium* strains CB0104 and CB0203 and four oceanic stains (CB0201, WH7803, WH7805 and WH8101) were screened for possible cross-infection by isolated cyanophages. The lysate of clonal cyanophage isolate (10 μl of 10^9 PFU ml^{-1}) was added to 0.25 ml of exponentially growing host cultures in 96-well microtiter plates, while controls received only SN medium. All plates were placed under the same culturing conditions as described above, and monitored daily for cell lyses. Triplicates for each phage-host combination were performed.

Cyanophage purification and DNA extraction

Clonal cyanophage isolates were propagated in 1-liter host cultures. Upon lysis, phage particles in the lysates were harvested and purified with CsCl as previously described (Wilson et al. 1993). The phage DNA was then extracted using

phenol and chloroform method (Maniatis et al. 1982). The phage DNA pellets were dissolved in TE buffer (10 mM Tris, 1 mM EDTA, pH 7.4) and stored at 4°C.

Transmission electron microscopy (TEM)

Phage particles were mounted the 200-mesh Formvar/Carbon coated copper grids followed by staining with 1% aqueous uranyl acetate and then examined with a Zeiss CEM902 transmission electron microscope. The TEM images of cyanophages were acquired using a Mega ViewII camera and dimensions of phage structure were analyzed using the analySIS software (Soft Imaging System Corp.).

PFGE and restriction digestion of viral genome

Pulsed-field gel electrophoresis (PFGE) analysis, including gel plug preparation and proteinase K treatment, followed the method described by Wommack et al. (1999b). PFGE was performed using a clamped homogeneous electric field system (CHEF DR-III, Bio-Rad, Richmond, Calif.) under the following conditions: 1% agarose in 1× Tris-borate-EDTA (TBE) gel buffer (90 mM Tris-borate, 1 mM EDTA; pH 8.0), 0.5 × TBE tank buffer, 1 to 12 s pulse ramp, 6.0 V cm^{-1} with an included angle of 120° at a constant temperature of 14°C, and a run time of 16 h. Genomic DNA of cyanophages in the gel plugs was digested overnight with three restriction endonucleases: *Acc* I, *Xho* I and *Cfo* I (Invitrogen) respectively. The digested viral genomic DNA was separated by gel electrophoresis (1% agarose, 0.5 × TBE buffer, 100 V for 6 h). Gels were stained with SYBR Gold (Invitrogen) according to the manufacturer's instructions and were visualized with Kodak EDAS 290 gel documentation system (Eastman Kodak Company, New Haven, CT).

PCR amplification of cyanophage *psbA*, *pol* and *g20* genes

The partial photosynthetic protein D1 gene (*psbA*) was amplified using our newly designed primers psbA-93F and psbA-341R. This primer set was designed based on the conserved sequence regions of *psbA* gene from eight picocyanobacteria strains and nine cyanophage isolates (Fig. 4-1). To amplify the *psbD* gene, we used the primer sets and protocols described by Sullivan et al. (2006). To characterize cyanophage DNA polymerase gene (*pol*), two novel degenerate PCR primer sets were designed based on the conserved amino acid sequences of nine podoviruses and four environmental podovirus sequences. Since the first primer set Podo-85F/-539R was only able to amplify S-CBP2 but none of the other isolates. The second primer set Podo-260F/-387R was then designed to amplify the cyanopodovirus isolates (Fig. 4-2). Based on the sequence information from PCR-amplified fragments, the full sequences of cyanopodoviruses DNA *pol* gene were obtained by primer walking from both directions. The DNA templates for primer walking were prepared with a GenomiPhi V2 DNA amplification kit (GE Healthcare, UK), according to the manufacturer's protocol. The partial cyanomyovirus viral capsid assembly gene (*g20*) was PCR-amplified as described previously (Wang and Chen 2004). Sequence information and annealing temperature of PCR primers used in this study were summarized in Table 4-1. All PCR reactions were performed either in 25-μl or 50-μl volume containing 1 × reaction buffer (Promega) with 1.5 mM $MgCl_2$, 100 μM dNTPs, 10 pmole of each primer, 1 unit of *Taq* DNA polymerase (Promega, Madison, WI) and 5-10 ng phage DNA as templates. The PCR program for all reactions included an initial denaturing at 94°C for 3 min, followed by 30 cycles of 94°C for 1

min, annealing at different temperatures for different primer sets used (Table 4-2) for 30 s, and 72°C for 1 min. The final extension of PCR amplification was at 72°C for 10 min and all the PCR reactions were concluded at 4°C. PCR fragments were gel-purified before they were sequenced bi-directionally on an automated sequencer ABI310 (PE Applied Biosystems) at Center of Marine Biotechnology, University of Maryland Biotechnology Institute.

Table 4-1. PCR primers used in this study.

Target Gene	Primer Name	Primer Sequence (5'→ 3')	Corresponding Amino Acid sequence	PCR Amplicon Size (bp)	Annealing Temp. (°C)	Source
psbA	psbA-93F	TAYCCNATYTGGGAAGC	YPIWEA	745	55	This study
	psbA-341R	TCRAGDGGGAARTTRTG	HNFPLD			
psbD	psbD-26Fa	TTYGTNTTYRTNGGNTGGAGYGG	FVFV/IGWSG	N.D.	51	Sullivan et al., 2006
	psbD-26Fb	TTYGTNTTYRTNGGNTGGTCNGG	FVFV/IGWSG			
	psbD-54Fa	GTNACNAGYTGGTAYACNCAYGG	VTSWYTHG			
	psbD-54Fb	GTNACNTCNTGGTAYACNCAYGG	VTSWYTHG			
	psbD-308Ra	YTCYTGNGANACRAARTCRTANGC	AYDFVSQE			
	psbD-308Rb	YTCYTGRCTNACRAARTCRTANGC	AYDFVSQE			
pol	Podo-85F	GAYACNYTBRTRCTSTC	DTLV/ILSRL	1360	52	This study
	Podo-539R	TCRTCRTGHAYMMABGC	AF/WV/IHDE			
	Podo-260F	CGNSABMACATHGCVTGG	RH/D/QH/NIAW	381	52	
	Podo-387R	TARTGVSCVAGVMDHCKBARYTC	ELRC/L/MLG/AH			
g20	CPS1	GTAGWATTTTCTACATTGAYGTTGG	RIFYIDV	590	46	Zhong et al., 2002
	CPS8	AAATAYTTDCCAACAWATGGA	PY/FVGKY			

N.D.: No data.

```
                                    psbA-93F→                ← psbA-341R
                          5′-TAYCCNATYTGGGAAGC-3′  3′-GTRTTRAAGGGDGARCT-5′
Synechococcus WH8102      TATCCCATCTGGGAAGC--------CACAACTTCCCCCTCGA
Synechococcus WH7803      TATCCCATCTGGGAAGC--------CACAACTTCCCCCTCGA
Synechococcus RS9901      TATCCCATCTGGGAAGC------------------------
Synechococcus elongatus   TACCCCATTTGGGAAGC--------CACAACTTCCCCCTCGA
Prochlorococcus CCMP1375 TACCCAATTTGGGAAGC--------CATAATTTCCCTCTTGA
Prochlorococcus MED4      TACCCAATTTGGGAAGC--------CACAACTTCCCACTTGA
Prochlorococcus MIT9313   TATCCCATTTGGGAAGC--------CATAACTTCCCTCTCGA
Prochlorococcus NATL1A    TACCCAATCTGGGAAGC------------------------
S-WHM1                    TATCCCATCTGGGAAGC--------CACAACTTCCCTCTCGA
S-PM2                     TATCCCATCTGGGAAGC--------CACAACTTCCCTCTTGA
S-RSM88                   TATCCCATCTGGGAAGC--------CACAACTTCCCTCTTGA
S-BM4                     TATCCCATCTGGGAAGC--------CACAACTTCCCTCTCGA
S-RSM2                    TATCCCATCTGGGAAGC--------CACAACTTCCCACTCGA
P-SSP7                    TATCCGATCTGGGAAGC--------CATAACTTCCCACTTGA
P-SSM2                    TATCCTATCTGGGAAGC--------CACAACTTCCCACTTGA
P-RSM28                   TATCCCATCTGGGAAGC--------CACAACTTCCCTCTCGA
P-SSM4                    TATCCCATCTGGGAAGC--------CACAACTTCCCACTAGA
```

Fig. 4-1. DNA sequence alignment of *psbA* gene for PCR primer design. Consensus sequences were underlined and shown in the bold type. The primer name is numbered according to the corresponding amino acid location in the reference gene of *Synechococcus* WH8102.

```
                          Podo-85F→      Podo-260F→    ←Podo-381R    ←Podo-534R
                          DTLV/ILSRL     RH/D/QH/NIAW  ELRC/L/MLG/AH AF/WV/IHDE
Cyanophage P60    --DTLILSRL--------RHHIAW-------ELRCLGH-------AFVHDE--
Cyanophage S-CBP2 ------------------RDHIAW-------ELRCLAH--------------
GOM-W-I(AY258466) --DTLVLSRL--------RDHIAW-------ELRCLAH--------------
SOG-S-G(AY258464) --DTLVLSSL--------RQNIAW-------ELRLLGH--------------
SOG-S-L(AY258465) --DTLVMSRL--------RQNIAW-------ELRLLGH--------------
SOG-S-D(AY258463) --DTLVLSRL--------RQNIAW-------ELRLLGH--------------
Cyanophage P-SSP7 ------------------RDHIAW-------ELRMLAH-------AFVHDE--
Coliphage T3      --DTLVLSRL--------RDHIQK-------ELRCLAH-------AWIHDE--
Coliphage T7      --DTLVLSRL--------RDHIQK-------ELRCLAH-------AWVHDE--
Phi-YeO3-1        --DTLVLSRL--------RDHIQK-------ELRCLAH-------GWIHDE--
YePhiA1122        --DTLVLSRL--------RDHIQK-------ELRCLAH-------AWVHDE--
Phage gh-1        --DTLVLVRL--------RPHIIK-------ELRCLGH-------AWVHDE--
Roseophage SIO-1  ---VIGFDIP--------PAQVNK-------QLRVLAH-------AWVHDE--
```

Fig. 4-2. Amino acid sequence alignment of DNA *pol* gene for PCR primer design. Consensus sequences were underlined and shown in the bold type. The primer name is numbered according to the corresponding amino acid positions in the reference gene of cyanopodovirus P60.

Phylogenetic analysis

Sequences were aligned using Mac Vector 7.2 program (GCG, Madison, WI.). Phylogenetic reconstruction was performed by using MEGA 4.0 software (Tamura et al. 2007). Maximum Likelihood (ML) analysis was performed using PHYLIP program (Felsenstein 1993). For deduced amino acid sequences of DNA *pol* gene and g20 gene, evolutionary distances were calculated by using neighbor-joining (NJ) method under the Jones-Taylor-Thornton (JTT) model with rate variation among sites and complete deletion for gaps. To obtain a better resolution for the *psbA* gene phylogeny, the DNA sequences instead of the deduced amino acid sequences were used for phylogenetic reconstruction (Zeidner et al. 2005). Sequences detected with strong signal for intragenic recombination (Sullivan et al. 2006) and the intron sequence of S-PM2 were excluded from phylogenetic analysis. To accommodate the compositional heterogeneity (particularly highly variable GC% at 3rd codon positions) among the sequence dataset, a modified LogDet method (Tamura and Kumar 2002) and Maximum Composite Likelihood (MCL) method (Tamura et al. 2007) were used to construct the phylogenetic tree, under the mode of heterogeneous substitution pattern among lineages and gamma correction [α=2.07, from Zeidner et al. (2005)] for rate variation among sites. The topologies of inferred trees obtained by the two methods were essentially the same, and the final tree was inferred from MCL analyses. Bootstrap resamplings of the NJ, MP, ML, LogDet and MCL were performed to obtain the confidence estimates for inferred tree topologies. 1000 resampling were carried out for the NJ, MP, LogDet and MCL analyses, while 100 replicates were tested for ML analysis.

Results and Discussion

Isolation and morphological description of estuarine *Synechococcus* phages

After challenging 11 estuarine *Synechococcus* strains (all belong to Marine cluster B lineage) with 10 different viral assemblages collected from the Chesapeake Bay estuary, four host strains (CB0101, CB0202, CB0204 and CB0208) were sensitive to phage infection as evident by plaque assay. Among them, *Synechococcus* CB0101 appeared to be most susceptible to the infection by various types of phage (Table 4-2). In contrast, for the remaining three CB strains, only one particular phage type (either podo- or siphovirus but not myovirus) could be isolated (Table 4-2). It is noteworthy that only the viral assemblages collected from the warm seasons (ca.17-27°C water temperature) resulted in visible plaques for the estuarine *Synechococcus*. It usually took 3 to 10 days for visible plaques to appear on the host *Synechococcus* lawn. Among the seven phages obtained, one (S-CBM2) is a myovirus, three are podoviruses (S-CBP1, S-CBP2 and S-CBP3) and three are siphoviruses (S-CBS2, S-CBS3 and S-CBS4) (Fig. 4-3). The three podoviruses yielded clear plaques with diameter greater than 5 mm; three siphoviruses produced plaques with diameter from 2 to 5 mm, and with either clear or fuzzy edges; and the myovirus gave the smallest but clear plaques (1-1.5 mm in diameter).

Table 4-2. Description of seven cyanophages of estuarine *Synechococcus* strains.

*Cyanophage clone	Host strains	Capsid Size (mean ± SD, nm)	Genome size (Kb)	Latent Period (h)	Burst size	Source of Viral concentrates (VCs)	*Synechococcus* cell density (ml^{-1})	*Phage Titer ($PFU\ ml^{-1}$)
S-CBM2	CB0101	90±5	180	15-18	28	Baltimore Inner Harbor (June 2001)	2.20E+03	6±3
S-CBP1	CB0101	52±3	48	6-8	86	Baltimore Inner Harbor (July 2002)	8.51E+05	240±33
S-CBP2	CB0208	55±5	48	8-10	92	Stn. 804 (September 2002)	2.49E+05	23±8
S-CBP3	CB0101	55±4	48	8-10	75	Stn. 858 (July 2004)	1.23E+06	72±6
S-CBS2	CB0204	90±3 × 50±2	70	24-36	65	Stn. 804 (September 2002)	2.49E+05	27±4
S-CBS3	CB0202	56±4	30	24-36	175	Stn. 804 (June 2003)	1.59E+04	18±3
S-CBS4	CB0101	72±4	65	24-36	57	Stn. 804 (July 2004)	1.44E+06	320±26

*Phage nomenclature: S: *Synechococcus* phages; CB: Chesapeake Bay; M: Myovirus; P: Podovirus; S: Siphovirus.
*Phage titer was estimated from triplicate plates using the original *Synechococcus* host strain for isolation.

Induction of estuarine *Synechococcus* strains with mitomycin C, following the method described by Chen et al. (2006b), did not yield detectable viral like particles. This result is consistent with no inducible lysogeny in marine *Synechococcus* observed by Waterbury and Valois (1993). In addition, no intact prophage genome was identified in both coastal and oceanic *Synechococcus* genomes (Palenik et al. 2003; Palenik et al. 2006). On the other hand, other studies have reported detectable lysogeny in natural *Synechococcus* populations (McDaniel et al. 2002; McDaniel and Paul 2005; Hewson and Fuhrman 2007). Given the fact that *Synechococcus* are relatively easy to cultivate, it remains puzzling why the culture based lysogenic induction does not support what has been observed in the natural waters. Perhaps

Synechococcus cultivated or maintained at room temperature have already lost prophage? Further studies should be done to understand such a controversy.

The specific phage titers of four estuarine *Synechococcus* strains could vary from 6 to 320 PFUs (plaque forming unit) ml^{-1} in the warm season (Table 4-2), but remained undetectable in wintertime. In contrast, cyanophage titers of the oceanic *Synechococcus* strain WH7803 in the same samples could be at least 100 fold higher (Chapter 6), indicating that sensitivity to phage infection varies with host types.

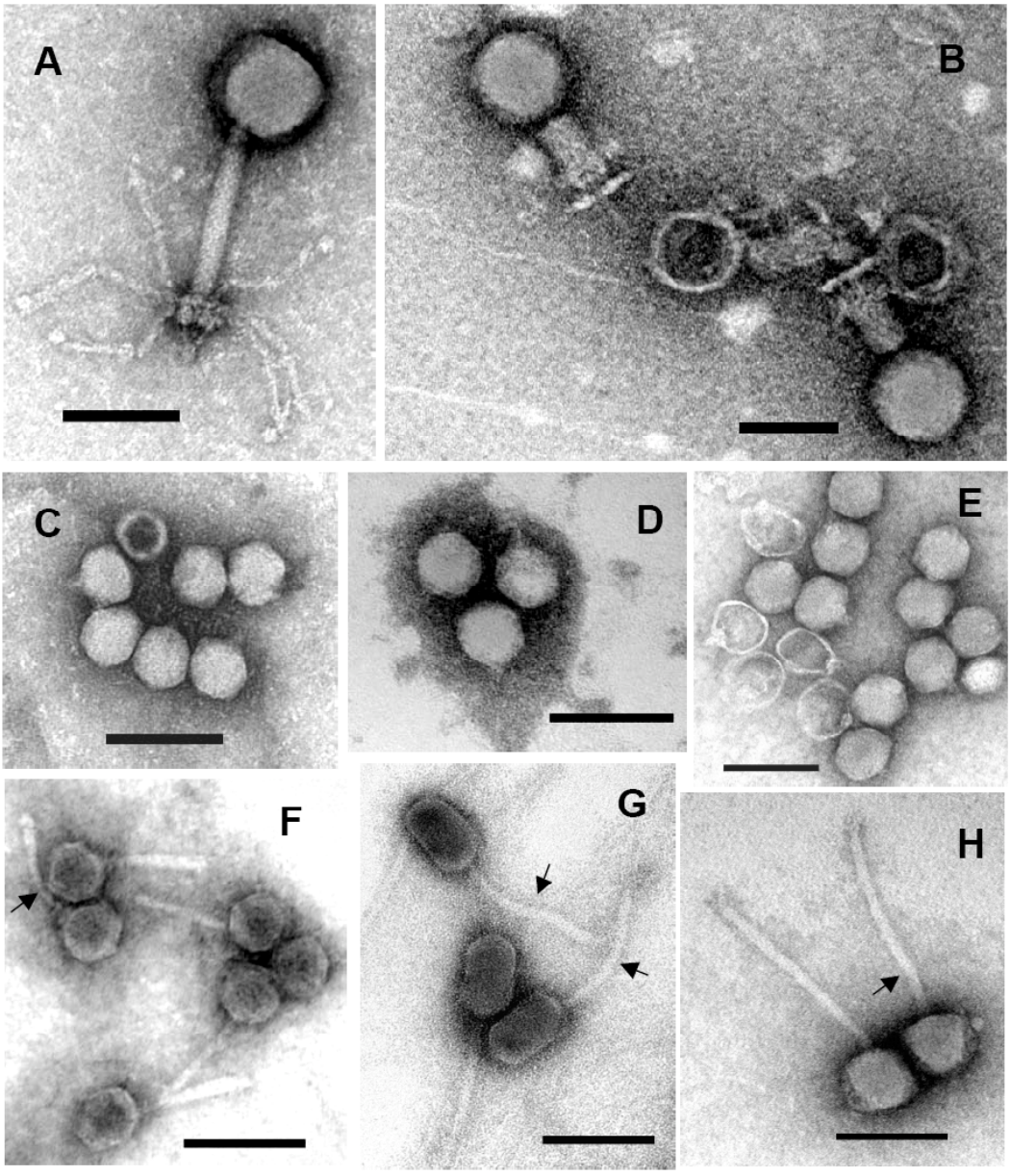


Fig. 4-3. Transmission electron micrographs of seven negatively stained estuarine *Synechococcus* phages. Myovirus S-CBM2 with non-contracted tail (A) and contracted tail (B). Short-tailed podovirus S-CBP1(C), S-CBP3 (D) and S-CBP2 (E). Non-contractile siphovirus S-CBS3 (F), S-CBS2 (G) and S-CBS4 (H), respectively. The bar is equivalent to 100 nm in all panels. Arrows indicate the flexible tails in panel F, G and H.

Host specificity of estuarine *Synechococcus* phages

All the podoviruses and siphoviruses were exclusively host specific. In contrast, the myovirus S-CBM2 could cross infect between MC-A and MC-B *Synechococcus* strains (Fig. 4-4). The estuarine *Synechococcus* strains used for the cross-infectivity assay are closely related based on ITS sequence phylogeny (as high as 98% sequence identity between CB0101 and CB0210) (Chen et al. 2006a). However, none of the podo- and siphoviruses could lyse other estuarine *Synechococcus* rather than their original host. Ecological adaptation of hosts could influence the distribution of different viral types. A recent study also found that only host-specific podoviruses were isolated from high-light-adapted *Prochlorococcus* in the ocean (Sullivan et al. 2003). Our results indicated that host specific phage (podo- or siphovirus) could be common to *Synechococcus* adapted to the estuary. It is likely that different ecotypes of marine picocyanobacteria (i.e. estuary-adapted, ocean-adapted, high-light adapted, or low-light adapted) select for the certain type(s) of co-existing cyanophages in a given habitat.

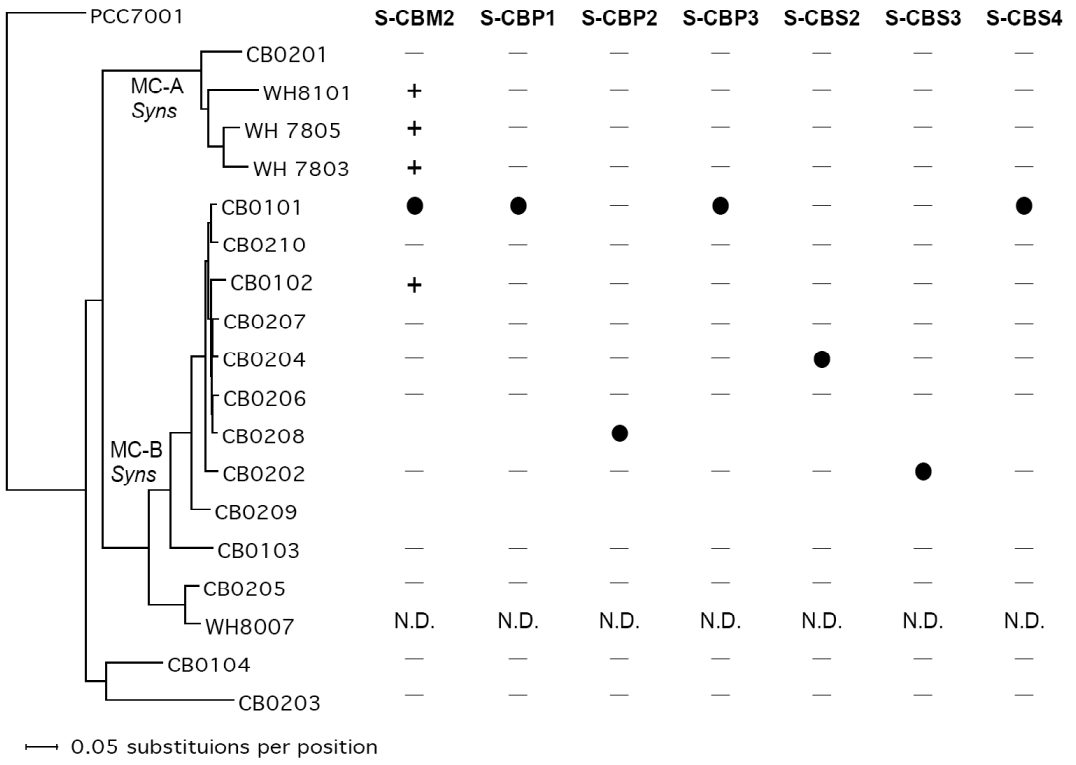


Fig. 4-4. Cross infectivity of seven estuarine *Synechococcus* phages tested with three Woods Hole *Synechococcus* strains (with "WH" prefix) and 14 Chesapeake Bay *Synechococcus* isolates (with "CB" prefix). Filled circle (●) represents the original host strain used for isolation. Plus symbol (+) indicates positive cross-infection of cyanophage with another host strain. Minus symbol (-) shows no infection. The phylogenetic relationship of 19 *Synechococcus* strains were constructed based on 16S-23S rDNA ITS spacer sequences (Chen et al. 2006a). N.D.: no data.

Genome size and genetic fingerprints

The genome sizes of these synechophages varied from ca. 30 to 180 kb (Table 4-1 and Fig. 4-5). The three podoviruses had similar genome sizes (ca. 48 kb), while the genome size of three siphoviruses was more variable (30-75 kb). Although the genomes of the three podoviruses had a similar size, their genetic fingerprints were distinct from each other (Fig. 4-6), indicating that they were different phages. Among the three restriction enzymes (*Acc* I, *Xho* I and *Cfo* I) tested, only *Acc* I was able to digest all seven synechophage genomes and yield distinctive restriction patterns (Fig. 4-6).

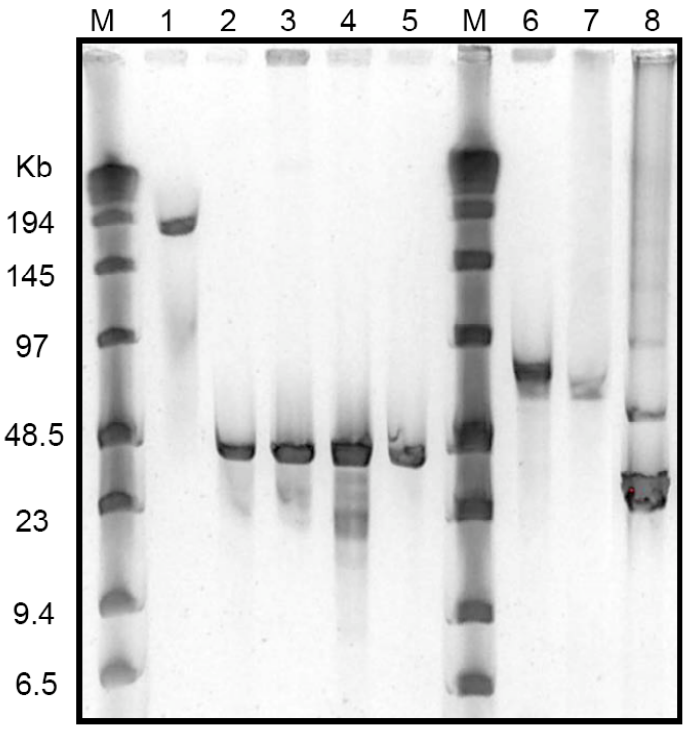


Fig. 4-5. PFGE analysis of estuarine *Synechococcus* phages. Lane 1 to 8: S-CBM2, S-CBP1, S-CBP2, S-CBP3, P60, S-CBS2, S-CBS4 and S-CBS3. Concatemers comprised of multiple phage genomes were evident for a siphovirus S-CBS2. Cyanopodovirus P60 with known genome size of 47.9 kb (lane 5) was included as an experimental control. Lane M: Low range PFGE marker (New England Biolabs).

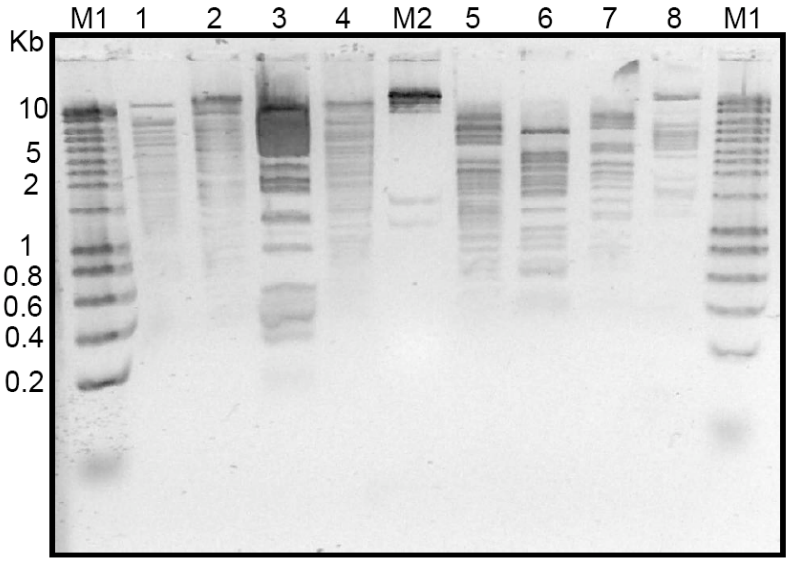


Fig. 4-6. *Acc* I restriction patterns of estuarine *Synechococcus* phages. Lane 1 to 8: S-CBM2, S-CBP1, S-CBS3, S-CBS2, S-CBP2, S-CBP3, P60 and S-CBS4. Lane M1: 1-10 kb ladder (Qiagen); Lane M2: λ/Hind III ruler (Qiagen)**.**

Growth and burst sizes of estuarine *Synechococcus* phages

The latent period of these seven synechophages ranged from 6 to 36 h (Table 4-2 and Fig. 4-7). The three podoviruses (S-CBP1, S-CBP2 and S-CBP3) have significantly shorter latent periods (ca. 6-8 h) compared to myovirus S-CBM2 and three siphoviruses (S-CBS2, S-CBS3 and S-CBS4). The latent period of these three *Synechococcus* podoviruses is similar to that of *Prochlorococcus* podovirus P-SSP7 (8h) reported by Lindell et al. (2005). In contrast, podovirus P60, which infects an oceanic *Synechococcus* has a much shorter latent period (less than 1.5 h, Fig. 4-7A), indicating that the virulence could vary among the cyanopodoviruses. The burst size of these estuarine *Synechococcus* phages varied from 28 to 175 (Table 2 and Fig. 4-7A). Myovirus S-CBM2 had the smallest burst size of 28, similar to the burst size (ca.

22) of an oceanic *Synechococcus* myovirus S-PM2 (Wilson et al. 1996). The burst size of the three podoviruses (75 to 92) was similar to the podovirus P60 burst size (80), despite the difference in their latent time. The burst size of the three siphoviruses varied from 57 to 175. Siphovirus S-CBS3 had a large burst size of ca. 175, probably due to its small genome size (ca. 30 kb). In contrast, myovirus S-CBM2 had the largest genome size (~180kb), but produced the smallest viral burst size among these cyanophages. It appears that the burst size of these synechophages has inverse relationship with their genome size (Fig. 4-7B). The product of synechophage genome size and its burst size (3.6-5.3 Mb) is roughly 1.5 to 2.0 fold of *Synechococcus* genome size (ca. 2.5 Mb). These observations further support the correlation between viral burst size and their host genome size (Brown et al. 2006).

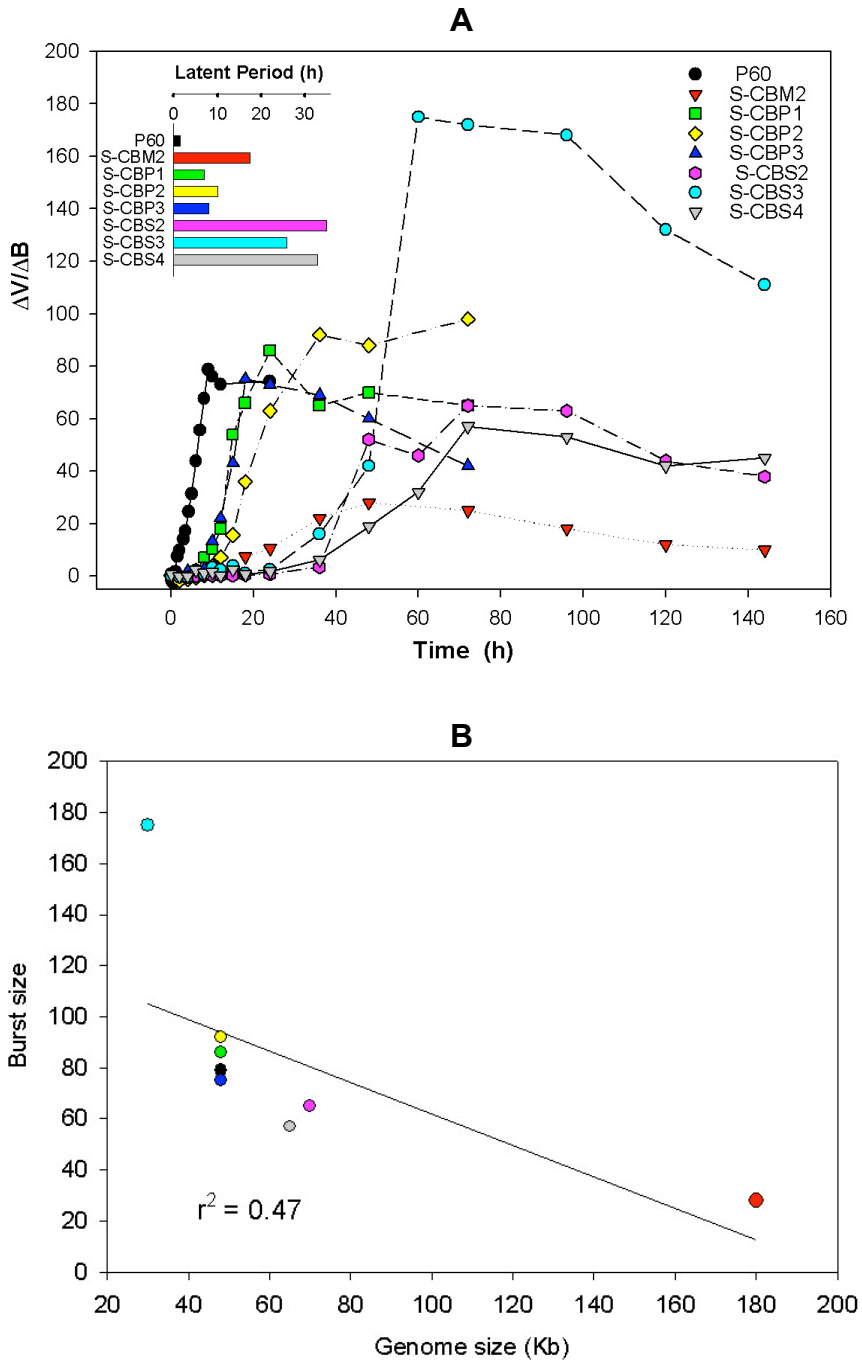


Fig. 4-7. A: One-step growth curves of eight cyanophages. The latent period of each phage isolate (upper left in the figure) was estimated by the time interval of first wave of increasing VLPs observed in the samples. The Y-axis was plotted by ΔV/ΔB (ΔV= Vt-V0, while ΔB=B0-Bt), as estimated by the ratio between increases in VLPs counts (ΔV) to the decreases of host cell counts (ΔB) at given time. *Synechococcus* podovirus P60 was included for comparison.
B: Relationships between burst size and genome size of eight cyanophages.

Identification of cyanophage signature genes

The cyanobacterial photosynthetic core gene (*psbA*) was detected in two podoviruses (S-CBP1 and S-CBP3) and one cyanomyovirus (S-CBM2). The *psbA* gene has been found in many *Synechococcus* and *Prochlorococcus* myoviruses (Mann et al. 2003; Lindell et al. 2004; Millard et al. 2004; Sullivan et al. 2006), and *Prochlorococcus* podoviruses (Lindell et al. 2004; Sullivan et al. 2005; Sullivan et al. 2006). Here, we first reported the presence of *psbA* gene in the cultivated *Synechococcus* podoviruses. Interestingly, the two *Synechococcus* podovirus *psbA* sequences clustered with several environmental *psbA* sequences as a group, which is distantly related to the *psbA* sequences from marine *Synechococcus*, *Prochlorococcus* and their known cyanophages (Fig. 4-8). Among the 10 environmental *psbA* sequences clustered with our isolates S-CBP1 and S-CBP3, six were from Chesapeake Bay (Bench et al. 2007), two from coastal Mediterranean waters (V31 and V141 viral fractions), one from Red Sea (BAC9D04) (Zeidner et al. 2005), and one from the Northern Gulf of Maine [GOS007, from The Sorcerer II Global Ocean Sampling project (Rusch et al. 2007)]. It has been speculated that the Red Sea clone (BAC9D04) contains the *psbA* gene originated from a *Synechococcus* podovirus because the same cloned fragment also contains cyanopodovirus like genes (Zeidner et al. 2005; Bench et al. 2007). Our results are consistent with this hypothesis.

The GC content of *psbA* genes in different lineages of cyanobacteria and their phages could vary significantly (Zeidner et al. 2005; Sullivan et al. 2006). Percentage GC content at the third codon position varied from 36.6% (*Prochlorococcus* myovirus P-SSM10) to 86.2% (*Synechococcus* CB0208) among the sequences

analyzed in this study (Fig. 4-8). Each major phylogenetic group appeared to have distinguishable percent GC content at the third codon. The *Synechococcus* group had the highest % GC; *Prochlorococcus* and their phage group had the lowest; and *Synechococcus* phage groups appeared to have intermediate % GC content (Fig. 4-8). Compared with the variable triplet sequences residing within the lately defined D1 protein motif $^{R}/_{K}$ETTXXXS$^{Q}/_{H}$ (Sharon et al. 2007), the *Synechococcus* podovirus and environmental viral *psbA* sequences contain the most diverse "viral-like" signatures assigned by Sharon et al. (2007) (Fig. 4-8). Among the eight "viral-like" triplets identified in this cluster, four of them (EDI, ETV, EDM and EDV, shown with asterisk in Fig. 4-8) appeared to be unique to virus. Despite that not all the *Synechococcus* podoviruses contain the *psbA* gene, our analysis suggests that diverse *Synechococcus* podovirus-like *psbA* gene sequences are present in various marine environments. Further effort is needed to isolate and characterize more *psbA*-possessing cyanopodoviruses in order to understand the distribution and evolution of this *psbA* group.

It is noteworthy that *Synechococcus* podovirus *psbA* sequences did not cluster with *Prochlorococcus* podovirus *psbA* sequences. Given the origin of viral *psbA* gene, a host-range-limited gene transfer mechanism (Sullivan et al. 2006) could primarily explain such observation. The *psbA* gene could not be amplified from three estuarine *Synechococcus* siphoviruses, and this is consistent with the absence of *psbA* gene in the two *Prochlorococcus* siphoviruses (Sullivan et al. 2006). Combining our results with previously published data, the presence of *psbA* gene in *Cyanomyoviridae* (71 out of 112 isolates, 63%) and *Cyanopodoviridae* (7 out of 11 isolates, 64%) but

the absence in *Cyanosiphoviridae* (0 out of 5) indicates that the distribution of the *psbA* gene varied among different cyanophage families. Further studies are needed to understand why cyanosiphoviruses do not carry the *psbA* gene. It was speculated that the virus-encoded *psbA* gene might preserve or enhance the host photosynthesis during viral infection and therefore benefit the phage production (Lindell et al. 2005; Sullivan et al. 2006). However, no significant increase in viral burst size (as an indicator for viral production) was seen among the four cyanopodoviruses we isolated regardless whether they contain the *psbA* gene or not. The burst size of two *psbA*-processing cyanophages (S-CBP1 and S-CBP3) is similar to that of non *psbA*-containing cyanophages (S-CBP2 and P60). To truly understand whether possessing a *psbA* gene will enhance the viral production, a gene knock-out experiment is necessary.

The *psbD* gene could not been detected in all the seven estuarine synechophages. This is consistent with the observation that the *psbD* gene is commonly absent among the cyanophages with narrow host range (Sullivan et al. 2006). Cyanomyoviruses tend to have a broad host range, while some phages are even able to infect both *Synechococcus* and *Prochlorococcus* (Sullivan et al. 2003). About 50% of *psbA*-containing cyanomyoviruses also harbor the *psbD* gene (Sullivan et al. 2006), but none of the cyanopodoviruses and cyanosiphoviruses tested so far contains the *psbD* gene. It appears that the presence of the *psbD* gene depends on the types of cyanophage.

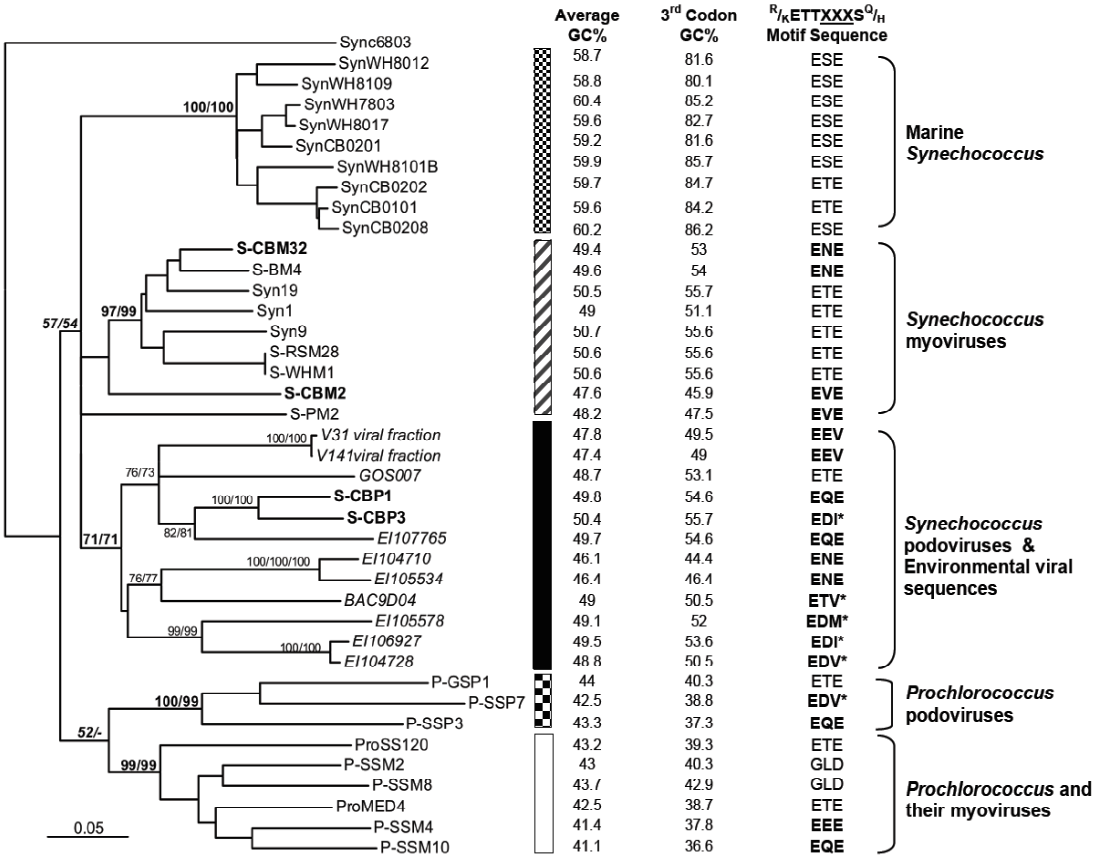


Fig. 4-8. Phylogenetic relationship, % GC contents and the triplet sequences from D1 protein motifs of partial *psbA* genes from marine picocyanobacteria and their cyanophages. The NJ tree was constructed based on 588-nt DNA sequences. Bold-types represent estuarine *Synechococcus* phages including S-CBM2 (DQ206827), S-CBM32 (DQ206829, host strain WH7805), S-CBP1 (DQ206826) and S-CBP3 (DQ206828). The partial *psbA* gene sequences from four estuarine *Synechococcus* strains CB0101 (DQ212909), CB0201 (DQ212912), CB0202 (DQ212910) and CB0208 (DQ212911) were also included. Additional sequences reported by Millard et al. (2004), Lindell et al. (2004), Zeidner et al.(2005), Sullivan et al. (2006) and Bench et al. (2007) were included here. Environmental clone GOS007, was from The Sorcerer II Global Ocean Sampling project (Rusch et al. 2007), sample ID: JCVI-READ-1092216070160 found in CAMERA (Seshadri et al. 2007). Bootstrap values (>50%) obtained from MCL and LogDet analyses are shown in the order of MCL/LogDet for major branches for clarity. The % GC content of all sites (in average) and the 3rd codon were listed in pair for all sequences. The triplet sequences of D1 protein motifs for these sequences were listed to the right. Bold-types are those defined as "viral-like" sequences while sequences with asterisk label are those found only for viruses, details can be found in Sharon et al. (2007).

The T7-like DNA polymerase (*pol*) gene was detected by PCR in all the three podoviruses, but not in the siphoviruses or myovirus. The *pol* phylogeny showed that the three estuarine *Synechococcus* podoviruses (S-CBP1, S-CBP2 and S-CBP3) cluster with oceanic *Synechococcus* podoviruses (P60, Syn5 and S-CBP42) and *Prochlorococcus* podovirus (P-SSP7) (Fig. 4-9). These podoviruses all contain a conserved family A DNA polymerase. Overall, the *pol* gene sequences among these podoviruses are 41-68% identical to each other, but only 29-31% identical to T7. The *pol* gene sequences from marine cyanopodoviruses formed a well supported monophyletic cluster (Marine Picocyanobacteria Podovirus, MPP subgroup, Fig. 4-9), distinguishable from the T7 subgroup (Scholl and Merril 2005; Molineux 2006), SP6 subgroup (Scholl et al. 2004), phi-KMV subgroup (Ceyssens et al. 2006). The T7, SP6 and phi-KMV subgroups belong to the T7 supergroup (Hardies et al. 2003). In addition, the MPP subgroup was distantly related to a marine roseobacter podovirus SIO-1 and a freshwater cyanopodovirus Pf-WMP4 infecting a freshwater filamentous cyanobacterium *Phormidium foveolarum* (Liu et al. 2007) (Fig. 4-9). The conservation of the DNA polymerase gene among marine cyanopodoviruses, which infect a broad range of hosts including MC-A, MC-B *Synechococcus* and *Prochlorococcus*, suggests the DNA replication modules are important to this group of marine cyanophages (podoviruses) and they could arise from a common ancestor. The virus-encoded DNA replication system (i.e. *pol* and DNA primase-helicase) allows DNA viruses to replicate their genomes quickly and faithfully, therefore, it is essential for many lytic podoviruses (Chen and Lu 2002). DNA polymerase is present

in all the currently sequenced T7-like podovirus genomes (22 viruses, as of July 2007 in GenBank) as well as other groups of podovirus (N4-like and Phi29-like).

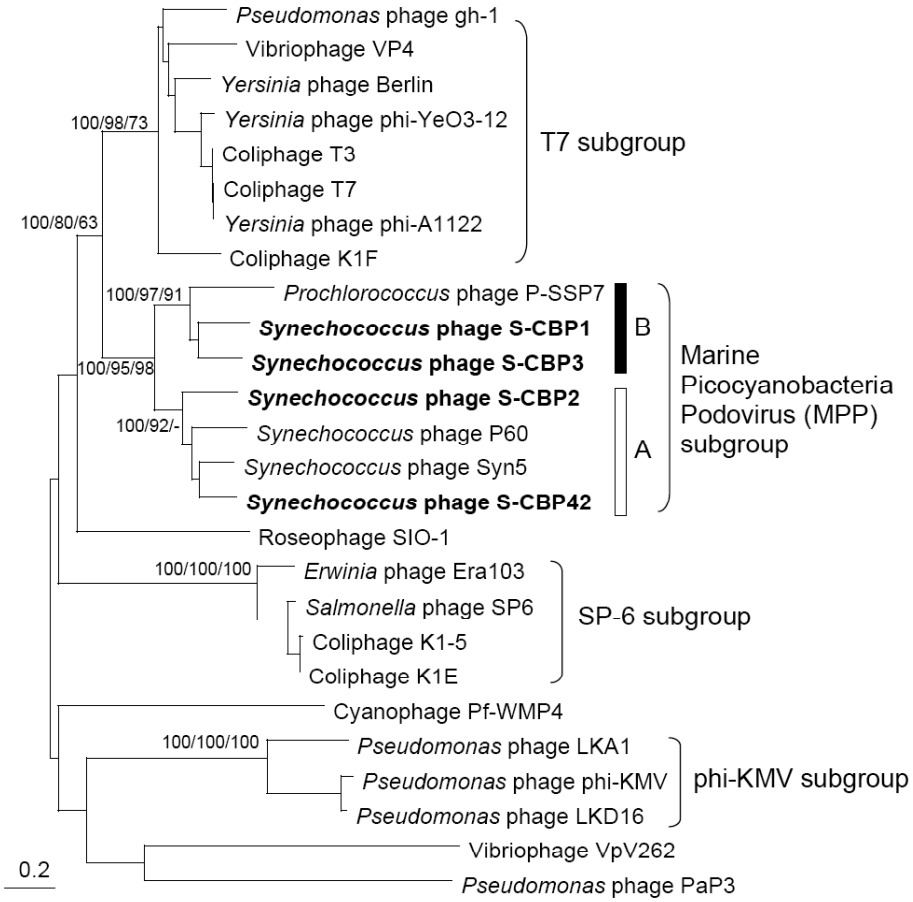


Fig. 4-9. Phylogenetic tree of *pol* gene amino acid sequences from 26 podoviruses within the T7 supergroup. The NJ tree was obtained based on aligned 210 amino acid residue sequences. Cyanopodovirus S-CBP1 (EF535232), S-CBP2 (DQ206830), S-CBP3 (EF535233) and S-CBP42 (EF535234, host strain WH7803) are shown in boldface. Other sequences obtained from GenBank are: cyanophage P60 (AF338467), P-SSP7 (AY939843), Syn5 (YP001285436), Pf-WMP4 (NC008367); Roseophage SIO-1 (AF189021), Coliphage T3 (AJ318471), T7 (V01146), K1-5 (YP654112), K1E (YP424986), K1F (YP338108); *Yersinia* phage phi-YeO3-12 (AJ251805), Berlin (YP919001), phi-A1122 (AY247822); Vibriophage VP4 (YP249581), VpV262 (AY095314); *Pseudomonas* phage gh-1 (AF493143), LKA1 (CAK24997), LKD16 (CAK25954), phi-KMV (NP877458), PaP3 (AY078382); *Salmonella* phage SP6 (NP853574) and *Erwinia* phage Era103 (EF160123). Bootstrap values (>50%) for neighbor-joining, maximum parsimony and maximum likelihood analyses are shown on the main branches in the order of NJ/MP/ML.

The MPP subgroup consists of at least two sister clusters A and B, which were well supported by the bootstrap values (Fig. 4-9). Cluster A includes three podoviruses infecting MC-A *Synechococcus* spp. and one podovirus (S-CBP2) infecting MC-B *Synechococcus* CB0208, while cluster B consists of two MC-B *Synechococcus* podoviruses (S-CBP1 and S-CBP3) and a *Prochlorococcus* podovirus (P-SSP7). Several interesting features within the MPP subgroup were noticed: 1) the *Synechococcus* hosts of podoviruses in the MPP-A cluster all contain phycoerythrin (PE), while the *Synechococcus* hosts for the MPP-B cluster lack PE; 2) the three cyanopodoviruses in MPP-B cluster contain the *psbA* gene, but none in the MPP-A cluster does; 3) members of the MPP-B cluster have a similar latent time of ca. 8 h, significantly longer than cyanophage P60, a member of MPP-A podovirus. However, whether there is a correlation between the length of the latent period and the presence of "photo genes" in the cyanopodoviruses still remains to be investigated.

The viral capsid assembly gene (g20) was also detected in many estuarine *Synechococcus* myoviruses including S-CBM2. Their g20 gene sequences were clustered with cyanomyoviruses isolated from oceanic *Synechococcus* and *Prochlorococcus*, forming a well-supported lineage apart from those environmental sequences as well as other bacterial myoviruses (Fig. 4-10). The g20 gene has been widely used to investigate the phylogenetic diversity of cyanomyoviruses in various aquatic environments (Wilson et al. 1999; Zhong et al. 2002; Dorigo et al. 2004; Wang and Chen 2004; Short and Suttle 2005; Sandaa and Larsen 2006; Wilhelm et al. 2006). However, a great deal of g20 sequences recovered from natural environments was not closely related to known cyanomyoviruses infecting oceanic *Synechococcus*

(Zhong et al. 2002; Short and Suttle 2005; Wilhelm et al. 2006). Our analyses showed that the g20 gene sequences from the cyanomyoviruses that infect estuarine, coastal, oceanic *Synechococcus*, as well as *Prochlorococcus* still cluster together (Fig. 4-10), suggesting that the majority of unmatched environmental g20 sequences are likely not from cyanomyoviruses. To explore the genetic diversity of cyanomyoviruses in the aquatic environment, the current PCR primer sets for the g20 gene need to be modified by incorporating the newly available g20 gene sequences.

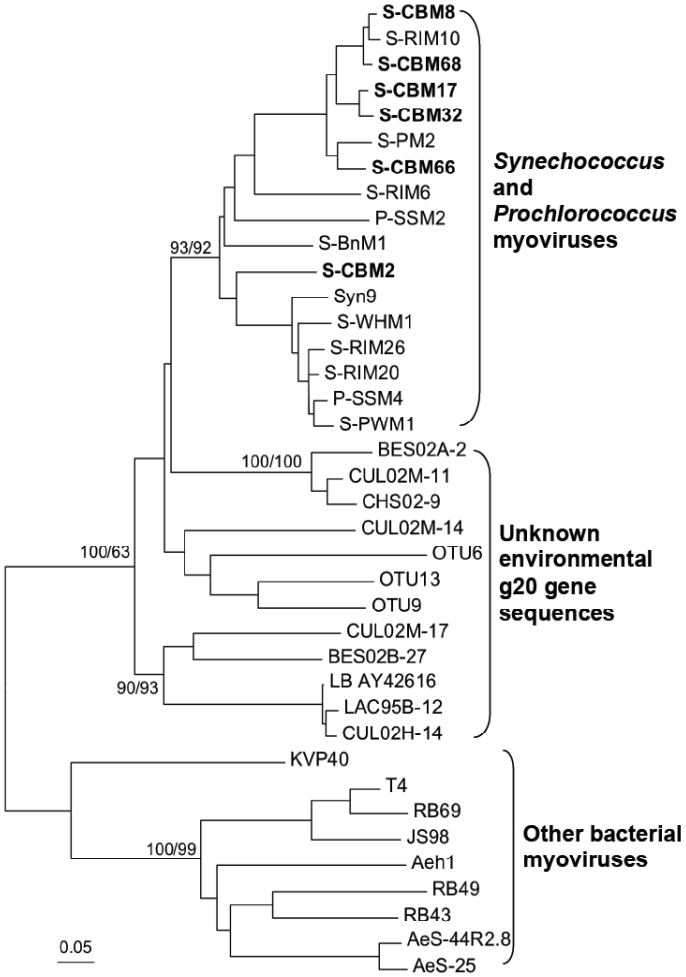


Fig. 4-10. Neighbor-joining tree of g20 gene based on 164 amino acid sequences of 26 myovirus isolates and 12 environmental clones from various studies. Bold-types represent *Synechococcus* myovirus isolated from Chesapeake Bay, including S-CBM2 (DQ212908), S-CBM8 (DQ212903), S-CBM17 (DQ212905), S-CBM32

(DQ212907), S-CBM66 (DQ212906) and S-CBM68 (DQ212904). Among them, S-CBM32 was isolated from oceanic *Synechococcus* WH7805, while all the rest phages were isolated using estuarine *Synechococcus* CB0101. Sequences of cyanomyovirus isolates S-PM2, S-BnM1, S-WHM1, and S-PWM1 were obtained from Zhong et al. (2002), S-RIM6, S-RIM10, S-RIM20 and S-RIM26 were from Marston and Sallee (2003). Myoviruses infecting other group of bacteria are: Vibriophage KVP40 (AB020525), coliphages T4 (AF158101), JS98 (AY746497), RB43 (AY343333), RB49 (NC005066), RB69 (AY303349), Aeromonas phages Aeh1 (AY266303), AeS-25 (DQ529280), AeS-44R2.8 (AY375531). The environmental clone sequences of OTU 6, 9 and 13 were from our lab (Wang and Chen, 2004); LB AY426168 was from Dorigo et al. (2004); BES02A-2, BES02A-27, CHS02-9, LAC95B-12, CUL02M-11, CUL02H-14 and CUL02M-17 were from Short and Suttle (2005). Bootstrap values (>50%) for Neighbor-joining and Maximum likelihood analyses are shown on the main branches in the order of NJ/ML.

Conclusions

Our work represents the first study on cyanophages infecting an autochthonous group of estuarine *Synechococcus* (MC-B cluster). Upon extensive viral screening efforts, four out of 11 estuarine *Synechococcus* strains tested were susceptible to phage infection. Six out of seven phages obtained from these hosts were highly host specific. We first reported the presence of the photosystem II core gene (*psbA*) in two *Synechococcus* podoviruses. The *psbA* genes of *Synechococcus* podoviruses are distinct from those of marine *Synechococcus* and their myoviruses. We also demonstrated that the DNA polymerase gene (*pol*) is conserved among the cyanopodoviruses infecting marine *Synechococcus* and *Prochlorococcus*. The *pol* gene can be an ideal gene marker to investigate the genetic diversity of cyanopodoviruses in nature. The g20 gene is conserved among the cyanomyoviruses that infect distantly related members of marine picocyanobacteria. Compared to cyanomyoviruses and cyanopodoviruses, much less is known about

cyanosiphoviruses. Perhaps genome sequencing of a few cyanosiphoviruses could provide new insight to the lysogenic potential of *Synechococcus*. In short, estuarine *Synechococcus* and their phages appear to develop more stringent host-phage relationship compared to their oceanic counterparts.

Chapter 5: Genetic diversity and population dynamics of cyanophage communities in the Chesapeake Bay

Abstract

In order to understand the genetic diversity and population dynamics of cyanophages in the estuarine water, the viral capsid assembly gene (g20) gene was used as a gene marker to monitor the genetic variations of natural cyanomyovirus communities in the Chesapeake Bay. Unique and diverse g20 sequences were found in the Chesapeake Bay. Only one of 15 g20 genotypes was closely related to the known cyanomyovirus isolates. Most of the g20 genotypes in the Chesapeake Bay were not related to the g20 clonal sequences recovered from the open ocean water. Terminal-Restriction Fragment Length Polymorphism (T-RFLP) based on the g20 gene was developed to investigate temporal and spatial distribution of cyanomyovirus communities in the Bay. The T-RFLP profiles of the g20 gene demonstrated that the cyanomyovirus population structures in the Bay were more dynamic seasonally rather than spatially. The seasonal variation of the cyanophage community appeared to correspond to the change in host cell density, which in turn was mainly affected by the water temperature. This study presented the first effort to monitor both cyanophage titer and genetic diversity over time and space. Our study suggested that cyanophages could play a significant role on regulating the *Synechococcus* biomass and population structure in the Chesapeake Bay.

Introduction

Discovery of highly abundant viruses (i.e. 10^7 viral particles ml^{-1}) in marine environments re-initiated the investigation into ecological roles of marine viruses (Bergh et al. 1989; Proctor and Fuhrman 1990; Fuhrman 1999). Viruses are now known to be ubiquitous biological components that could regulate element cycling in the microbial loop, altering the nutrient cycling and energy flow (Suttle et al. 1990; Fuhrman and Suttle 1993; Fuhrman et al. 1993; Thingstad et al. 1993; Bratbak et al. 1994; Wilhelm and Suttle 1999). Viruses are also thought to mediate gene transfer among microorganisms in natural aquatic environments, and shape the genetic diversity of the microbial community, by means of virus-mediated genetic exchange such as transduction, transformation and conversion between lytic and lysogenic cycles (Fuhrman 1999; Paul 1999; Wommack and Colwell 2000). Unicellular cyanobacteria of the genus *Synechococcus* are among the most abundant and ubiquitous members of the picoplankton in the open ocean, and it has been estimated that they are responsible for substantial portion of the primary production in the open sea (Waterbury et al. 1986; Li 1994). Cyanophages were found to be abundant (i.e. 10^3-10^6 ml^{-1}) and a significant factor in determining the dynamics of *Synechococcus* populations (Suttle and Chan 1993; Waterbury and Valois 1993; Wilson et al. 1993; Suttle and Chan 1994; Lu et al. 2001). Cyanophages that infect *Synechococcus* spp. sometimes can reach concentrations in excess of 10^6 ml^{-1} in marine environments (Suttle 2000), (Suttle and Chan 1994).

Cyanophage isolates that infect marine *Synechococcus* spp. are diverse in terms of morphology. All known cyanophages belong to three phage families, *Myoviridae*,

Podoviridae and *Siphoviridae* (Safferman et al. 1983; Ackermann and DuBow 1987b; Martin and Benson 1988). Molecular characterization of natural cyanophage assemblages has so far revealed much greater diversity than ever expected. Cyanophage genotypes revealed by restriction fragment length polymorphism (RFLP) exhibited even greater complexity, when compared with their morphotypes (Wilson et al. 1993; Lu et al. 2001). Cyanophages of *Myoviridae* (cyanomyoviruses) are commonly found among cyanophage isolates (Suttle and Chan 1993; Waterbury and Valois 1993; Lu et al. 2001). Characterization of a conserved viral capsid assembly protein gene (g20) in three cyanomyoviruses allowed the design of PCR primers specific to cyanomyoviruses and therefore greatly facilitated the investigation of the genetic diversity of natural cyanophage assemblages (Fuller et al. 1998). Recently, denaturing gradient gel electrophoresis (DGGE) analysis based on 165-bp DNA fragments of g20 gene amplified by cyanomyovirus specific PCR primers (CPS4 and CPS5) has been used to examine the population structure of cyanophages along a south-north transect in the Atlantic Ocean. High genetic diversity of cyanophage was found through the depth profile and significant changes in the population structure were observed from surface to depth, while the maximum diversity was always correlated to the maximum *Synechococcus* abundance (Wilson et al. 1999; Wilson et al. 2000). More recently, cyanomyovirus-specific primers (CPS1 and CPS8) have been successfully used to amplify ca. 592 bp-fragments of g20 gene from many cyanomyovirus isolates and natural virus communities (Zhong et al. 2002). Phylogenetic analysis of 114 g20 gene sequences recovered from both coastal and oceanic water samples also revealed strikingly high genetic diversity. The g20

sequence diversity varied from coastal to oceanic waters and from surface water to the deep chlorophyll maximum depth (Zhong et al. 2002).

Genetic diversity of natural marine virus communities appears to be more complex than we expected. In order to understand better co-variation and co-evolution between marine phages and their host bacteria, it is necessary to study the genetic variation among marine viruses. Microbial diversity studies based on cloning and sequencing techniques are expensive and time consuming, and not suitable when a large amount of environmental samples are involved. Virioplankton communities in Chesapeake Bay revealed by the pulse-field gel electrophoresis (PFGE) demonstrated that the annual variation in viral community structure was correlated with time, geographical location and extent of water column stratification (Wommack et al. 1999a; Wommack et al. 1999b). A technique like PFGE is very useful for large-scale ecological studies when numerous environmental samples are involved. However, the PFGE technique is more suitable for studying the whole viral community than for a specific group of phages like cyanophage.

Currently, two rapid genetic fingerprinting techniques, denaturing gradient gel electrophoresis (DGGE) (Muyzer et al. 1993; Muyzer 1999) and terminal-restriction fragment length polymorphism (T-RFLP) (Liu et al. 1997b; Marsh 1999) are commonly used to investigate complex microbial communities. T-RFLP method takes advantage of the high resolution and high throughput of automated sequencing technologies to detect the terminal restriction fragments (T-RFs) after restriction digestion. This method was first used to identify complex bacterial communities in 1994 (Avaniss-Aghajani et al. 1994), and since then this rapid genetic fingerprinting

technique has been widely used to investigate complex community structure of bacterial, archaeal and eukaryal communities in various environments (Muyzer 1999). Compared with the PCR-DGGE fingerprinting technique, the T-RFLP method can provide simple and rapid T-RF data by using standard sequencing equipment. The T-RF data can be easily analyzed with a variety of statistical techniques. The comparison between observed T-RFs with simulated T-RFs from clone library or sequence database allows the identification of specific genotypes. The use of DGGE however, allows the subsequent sequencing of specific genotypes present in the gel, which is not possible for T-RFLP technique. T-RFLP has been demonstrated to be an automated and sensitive tool for characterization of complex microbial community structure and dynamics (Liu et al. 1997b; Marsh 1999; Muyzer 1999; Osborn et al. 2000; Kitts 2001).

In this study, we first examine the genetic diversity of cyanomyovirus in the Chesapeake Bay based on the g20 gene RFLP patterns and their sequences. Secondly, the T-RFLP method based on the g20 gene was developed to investigate the spatial and temporal variations of the natural cyanomyovirus population in the Chesapeake Bay.

Materials and Methods

Location and sampling

For spatial analysis, the Chesapeake Bay water samples were collected using Niskin bottles on board the R/V *Cape Henlopen* on November 1-3, 2000. Water samples from Pier 5 in the Inner Harbor and from three stations (Station 908, 818 and 707) in the Chesapeake Bay were collected (Fig. 5-1 and Table 5-1). For temporal

analysis, water samples were collected monthly at Pier 5 in the Inner Harbor from March 2001 through May 2002. For direct counts of viruses and picocyanobacteria, 10 ml of surface water sample from Pier 5 was collected and fixed with glutaradehyde (final concentration of 2.5%) and stored at 4°C in the dark until use.

Table 5-1. Stations in the Chesapeake Bay where viral concentrates were collected.

Location	Date	Latitude	Longitude	Depth (m)	Salinity (‰)	Temp (°C)
Inner Harbor	1-Nov-2000	39˚17' N	76˚36' W	Surface	20.00	18.5
CB908	3-Nov-2000	39˚08' N	76˚20' W	3	18.15	16.6
CB818	3-Nov-2000	38˚18' N	76˚17' W	3	19.45	16.0
CB707	2-Nov-2000	37˚07' N	76˚07' W	3	29.03	15.4

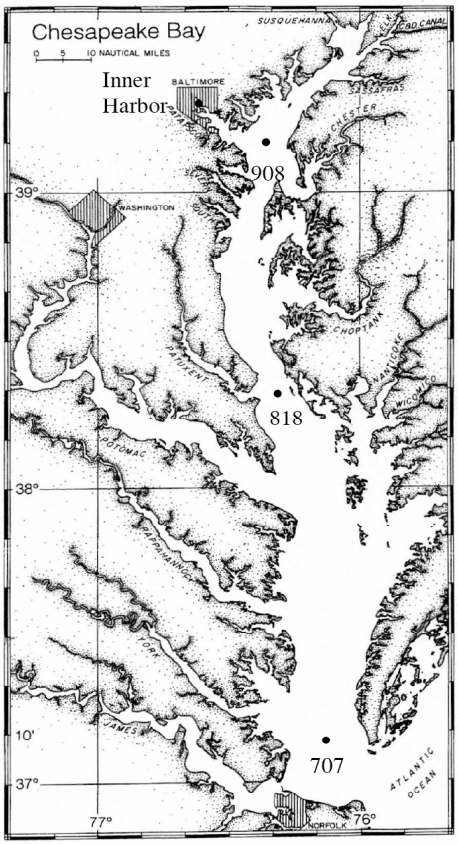


Fig. 5-1. Sampling sites in the Chesapeake Bay

Ultrafiltration

Viral communities from the Chesapeake Bay were concentrated during the cruise, while viral concentrates (VCs) from Inner Harbor were prepared in the laboratory. Viral communities were concentrated using the ultrafiltration protocols described by Chen et al. (1996). In brief, 40 to 80 liters of water was filtrated through A/E glass fiber filters (Gelman Sciences; nominal pore size, 1.2 μm) and 0.45 μm pore-size low-protein-binding Durapore membranes (Millipore). The filtrate was then concentrated by ultrafiltration through 30,000 MW cutoff Amicon S10Y30 spiral cartridge (Millipore) in the ProFlux M-12 system (Millipore), 30% of pump speed and 16-18 kPa of backpressure. The final water sample concentrates contained particulates range in size between 20 nm (approximately 30 kDa) to 450 nm. Most viruses range in size from 20 to 400 nm with molecular weight greater than 30 kDa. The final volume of each viral concentrate ranged from 400 to 800 ml.

Direct counts of viruses and cyanobacteria

Virus-like particles (VLP) were enumerated following the protocol described by Chen et al. (2001). Briefly, 100-300 μl of the water samples were brought up to 2 ml in final volume with Tris EDTA (10 mM / 1 mM, pH 7.5) and filtered onto a 0.02 μm pore size Al_2O_3 Anodisc 25 mm membrane filter (Whatman) under approximately 150-200 mm Hg vacuum. The sample was stained with 2 × SYBR Gold solution (final concentration) for 15 min in the dark. The viral-like particles were counted under blue excitation (excitation BP 485, emission LP 520) with a Zeiss Axioplan epifluorescence microscope at ×1000 magnification. At least 200 viral particles from 10 random fields were counted per sample.

The *Synechococcus* cell abundance in a water sample was counted following the method described in Chapter 2.

Plaque Assay

To enumerate cyanophage that infect indigenous *Synechococcus* in the Chesapeake Bay, a plaque assay was developed using the *Synechococcus* sp. CB0101 isolated from the Bay. Among many *Synechococcus* strains isolated from the Bay, CB0101 was most sensitive to viral infection (Chapter 4). The strain CB0101 also represented a common genotype in the *Synechococcus* community revealed by the RuBisCO gene sequences (Chapter 2). The strain CB0101 was grown in SN medium (Waterbury et al. 1986) with 12‰ salinity and 900 μM of $NaNO_3$ as a nitrogen source at 22 to 24°C. Strain CB0101 was incubated with light intensity of 20 to 30 microeinsteins m^{-2} s^{-1} with a cycle of light: dark =16 h: 8 h. Pour plating procedure for plaque assay followed the protocol described by Brahamsha (Brahamsha 1996b). For each plate, 100 μl of viral concentrate was screened against CB0101, and triplicate plates were analyzed for each sample.

PCR amplification

The oligonucleotide primers CPS1 (5'-GTAGWATTTTCTACATTGAY GTTGG-3') and CPS8 (5'-AAATAYTTDCCAACAWATGGA-3') (Zhong et al. 2002) were used to amplify ca. 592 bp fragments from cyanomyoviruses. PCR amplification followed the protocol described by Zhong et al. (2002), except that the annealing temperature was increased to 46°C to increase the PCR specificity. We tested the annealing temperature ranges from 36 to 52°C, and found that 46°C was the optimal annealing temperature in order to avoid unspecific amplification. The PCR

product was examined by electrophoresis in 2% agarose gels stained with ethidium bromide.

Clone library construction

PCR amplicons from each VC were purified by using Wizard PCR Prep DNA Purification System (Promega). The purified DNA fragments were cloned into the pGEM-T Easy cloning vector (Promega) and then transformed into JM109 competent cells (Promega) according to the manufacturer's instructions. Positive clones (white-colonies) were picked randomly and transferred onto a new agar plate for further use.

Restriction Fragment length polymorphism (RFLP) analysis

To avoid redundant sequencing, g20 clones amplified from four Chesapeake Bay virus communities were pre-screened with RFLP. About 60 positive clones from each clone library were randomly picked and the plasmid inserts were PCR-amplified with vector-specific primers T7 (5'-TAATACGACTCACTATAGGGCGA-3') and SP6 (5'-ATTTAGGTGACACTATAG-3') along with Taq DNA polymerase (Promega). PCR amplification cycles involved a 3-min initial denaturation at 94°C and followed by 35 cycles of 94°C for 1 min, 56°C for 1 min, 72°C for 2 min, and a 5-min final extension at 72°C. Once a clone was confirmed to contain an insert of the appropriate size (ca. 772 bp), the insert was subjected to RFLP analysis. Forty-four clones from Inner Harbor, 46 clones from station 908, 42 clones from station 818 and 48 clones from station 707 were confirmed to contain the correct inserts. The commonly used restriction enzymes *Hae* III (GG'CC) and *Rsa* I (GT'AC) (Promega) for T-RFLP analysis were used to digest the PCR products and their resolutions were

compared in the preliminary trial. It appeared that *Rsa* I could yield more RFLP patterns among tested clones than *Hae* III did (data not shown) and therefore, *RsaI* was chosen to digest PCR products in the subsequent RFLP and T-RFLP analysis. A subsample (10 μl) from each PCR mixture was digested with 5 U *Rsa*I in 1 × Buffer C (Promega) at 37°C overnight in a final reaction volume of 25 μl. Digested DNA was separated by 2% agarose gel electrophoresis as described above. The resulting RFLP patterns were examined and compared by visual check.

Sequencing and phylogenetic analysis

Representative clones (eight from Inner Harbor, three from station 908, three from station 707, one from station 818) that could be distinguished by RFLP analysis were sequenced. The plasmid inserts from selected clones were PCR-amplified with vector-specific primers T7 and SP6 using Expand^{TM} High Fidelity DNA polymerase (Roche) with conditions described above. The purified DNA from each representative genotype was sequenced bi-directionally with primers T7 and SP6 using the ABI model 310 automated DNA Sequencer (Applied Biosystems). Sequence alignment and phylogenetic analysis were performed using the Mac Vector 7.1 program (GCG, Madison, WI). Pairwise alignment was calculated by using Blosum 30 as matrix with open gap penalty of 10.0 and extend gap penalty of 0.1. The phylogenetic tree was constructed by using neighbor-joining method based on ca. 197 amino acid residues inferred from their nucleotide sequence alignment with T4 as the outgroup. The protein distances were determined by using Poisson-correction method.

T-RFLP analysis

To obtain rapid fingerprints of cyanomyovirus communities, CPS1 and 5' Hex-labeled CPS8 primers were used for T-RFLP analysis, following the same PCR conditions described above. Purified PCR products were digested with *Rsa*I overnight at 37°C. Each 20-μl-digestion mixture contained ca. 300 ng PCR products, 5 U *Rsa*I enzyme and acetylated BSA (final concentration 0.1 mg ml^{-1}) as recommend by the manufacturer (Promega). The digested DNA was precipitated with a 0.1 volume of 3 M sodium acetate and 2.0 volumes of 95% ethanol, followed by centrifugation at 16,000 × *g* for 20 min. The DNA pellet was washed with 75% ethanol, dried, and resuspended in a mixture of 12 μl deionized formamide and 0.5 μl Internal Lane Standard 600 (asymmetrically labeled with carboxy-x-rhodamine, Promega). Fluorescently labeled terminal restriction fragments (T-RFs) were size separated on an ABI 310 automated sequencer (Applied Biosystems). T-RFLP profiles were generated and analyzed using GENESCAN 2.1 software (Applied Biosystems). The size, in basepair, of T-RFs was analyzed by comparison with the internal standard using the Local Southern Method, GeneScan 2.1 software (Applied Biosystems). For each viral concentrate, at least duplicate samples were analyzed. To avoid detection of primers and other uncertainties, T-RFs smaller than 50 bp and larger than 600 bp were excluded from the analysis and only peaks over a threshold of 50 units above background fluorescence were analyzed.

Computer-based predictions of the expected T-RFs from g20 gene sequences were performed using the Mac Vector 7.0 program (GCG, Madison, WI). To calibrate bias between the predicted size and observed size, the T-RFs of g20 amplicons from

representative clones were examined individually. The model community constructed with these amplicons was used as a reference for the subsequent T-RFLP profile analysis.

Nucleotide sequence accession numbers

The nucleotide sequences obtained in this study have been deposited in GenBank and assigned the accession numbers from AY152732 to AY152746.

Results

Seasonal variation of *Synechococcus* and their phages

A strong seasonal variation in *Synechococcus* spp. and their phage abundance was observed in the Inner Harbor (Fig. 5-2). The *Synechococcus* density ranged from 2.6 ± 0.5 (standard deviation, the same as follows) × 10^2 cells ml^{-1} in February to 8.1 ± 1.0 × 10^4 cells ml^{-1} in July. Meanwhile, the *Synechococcus* phage titer as determined by using the indigenous strain CB0101 increased from the lowest value of 2 PFU ml^{-1} in April to the highest value of 560 ± 172 PFU ml^{-1} in September. Phage titers in the summer water samples were about 260 fold higher than those in the winter water samples (Fig. 5-2 B). The diverse plaque morphologies were observed by using *Synechococcus* CB0101 as host. The diameters of the plaques ranged from ca. 1 mm to larger than 1 cm. Both clear and turbid plaques with roughly circular or irregular shapes were shown on the host bacterial lawn. More diverse plaque morphotypes and wider ranges of plaque sizes were seen in the summer than in the winter. In general, large and clear plaques (greater than 5 mm in diameter) were dominant in warm seasons while only a few clear plaques (ca. 2 mm in diameter) were seen in cold seasons. Among 56 cyanophages recently isolated from the Bay, 32

(approximately 60%) of them were tested positive with the g20 primers. The concentration of viral like particles during the course of this study was observed with a maximum abundance of $4.9 \pm 0.5 \times 10^7$ ml^{-1} in September and the least value of $9.3 \pm 1.9 \times 10^6$ ml^{-1} in April. In general, the seasonal variation of *Synechococcus*, cyanophage and total viral particles in the Chesapeake Bay appeared to follow a similar pattern. The *Synechococcus* cell density increased dramatically from June to July and remained high ($> 10^4$ cells ml^{-1}) until November. Correspondingly, the cyanophage titer and direct virus counts are low in winter and high in summer (Fig. 5-2 B). The temporal variation of *Synechococcus* and their phages appeared to be correlated with water temperature since salinity did not change much seasonally in the Inner Harbor (Fig. 5-2 A).

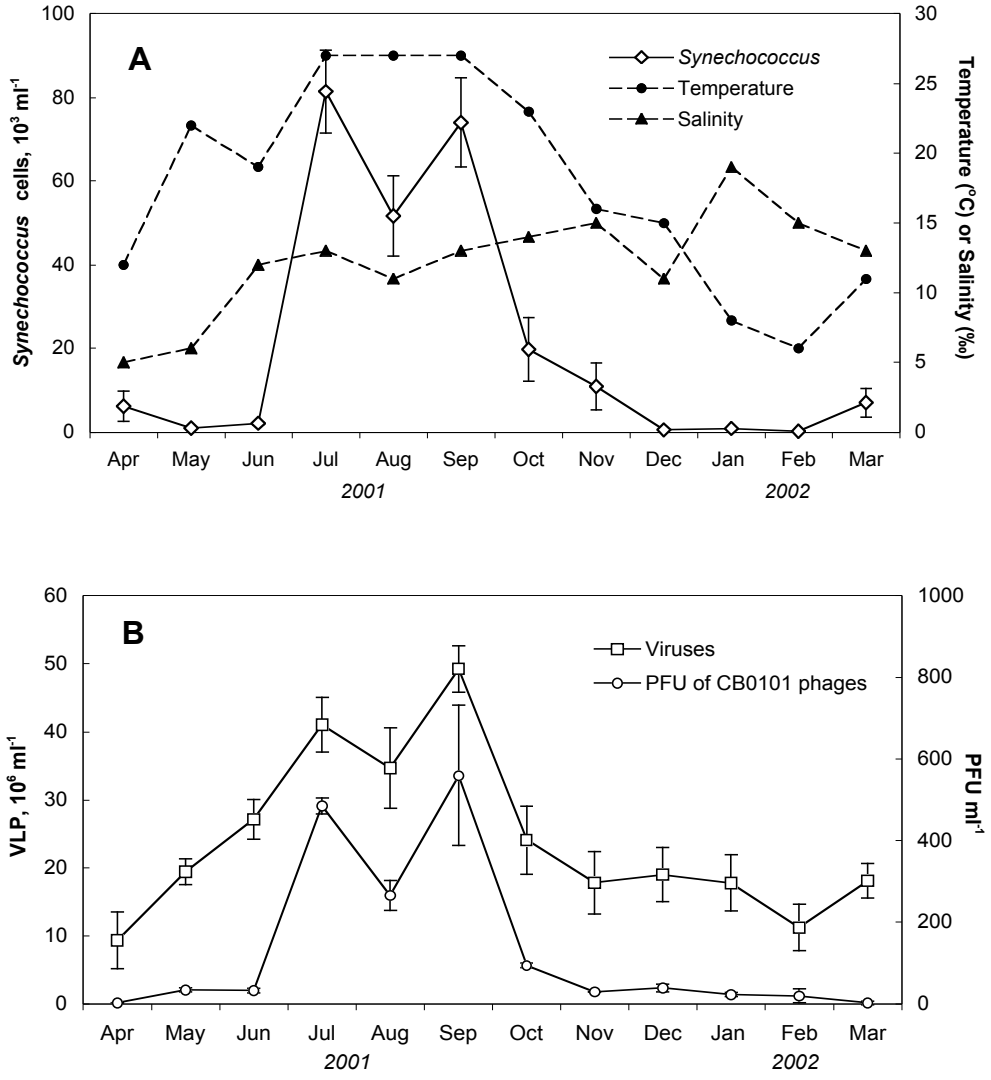


Fig. 5-2. A: Mean (±SD) monthly variation in *Synechococcus* cell density (cells ml^{-1}), water temperature and salinity and B: total VLP counts and *Synechococcus* CB0101 phage titer at Pier 5 of Baltimore Inner Harbor, from April 2001 to March 2002. PFU: plaque-forming unit. Error bars represent standard error.

g20 genotypes in Chesapeake Bay

The g20 gene fragments with ca. 592 bp (from 569 to 599 bp) were successfully amplified from Chesapeake Bay VCs. Among 180 randomly selected g20 clones, 15 distinguishable RFLP patterns were initially identified by visual check and assigned as operational taxonomic units (OTUs) from OTU1 through OTU15. The relative abundance of these OTUs found in the Chesapeake Bay VCs is summarized in Figure 5-3. OTU2 and OTU6 were present in all the four VC samples from the Bay, while OTU6 was the most abundant pattern (31%) observed. Eleven different RFLP patterns were observed in the Inner Harbor VCs, while eight, eight and seven RFLP types were found in the station 908, 818, 707 respectively (data not shown).

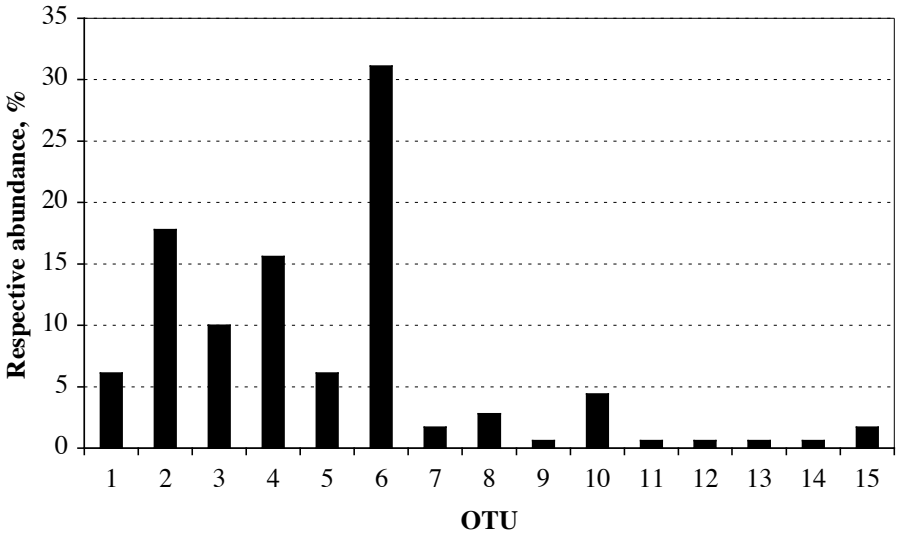


Fig. 5-3. Relative abundance (%) of g20 gene RFLP patterns digested by endonucleases *Rsa* I at four stations (combined data) in the Chesapeake Bay, showing 15 different RFLP patterns or OTUs identified among 180 g20 clones recovered from the Bay.

The representative clone sequences of these OTUs were different from each other. Notably, only one (OTU12) of 15 OTUs was closely related to known cyanomyovirus isolates. The g20 sequences recovered from Chesapeake Bay VCs appeared to be very diverse, and many of them formed their unique clusters by using the cut-off value of 0.1 substitution per site (Fig. 5-4). Cluster N1 to N4 contain only g20 clonal sequences recovered from the Bay. Cluster N1 (OTU6 and OTU15) and Cluster N4 (OTU1, 7, 10), which account for 45% of 180 clones, were found to be unique and significantly distant from all the currently known g20 sequences. Cluster N2 (including OTU2, 4, 5, 8, 13 and 14), which account for 42% of 180 clones could be related to the SE38 clone which was found in the Skidaway Estuary. The cluster N3, which consists of OTU3 and OTU9, was most closely related to a Sargasso Sea deep chlorophyll maximum clone (SS4705). OTU12 had greater than 97% amino acid sequence similarity to SE34 (Skidaway Estuary clone), and was only one that could be clustered with the known cyanomyovirus isolates. OTU11 was unique and therefore not grouped into any cluster of isolates. In general, the vast majority of g20 sequences recovered from Chesapeake Bay were unique and not closely related to currently known cyanomyovirus isolates.

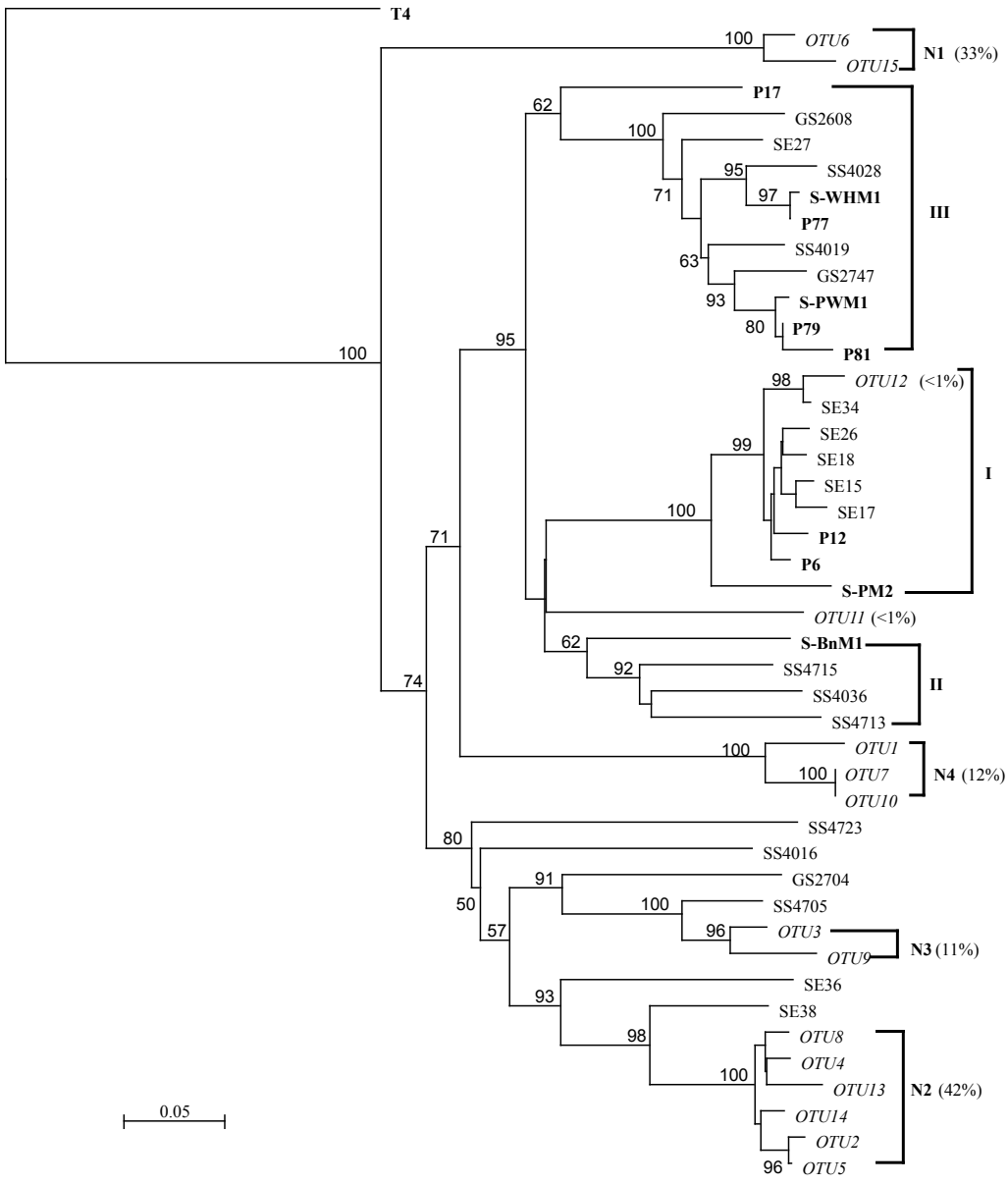


Fig. 5-4. Phylogenetic affiliation of Chesapeake Bay g20 genotypes (OTU1-OTU15) with cyanomyovirus isolates (bold face type) and representative g20 clones from various natural environments previously studied (Zhong et al. 2002). Each value in parentheses is the appearance frequency of each OTU of 180 clones screened. The tree was inferred by using the neighbor-joining method with Poisson correction based on ca. 197 amino acid residues sequence alignment with T4 as an outgroup. Bootstrap values were generated following 1000 replicates; values less than 50 are not shown. The scale bar indicates 0.05 substitutions per site. Clusters I, II, III was assigned based on the scheme of Zhong et al. (2002). Cluster N1 to N4 were assigned to include only the clone sequences recovered from the Chesapeake Bay (i.e. less than 0.1 substitutions per site).

Cyanomyovirus population dynamics

In order to identify the T-RFLP peaks using known g20 sequences, the predicted T-RFs of the representative g20 sequences were analyzed (Table 5-2). Forty-six representative sequences were chosen from more than 200 sequences of g20 gene in the GenBank, and analyzed for possible T-RFs patterns by computer analysis. Among the 46 g20 representatives which included 15 OTUs, 10 cyanomyovirus isolates and 21 environmental clones, 20 unique T-RFs were identified using computer simulation of restriction enzyme *Rsa*I. T-RFs of SS4019 and OTU10 contained 4 and 27 bp, respectively, and therefore were deleted from the analysis due to their small size (Table 5-2). The observed and predicted T-RFs of 15 OTUs were compared in Table 5-2. There were eight unique T-RFs generated from 15 representative clones of OTUs. Except for OTU9, where the observed T-RF was 8 bp shorter than the predicted size (144 bp vs. 152 bp), all other OTUs' observed T-RFs matched well with the predicted size (± 5 bp). The T-RFLP profile of this model community was therefore used as a reference for the subsequent analysis of spatial and temporal T-RFLP profiles of natural viral communities from the Chesapeake Bay.

Table 5-2. Description of T-RFs sizes of g20 gene fragments.

Sample	T-RFs Size Predicted	Observed	Host and/or Isolation Location	Source
Chesapeake Bay clones				
OTU1	140	140	Inner Harbor, Chesapeake Bay	This study
OTU2	485	481	Inner Harbor, Chesapeake Bay	This study
OTU3	221	216	Inner Harbor, Chesapeake Bay	This study
OTU4	485	480	Inner Harbor, Chesapeake Bay	This study
OTU5	485	480	Inner Harbor, Chesapeake Bay	This study
OTU6	592	592	Station 707, Chesapeake Bay	This study
OTU7	401	405	Station 908, Chesapeake Bay	This study
OTU8	346	342	Station 707, Chesapeake Bay	This study
OTU9	152	144	Station 908, Chesapeake Bay	This study
OTU10	27	ND	Inner Harbor, Chesapeake Bay	This study
OTU11	266	262	Inner Harbor, Chesapeake Bay	This study
OTU12	592	588	Inner Harbor, Chesapeake Bay	This study
OTU13	347	343	Station 707, Chesapeake Bay	This study
OTU14	346	342	Station 818, Chesapeake Bay	This study
OTU15	592	586-592	Station 908, Chesapeake Bay	This study
Cyanomyovirus isolates				
P6	592	ND	WH7805; Dauphin Island, Ala.	Lu et al. 2001
P12	592	ND	WH8101; Sayll Estuary, Ala	Lu et al. 2001
P17	457	ND	WH7803; Qingdao Coast, China	Lu et al. 2001
P77	150	ND	WH8007; Altamaha River Estuary, Ga	Lu et al. 2001
P79	404	ND	WH7805; Satilla River Estuary, Ga	Lu et al. 2001
P81	404	ND	WH7805; Altamaha River Estuary, Ga	Lu et al. 2001
S-PWM1	404	ND	WH7803; Gulf of Mexico	Suttle & Chan1993
S-PM2	592	ND	WH7803; Plymouth, UK	Wilson et al. 1993
S-WHM1	150	ND	WH7803; Woods Hole, Mass	Wilson et al. 1993
S-BnM1	115	ND	WH7803; Bergen, Norway	Wilson et al. 1993
Representative clones from various environments				
SE1	592	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE15	106	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE17	592	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE18	592	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE26	592	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE27	162	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE34	592	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE36	152	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
SE38	312	ND	Skidaway, Ga (31°59'N, 81°01'W)	Zhong et al. 2002
GS2608	162	ND	Gulf Stream (37°19'N, 71°37'W)	Zhong et al. 2002
GS2704	312	ND	Gulf Stream (36°24'N, 71°20'W)	Zhong et al. 2002
GS2747	312	ND	Gulf Stream (36°24'N, 71°20'W)	Zhong et al. 2002
SS4016	592	ND	Sargasso Sea (28°53'N, 65°04'W)	Zhong et al. 2002
SS4019	4	ND	Sargasso Sea (28°53'N, 65°04'W)	Zhong et al. 2002
SS4028	312	ND	Sargasso Sea (28°53'N, 65°04'W)	Zhong et al. 2002
SS4036	592	ND	Sargasso Sea (28°53'N, 65°04'W)	Zhong et al. 2002
SS4705	162	ND	Sargasso Sea (34°43'N, 68°07'W)	Zhong et al. 2002
SS4713	315	ND	Sargasso Sea (34°43'N, 68°07'W)	Zhong et al. 2002
SS4715	457	ND	Sargasso Sea (34°43'N, 68°07'W)	Zhong et al. 2002
SS4716	485	ND	Sargasso Sea (34°43'N, 68°07'W)	Zhong et al. 2002
SS4723	159	ND	Sargasso Sea (34°43'N, 68°07'W)	Zhong et al. 2002

Spatial distribution of cyanomyovirus populations

The T-RFLP profiles from the Inner Harbor and three stations in the Chesapeake Bay were strikingly similar to each other (Fig. 5-5) even though these sampling sites contained salinities ranging from 18 to 29‰ (Table 5-1). All OTUs differentiated by T-RFLP could be detected in the spatial profile. Compared with the model community profile consisting of eight T-RFs, more peaks (ca. 10 peaks on average) were present in the spatial profile. However, the relative abundances (as reflected by peak heights) of individual T-RFs vary from station to station in the spatial profile.

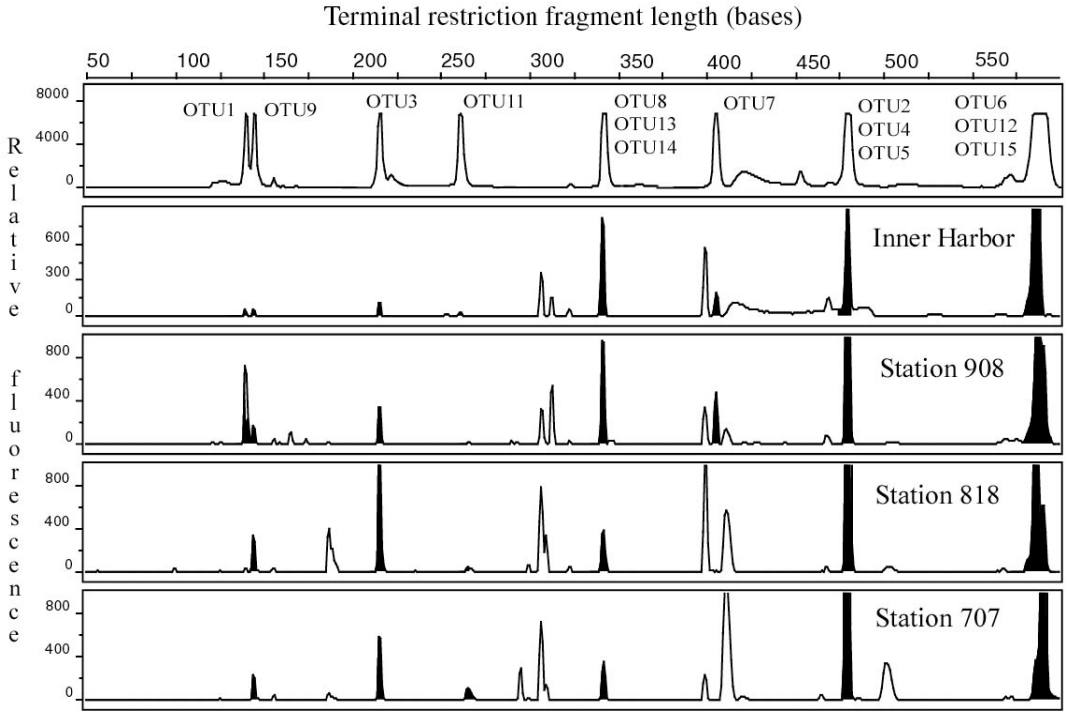


Fig. 5-5. Spatial T-RFLP profile of g20 genes of Chesapeake Bay stations (November 1-3, 2000). The shaded peaks shown here are the observed T-RFs those match with the expected sizes in the Table 5-2.

Temporal T-RFLP profile of cyanomyovirus populations in the Inner Harbor

PCR amplification for T-RFLP analysis was generally successful for samples collected during warm months other than the winter months. It was not possible to obtain the g20 amplification from VCs collected in April and May 2001 as well as February, April and May 2002 (data not shown). Although g20 amplification was obtained from VC prepared in March 2001, its T-RFLP profile was unclear. Therefore, only nine T-RFLP profiles were successfully obtained for samples with positive g20 amplifications. At least 25 T-RFs could be differentiated in the temporal profile, indicating the diverse g20 genotypes present in the Bay. In total, 18 peaks matching with predicted T-RFs (Table 5-2) were observed in the temporal profile. In contrast to spatial profiles, the temporal T-RFLP profiles of nine different months exhibited dramatic variation of genetic diversity (Fig. 5-6). About 12 major T-RFs were observed between July and September 2001 while only three major T-RFs were detected in January 2002. Correspondingly, both the host *Synechococcus* abundance and their phage titers (as determined by plaque assay using CB0101) reached the maximum level from July to September with a mean water temperature of 27°C and a salinity that varied only in a narrow range (ca. 11~13‰) during this period (Fig. 5-2 A). In general, more T-RFs were seen during summer time (from July to September 2001) than other seasons, while T-RFLP patterns of August and September 2001 were most similar to each other. Notably, the T-RFs of ca. 480 bp (as represented by OTU 2, 4, and 5 in November 2000 spatial profile) were not observed in the temporal profile from June 2001 to March 2002.

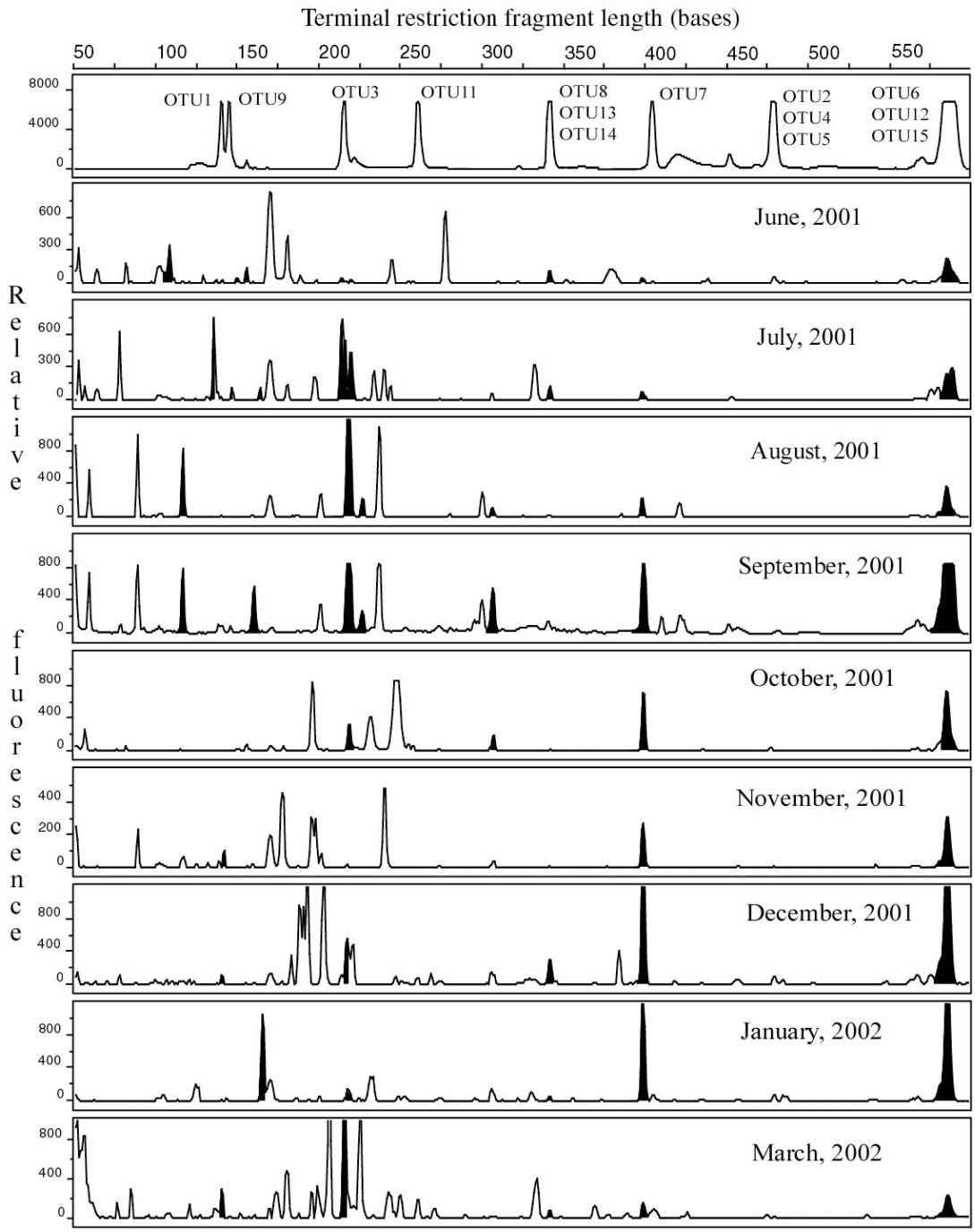


Fig. 5-6. Temporal T-RFLP profile of g20 gene at Baltimore Inner Harbor. Starting from June 2001, ended in March 2002. The shaded peaks shown here are T-RFs which match with the expected sizes in the Table 5-2.

Discussion

The unique but diverse composition of the cyanomyovirus population in the Chesapeake Bay was revealed by the g20 gene clone sequences and the subsequent T-RFLP analysis. It has been reported that the composition and structure of natural cyanophage communities in the estuary were different from those in the open ocean as determined by the g20 gene sequences (Zhong et al. 2002). About 43% of g20 OTUs (OTUs 2, 4 5, 8, 12, 13 and 14) in the Chesapeake Bay were clustered with the g20 clones recovered from the Savannah Estuary. In contrast, only a small percentage (11%) of g20 OTUs (OTUs 3 and 9) in the Bay appeared to be related to open ocean g20 clones. The g20 sequences recovered from environmental samples appeared to be much more diverse compared to those of cyanomyovirus isolates. The majority of identified cyanomyoviruses infecting *Synechococcus* spp. were isolated from pink or oceanic strains (Suttle and Chan 1993; Waterbury and Valois 1993; Suttle 2000). Recent studies in our lab suggested that *Synechococcus* communities in the Bay are more diverse than those in coastal and open ocean waters as revealed by *rbcL* gene phylogeny (Chapter 2 and 3). It is possible that the composition of host *Synechococcus* populations in the Bay is different from that in the oceanic water. Further characterization of cyanophages that infect indigenous *Synechococcus* strains may shed light on the g20 clonal sequences recovered from the Inner Harbor and Chesapeake Bay. In addition, we cannot rule out that some of the g20 environmental clones may come from phages that infect other bacteria. Much more effort is needed to characterize more g20 sequences from phages that infect different *Synechococcus* strains and other marine bacteria.

In the Chesapeake Bay samples, approximately 9 to 25 T-RFLP peaks or genotypes were visible for each sample. A previous study using DGGE analysis of g20 gene could differentiate 2 to 10 genotypes in each viral concentrate samples in the Atlantic Ocean (Wilson et al. 2000). It is difficult to compare the data obtained from DGGE and T-RFLP methods, because: 1) the samples were from different environments (estuary vs. open ocean); 2) PCR primers and amplicon length (ca. 592 bp vs. ca. 160 bp) were different; 3) detection methods for DGGE bands and T-RFLP peaks are different. The observation of many unpredicted T-RFs in both the spatial and temporal T-RFLP profiles also implied the possible presence of previously unidentified g20 genotypes in the Bay. The difference between expected and observed T-RFs together with some of the unpredicted T-RFs shown in the T-RFLP profiles could result from partial or unspecific restriction digestion and/or bias introduced during PCR amplification. This is indeed the inherent pitfall of the T-RFLP technique, which heavily relies on the PCR amplification accuracy and efficiencies of restriction enzymes. Nevertheless, care was taken to optimize the conditions for preparing T-RFLP samples in order to minimize these uncertainties. In this study, at least duplicate samples were analyzed and compared for each VC. The analysis and comparison between replicates of the same sample are therefore necessary to obtain statistically robust T-RFLP data.

In order to be consistent with T-RFLP analysis, the VC samples rather than raw seawater were used in the plaque assay. This may lead to the underestimations of natural viral titers due to the losses of viruses and/or their infectivity during the viral concentration processes. Failure to amplify the g20 gene from five VCs collected in

February (2002), April and May (2001 and 2002) could be due to the low titers of cyanophages during these time. The *Synechococcus* cell density and CB0101 cyanophage titers were very low during these months (Fig. 5-2). Other studies in Woods Hole Harbor (Waterbury and Valois 1993) and Tampa Bay (McDaniel et al. 2002) found that *Synechococcus* and their phages are more abundant in summer than in winter. The abundance of infectious cyanophages is known to correlate directly and strongly with the host concentrations rather than other environmental variables (Suttle 2000). The threshold in the abundance of *Synechococcus* was about 10^3 to 10^4 ml^{-1} in order for lytic virus production to occur (Suttle and Chan 1993; Suttle and Chan 1994). Therefore, the low abundance ($<10^3$ ml^{-1} observed in this study) of host *Synechococcus* during the winter and early spring seasons might not produce many lytic cyanophages. In this study, an indigenous *Synechococcus* isolate CB0101 was chosen as the host for plaque assay with the intention to better estimate the cyanophage abundance in the Chesapeake Bay.

The similar T-RFLP patterns of community fingerprints observed in four stations in Chesapeake Bay suggests that there is no significant variation of genetic diversity of cyanomyovirus communities in the surface water column of the Bay at the given time (November 2000). In contrast, the dramatic temporal change in g20 gene T-RFLP patterns observed in this study suggested that the cyanomyovirus community could be more diverse in summer than in winter. The DGGE profiles of the bacterial communities from the surface water column in the Chesapeake Bay also demonstrated a stronger seasonal pattern than spatial pattern (Kan et al. 2006).

The significant seasonal changes in both cyanophage titers and their genetic diversity in the Bay appeared to be correlated with their host *Synechococcus* populations. The maximum cyanophage diversity was also observed when their host *Synechococcus* abundance reached the annual maximum. This observation is consistent with a previous study in the Atlantic Ocean transect (Wilson et al. 2000). The dynamic interaction between cyanophage and *Synechococcus* communities observed in the Chesapeake Bay suggested that cyanophages could play important roles in regulating the *Synechococcus* biomass and population structure over time scales. Further investigation of spatial distribution of both cyanophages and their hosts across the Bay in different seasons will help us better understand the geographical variation of cyanophage populations.

Here, using the g20 gene, we demonstrated for the first time that T-RFLP of the g20 gene could be used as a rapid fingerprinting method to explore the population dynamics of a specific group of viruses in the aquatic environment. The PCR primers specific for *Synechococcus* spp. based on the RuBisCO gene are now available from our early work (Chapter 2). It is expected that a T-RFLP or DGGE based method can be used to obtain fingerprints of *Synechococcus* populations. This will allow us to study the co-variation and co-evolution of the cyanophage and host cyanobacterial communities in natural aquatic environments.

Chapter 6: Inter-annual survey in the Chesapeake Bay I: Abundance and distribution of *Synechococcus* and cyanophages

Abstract

Despite the increasing knowledge of *Synechococcus* and their co-occurring cyanophages in oceanic and coastal water, little is known about their abundance, distribution and interactions in the estuarine ecosystem. In this study, we found that *Synechococcus* and their phages were persistent and abundant microbial components in the Chesapeake Bay. The massive *Synechococcus* blooms (10^6 cells ml^{-1}) are often observed in summer throughout the Bay, and they could contribute on average 20-40% of total phytoplankton chlorophyll *a*. The distribution of phycoerythrin-containing (PE-rich) *Synechococcus* appeared to mostly correlate with salinity gradient. Cyanophages infectious to *Synechococcus* were also abundant (up to 6×10^5 MPN ml^{-1}) during summer months in the Bay. The co-variation in abundance of *Synechococcus* and cyanophage was evident but the correlation varied in different years, reflecting the changing environmental conditions. The impacts of cyanophage on host *Synechococcus* populations also varied spatially and temporally. Higher phage-related *Synechococcus* mortality was observed in drought years. These observations emphasize the influence of environmental gradients on natural *Synechococcus* and their phage population dynamics in the estuarine ecosystem.

Introduction

Unicellular cyanobacteria of the genus *Synechococcus* are among the most abundant (up to 10^5 ml^{-1}) and ubiquitous members of the picophytoplankton in the open ocean (Johnson and Sieburth 1979; Waterbury et al. 1979; Waterbury et al. 1986), and are responsible for ca. a quarter of primary production in the world oceans (Li et al. 1983; Waterbury et al. 1986; Li 1988; Partensky et al. 1999a).

Synechococcus are also abundant in various estuarine ecosystems including Chesapeake Bay (Ray et al. 1989; Affronti and Marshall 1993), Southampton estuary (Iriarte 1993; Iriarte and Purdie 1994), Bay of Blanes (Agawin et al. 1998), Florida Bay (Phlips et al. 1999), San Francisco Bay (Ning et al. 2000), Pensacola Bay (Murrell and Lores 2004) and Changjiang estuary (Pan et al. 2007). *Synechococcus* cell density typically ranges from 10^2 to 10^5 cells ml^{-1} in temperate estuaries, but often exceed 10^6 cells ml^{-1} in subtropical region (Murrell and Lores 2004). In temperate estuaries, the composition of phycoerythrin-rich (PE type) vs. phycocyanin-rich (PC type) *Synechococcus* appeared to vary with salinity gradient (Iriarte 1993) and temperature (Affronti and Marshall 1993). Chesapeake Bay differs from the above estuaries in its long residence time. The residence time of main Chesapeake Bay water can be 6-7 month long (Nixon et al. 1996) which allows the development of indigenous bacterial species in the Bay (Kan et al. 2006).

Estuarine ecosystems are so complex that each system may differ from the other in terms of environmental conditions. Chesapeake Bay is subject to various climatic forces that influence the growth and distribution of phytoplankton (Malone 1992). Chesapeake Bay is a temperate Bay where the annual temperature varies from

0 to 30ºC. The annual cycle of freshwater discharge typically demonstrates a spring maximum and a summer minimum. Over half of the annual nutrient input to Chesapeake Bay occurs during spring runoff (March through May). Consequently, the excess nutrient input often triggers phytoplankton (mainly diatom and dinoflagellate) blooms in Spring (Malone 1992). A few earlier studies reported that picophytoplankton (dominated by *Synechococcus*) could contribute 10 to 14% of Chesapeake Bay primary production during early summer (Ray et al. 1989). During late summer, *Synechococcus* can reach nearly10^6 cells ml^{-1} and account for 56% of primary production in the southern Bay (Affronti and Marshall 1994). Picoplankton productivity often exceeds heterotrophic bacterial productivity during summer months in the Bay (Malone et al. 1991). The relative contribution of PE type vs. PC type *Synechococcus* varied seasonally in the lower Chesapeake Bay (Affronti and Marshall 1993). The earlier studies on the Chesapeake Bay *Synechococcus* mainly focused on specific Bay regions and were not very systematic. A large scale and long term survey on the spatial and temporal dynamics of *Synechococcus* in the Chesapeake Bay has not been undertaken. Inter-annual variations on *Synechococcus* abundance and the mechanisms influencing such variations still remain unclear. No efforts have been taken to understand the interaction between *Synechococcus* and their phages in the Chesapeake Bay.

Grazing and viral lysis are the two main factors responsible for microbial mortality in aquatic environments (Pace 1988; Fuhrman and Noble 1995; Weinbauer and Hofle 1998a). Recent studies showed that *Synechococcus* appear to be a poor food source for nanoflagellates (Christaki et al. 2002). Only 0.2-1.0 % of natural

Synechococcus populations were consumed by grazing (Dolan and Šimek 1999). The lysis of *Synechococcus* by cyanophages (viral pathogens of cyanobacteria) appeared to have a significant impact on the mortality of *Synechococcus* and thereof their biomass and productivity (Suttle 2000; Mann 2003). It was estimated that viral lysis could cause up to 5-14% loss of *Synechococcus* cells in the open ocean on a daily basis (Suttle and Chan 1994). In the coastal water, cyanophage titers appeared to co-vary with *Synechococcus* abundance on temporal and spatial scales (Suttle and Chan 1993; Waterbury and Valois 1993; Marston and Sallee 2003). The abundance of infectious cyanophages was found to increase along the increasing salinity gradient in Savannah river estuary (Lu et al. 2001). Virioplankton could be an important factor affecting the biomass and population structure of bacterial community in the Chesapeake Bay (Wommack et al. 1992; Wommack et al. 1999a). The relationship between total viral and bacterial counts in the Chesapeake Bay has been studied (Wommack et al. 1992; Wommack et al. 1999b). No studies have been conducted to understand the interaction between specific virus and host systems in the Bay. *Synechococcus* and their phages are an ideal system for studying the host-virus dynamics at the population level, because; 1) they are abundant and dynamic in the marine environment; 2) total *Synechococcus* can be counted based on their autofluorescence; 3) cyanophages can be titered using representative host cultures.

To understand the ecological impacts of cyanophage on picocyanobacteria in the estuarine ecosystem, we co-monitored and the spatial and temporal distributions of *Synechococcus* and their co-occurring cyanophages in the Chesapeake Bay for five consecutive years (September 2002 to February 2007).

Materials and Methods

Sample collection

Water samples were collected from nine stations along the midstem of Chesapeake Bay (Fig. 5-1), on board the R/V *Cape Henlopen* and R/V *Hugh R. Sharp* during the research cruises for the MOVE (Microbial Observatory of Virioplankton Ecology) project from September 2002 to February 2007 (details of these cruises can be found at http://www.virusecology.org/MOVE/Home.html).

The location of nine stations is shown in Fig. 6-1. Their coordinates are as follows: Stn. 908 (39° 08' N, 76° 20' W), Stn. 858 (38° 58' N, 76° 23' W), Stn. 845 (38° 45' N, 76° 26' W), Stn. 834 (38° 34' N, 76° 26' W), Stn. 818 (38° 18' N, 76° 17' W), Stn. 804 (38° 04' N, 76° 13' W), Stn. 744 (37° 44' N, 76° 11' W), Stn. 724 (37° 24' N, 76° 05' W), Stn. 707 (37° 07' N, 76° 07' W). Stations 858, 804 and 707 are key stations representing of northern Bay, middle Bay and southern Bay region, respectively. During February 2004 cruise, it was not possible to sample Stn. 908 and 858 due the frozen condition in the northern Chesapeake Bay, and consequently there were no water samples collected from these two stations in this cruise. In addition, no water samples were collected from Stn. 845 during the July 2005 cruise.

Water samples were collected using 10-liter Niskin bottles mounted to a General Oceanics rosette (General Oceanics) from two depth: 1 m below the surface (designated as T) and 1 m above the bottom (designated as B). Middle layer water samples (designated as M) were also collected at the pycnocline or thermocline layers for most of stations, except for Stn. 908 where the water depth is relatively shallow. Upon collection, water samples were fixed immediately with glutaraldehyde (final

concentration 1%) and stored at 4 °C in the dark for *Synechococcus* counts; while 5 ml fixed water samples were stored at 4 °C in the dark (in 2002 to 2003 cruises) or in a liquid nitrogen reservoir (in 2004-2007 cruises) for bacterial and viral counts. Subsamples of 50 ml of surface water from each station were immediately frozen (-20°C) till later nutrient analysis.

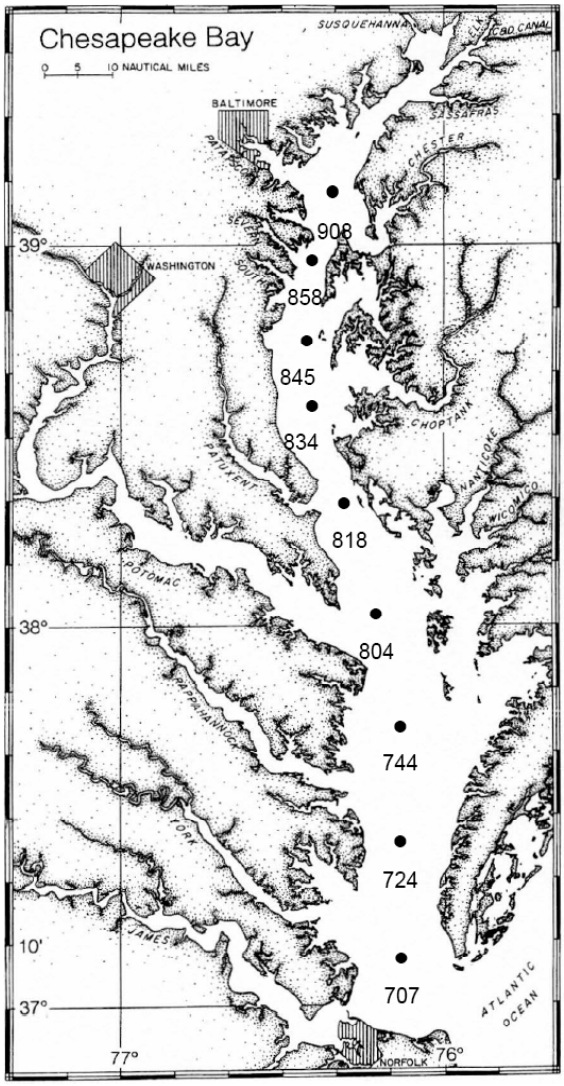


Fig. 6-1. Location of nine sampling stations in the Chesapeake Bay.

Chl *a* and nutrients concentrations

Dr. Wayne Coats at Smithsonian Environmental Research Center kindly provided Chl *a* data. Phytoplankton was size-fractionated into three classes: micro-size (>20 μm), nano-size (3-20 μm) and pico-size (< 3 μm). Duplicate samples (100 ml) from each station were vacuum filtered (<150 mm Hg) onto 25 mm Whatman GF/C54 filters and Chl *a* extracted with 90% acetone for 24 h at 4°C in the dark. Chl *a* concentration was determined using a Turner Designs 10-AUfluorometer. Chl *a* data from the March 2003 to October 2005 cruises were available for this study. The Chl *a* concentration of < 3 μm fraction from each water sample was assumed constituted of picocyanobacteria *Synechococcus*, and the data was compare with total Chl *a* concentration in this study.

Nutrient data including ammonia, nitrite and nitrate, and phosphate were determined using a Technicon AutoAnalyzer II at the Horn Point Analytical Services Laboratory (www.hpl.umces.edu/services/as.html). The analysis followed standard methods for chemical analysis of water and wastes proposed by USEPA (USEPA, 1983). Nutrient data included here are from the September 2002 to July 2006 cruises.

Direct counting of *Synechococcus*, bacteria and viruses

Synechococcus cells were counted using a Zeiss Axioplan (Zeiss, Germany) epifluorescence microscope equipped with 100-Watt mercury bulb (USH-102DH, USHIO Inc., Japan). Briefly, 3 to 50 ml of water sample was filtered onto a 0.2 μm pore-size 25 mm black polycarbonate membrane filter (Osmonics) and counted with a green light excitation filter set (Zeiss filter set 14: excitation BP 510-560, emission LP 590) at 1000 × magnification. The phycoerythrin-enriched (PC) *Synechococcus* cells autofluoresced in yellow to orange color while phycocyanin-enriched (PE)

Synechococcus cells autofluoresced in red under the above setting (Wood et al. 1985; Waterbury et al. 1986). At least 200 cells from 10 random fields were counted per sample.

Bacteria and viruses were enumerated following the protocol described in Chapter 5.

Most Probable Number (MPN) assay of cyanophage titers

Synechococcus WH7803 (MC-A strain) and CB0101 (MC-B strain) were used as host cells for cyanophage titer measurement. They were maintained and grown in SN medium as described previously (Chapter 2). The surface water samples of three key stations (Stn. 858, 804 and 707) from cruise August 2003 to February 2007 were assayed. MPN assay was performed following protocol described by Suttle and Chan (1994). Exponentially growing *Synechococcus* hosts (1.5 ml of 2×10^8 cells ml^{-1}) were mixed with 1.5 ml of 0.22 μm-filtered water samples. The 10-fold serial dilutions were performed in a 96-well microtiter plate. The controls received the same amount of microwave-treated water samples. Duplicates plates were performed for the same water sample and all the plates were incubated at 25 °C under 14 /10 h of light: dark cycle with 10-20 μE (microeinsteins) m^{-2} s^{-1} illumination. Cell lysis was monitored daily for up to 15 days and the number of observed host cell crash events (as determined by loss of pigments) at each dilution series were recorded. The cyanophage titers were estimated by using "MPN analyzer" software (available at http://www.geocities.com/cpsc319). Theoretically, the detection range of this MPN assay setting was from few to 10^7 cyanophages ml^{-1}.

Estimation of *in situ* contact rate of *Synechococcus* and their cyanophages

The contact rate at which cyanophage would encounter *Synechococcus* cells on a daily basis was calculated using the formula $R = (\mathrm{Sh}2\pi w D_v)VP$, where Sh is the Sherwood number (dimensionless) for *Synechococcus* cells (1.01), w is the cell diameter (mean = 1.5×10^{-4} cm), D_v is the diffusion constant for viruses (3.456×10^{-3} cm^2 day^{-1}); while V and P are *in situ* cyanophage abundance (measured by MPN assay for WH7803, see discussion below) and *Synechococcus* cell density (from direct counts), respectively (Suttle and Chan 1994). This formula can be simplified to $R = 3.29 \times 10^{-6}$ (ml day^{-1}) VP, when the above parameters were provided. The origin of these parameters were described in detail by Murray and Jackson (1992).

Estimation of impact of cyanophages on *Synechococcus*

Assuming that the standing stock of infectious cyanophage in natural waters resulted from the net balance between viral production and the removal of viruses due to viral decay at given time (Suttle and Chan 1994), the number of host cells (N_c) that would have to be lysed per day in order to produce the viral abundance measured by MPN can be calculated by introducing the average synechophage burst size of 83 (Chapter 4), by the formula $N_c = V(1+R_d) / 83$. V is the cyanophage abundance measured from MPN assay using *Synechococcus* WH7803 as host cell, the viral decay rates (R_d) used here are 0.03 day^{-1} (January, February and March) for winter months and 0.5 day^{-1} for the summer (June, July and August), while 0.12 day^{-1} is assumed here for the remaining month (Garza and Suttle 1998).

Phage type determination of the most abundant phage clones

To determine the type of the most abundant cyanophage clone at the end-point of serial dilutions from the MPN assay, PCR amplification with primers specific for

cyanomyovirus (based on the g20 gene, Chapter 5), and for cyanopodovirus (based on the DNA *pol* gene, Chapter 4) was used. The phage clones were assumed as siphoviruses if none of the above gene markers yielded a positive PCR amplification.

Results

Hydrographic conditions

As a temperate estuary, water temperature of the Chesapeake Bay varied dramatically in different seasons (Fig. 6-2 upper panel). The average surface water temperature from each cruise ranged from 1.7 °C (February 2007) to 28.3 °C (August 2005). The highest surface water temperature observed was 29.4 °C (Stn. 804, August 2005) while the lowest was -0.03 °C (Stn. 858, February 2007).

As a typical salt-wedge estuary, salinity in the Chesapeake Bay increased gradually from Stn. 908 to Stn. 707 (Fig. 6-2, middle panel). The contour plot of salinity in different seasons reflects the mixing of freshwater and oceanic water. The salinity of 9 sampling stations varied from 2.2 ppt (Stn. 908 in March 2003 and May 2004) to 26.8 ppt (Stn. 707, September 2002). After a prolonged drought period from 1999 to 2002, Chesapeake Bay received a large amount of precipitation through 2003 and 2004. In fact, year 2003 was the wettest year over a century (National Weather Service annul review report, http://www.nws.noaa.gov/). This dramatic change is reflected by the freshwater discharge from the three main rivers in the Chesapeake Bay (Fig. 6-2 lower panel). According to the USGS monthly water resources reports (http://waterdata.usgs.gov/nwis), freshwater flow in first five months of 2003 was double the amount in the same period of 2002. In 2003, the average freshwater flow

was about two fold of the average in July and August, and four fold above the average in September. Year 2003 and 2004 were featured as “deluge years” for the Chesapeake Bay, and increased freshwater flow consequently lowered the salinity throughout the Bay. The average salinity of 9 stations was 11.7 ppt for 2003 and 11.6 ppt for 2004, respectively, which were considerably lower than those in the “drought years” of 2005 (13.3 ppt) and 2006 (13.9 ppt).

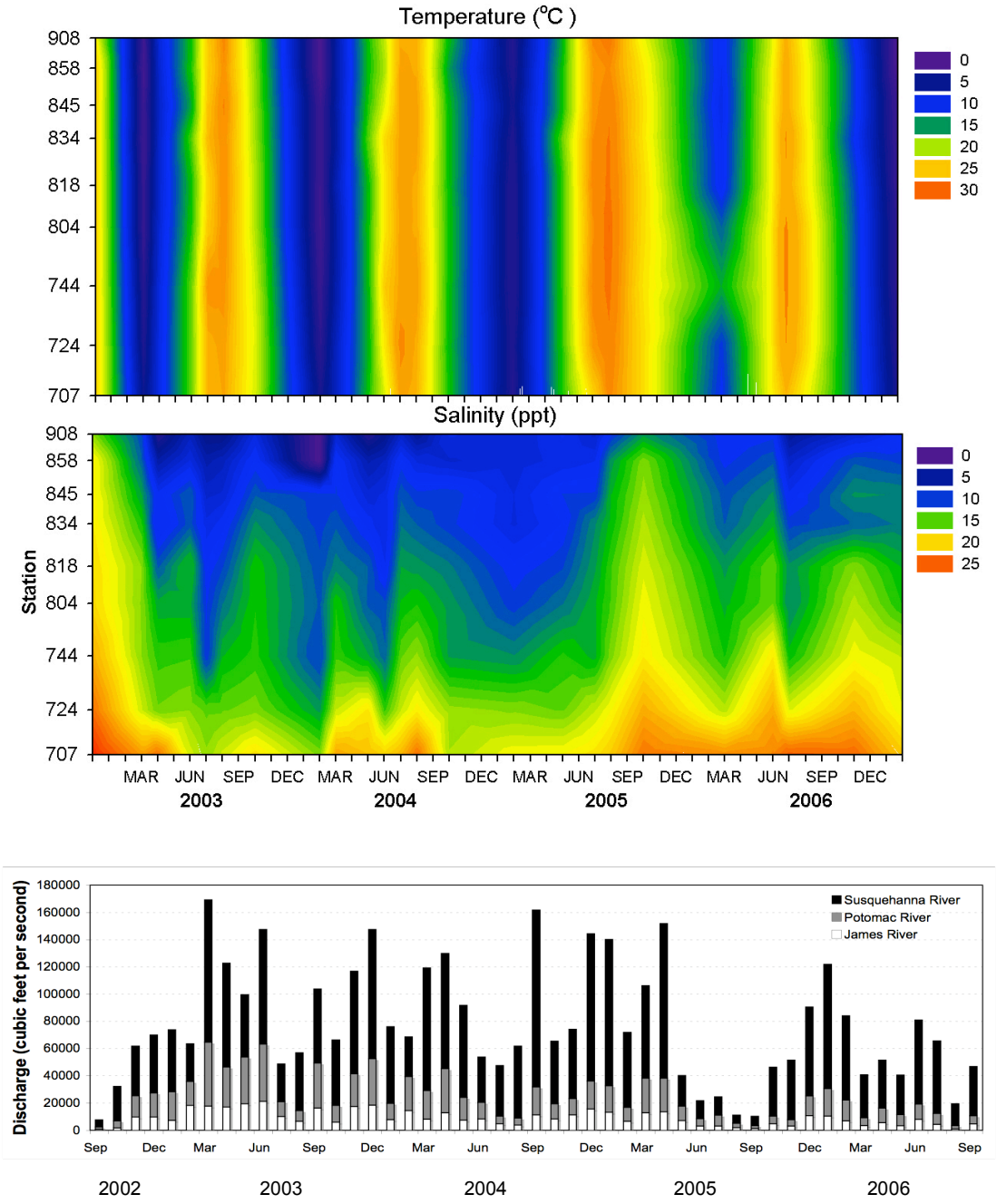


Fig. 6-2. Contour plots of Chesapeake Bay surface water temperature (upper panel) and salinity (middle panel) from September 2002 to February 2007. Lower panel: mean monthly freshwater flow from three main rivers to Chesapeake Bay from September 2002 to September 2006. The data were obtained from the USGS Water Resources Division (http://waterdata.usgs.gov/nwis); all the data are preliminary.

The nutrient concentration (N and P) in the Chesapeake Bay surface water also demonstrated significant seasonal and spatial variations (Fig. 6-3). The measured total nitrogen (nitrate, nitrite and ammonium) concentration was generally higher in the northern Bay (Stn. 858) compared with the middle (Stn. 804) and southern Bay (Stn. 707). The average concentration of nitrate and nitrite ranged from 26.2 μM (Stn. 858), 10.7 μM (Stn. 804) to 2.9 μM (Stn. 707) and appeared to have a negative linear relationship with water salinity (Fig. 6-4, upper panel). In contrast, there was no obvious correlation between ammonium or phosphorous and salinity (Fig. 6-4, middle and lower panel). The concentration of nitrate and nitrite appeared to be high in the cold seasons (winter and early spring) and low in summers (from June to August). Phosphorous (phosphate) concentrations were highest in the southern Bay February 2004, and also showed peaks in July 2003, August 2003 and October 2005.

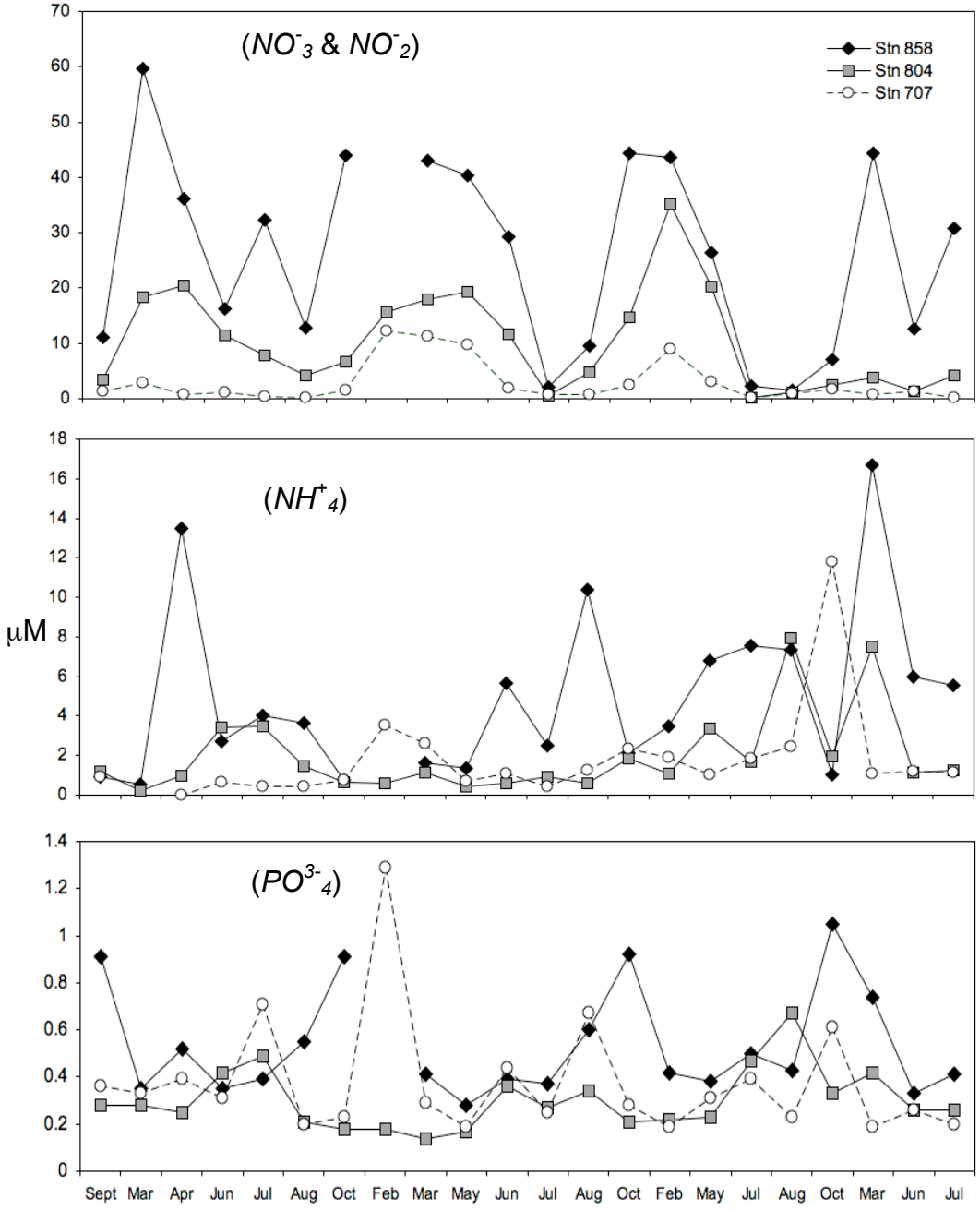


Fig. 6-3. Surface water nutrient concentrations at three stations representing northern (Stn 858), middle (Stn 804) and southern (Stn 707) Chesapeake Bay. Data were from September 2002 till July 2006. Note that no nutrient data for February 2004 for Stn. 858, due to the frozen condition in northern Bay.

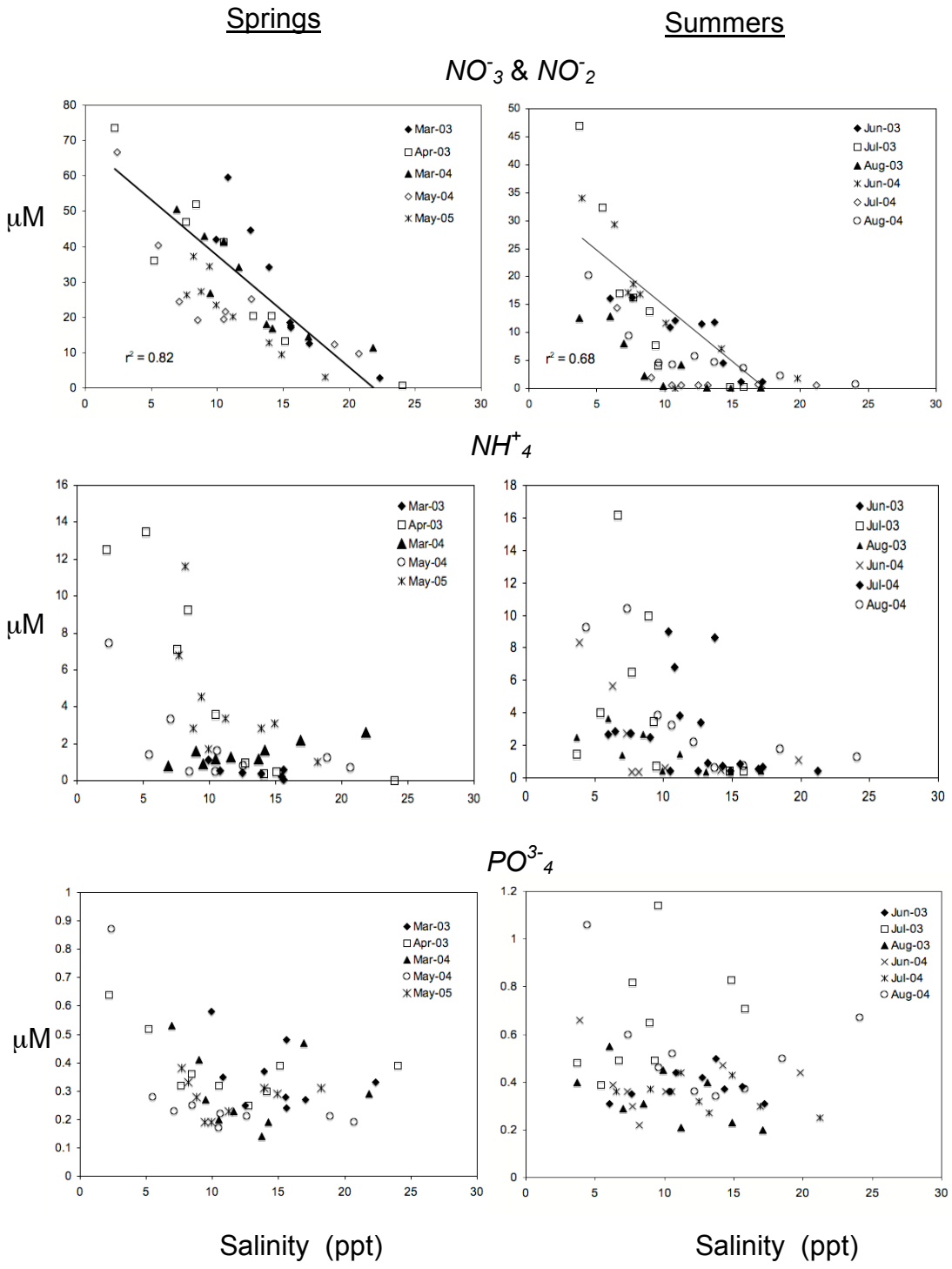


Fig. 6-4. Relationships between nutrient concentrations and salinity of Chesapeake Bay surface water.

Temporal and spatial variation of *Synechococcus* abundance

A temperature driven pattern of *Synechococcus* abundance was obvious and repeatable annually from 2003 to 2006 (Fig. 6-5. upper and middle panel). *Synechococcus* abundance is positively correlated with water temperature (Fig. 6-6), and often peaked in summers (June to August). The cell density of *Synechococcus* exceeded 1 million cells ml^{-1} in summers of 2003, 2004 and 2006. The highest cell density (3.3×10^6 cells ml^{-1}) was recorded at Stn.818 (surface water sample) in June cruise 2006. Despite the strong salinity and nutrient (particularly nitrate and nitrite) gradients along the Bay, total *Synechococcus* counts remained relatively stable on the spatial scale compared to the temporal variation. During the summer blooms, *Synechococcus* could make up to 32% of total bacterial abundance (Fig. 6-5. lower panel), and contributed 20-40% (up to 80 %, Stn. 804 in July 2005) of total phytoplankton chlorophyll concentrations (Fig. 6-7). *Synechococcus* were responsible for up to 50-60% of total primary production in the southern Bay stations (Dr. Wayne Coats, personal communication). In winter and early spring, *Synechococcus* cell density decreased dramatically to a few hundred cells per milliliter. The lowest *Synechococcus* (67 cells ml^{-1}) was recorded at Stn. 744 (bottom water) in April 2003. The annual average of *Synechococcus* cell density was 6.97×10^4 cells ml^{-1} for bottom waters, and 3.00×10^5 cells ml^{-1} for surface waters. Notably, *Synechococcus* abundance was significantly lower in 2005 summers (on average three fold) compared with other summers (Fig. 6-7). This resulted in the lowered percentage of *Synechococcus* to total bacterial counts (Fig. 6-5. lower panel), but not the contribution to total phytoplankton Chl *a* concentration (Fig. 6-7).

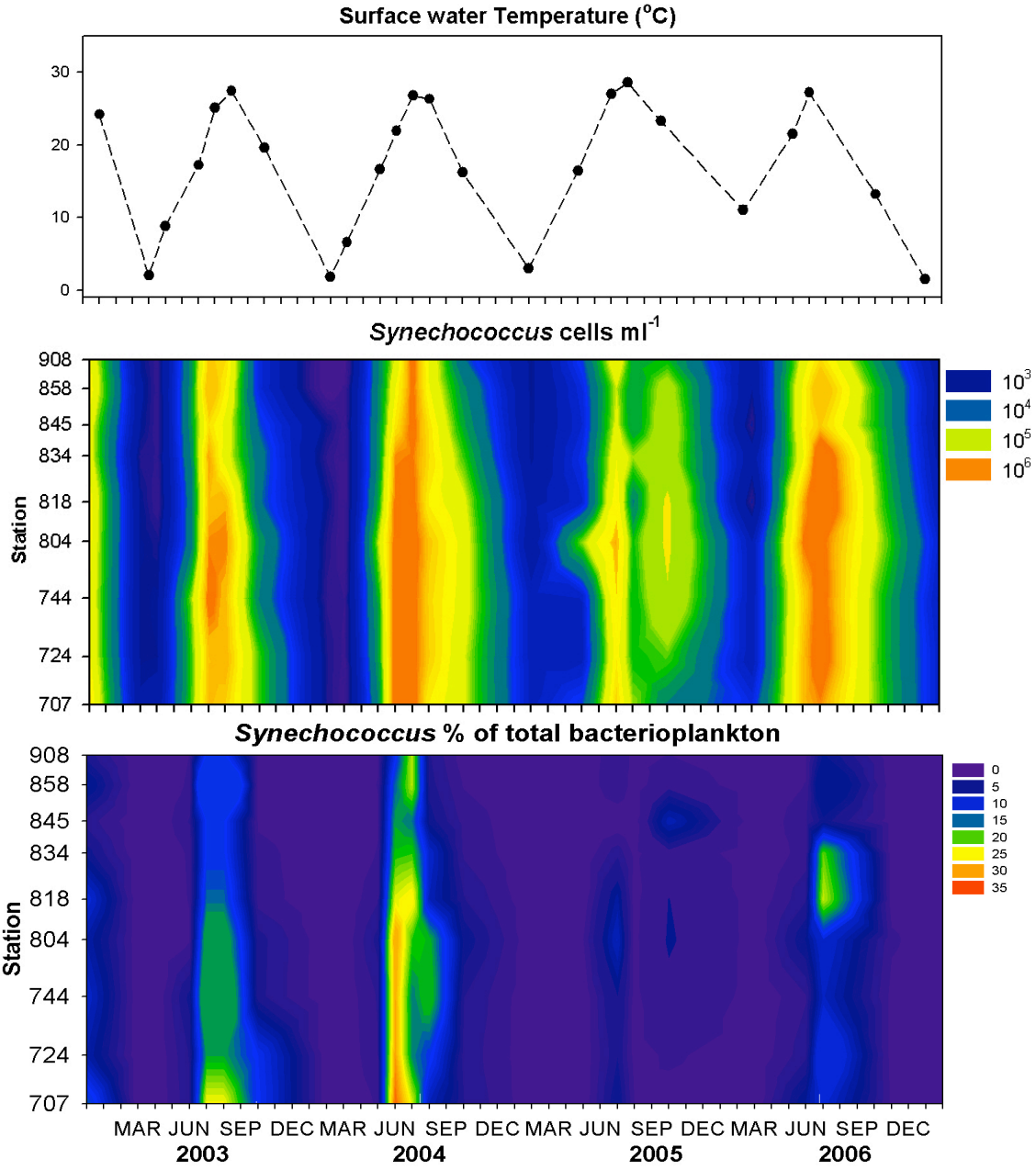


Fig. 6-5. Average surface water temperature (upper panel), *Synechococcus* cell density (middle panel) and its percentage to total bacterial counts (lower panel) in the Chesapeake Bay.

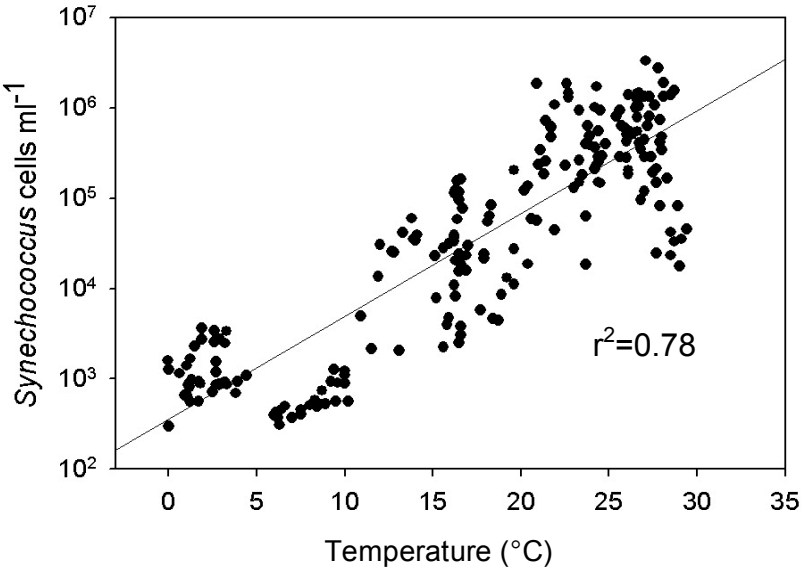


Fig. 6-6. Relationship between *Synechococcus* abundance and water temperature (surface water samples, n=213).

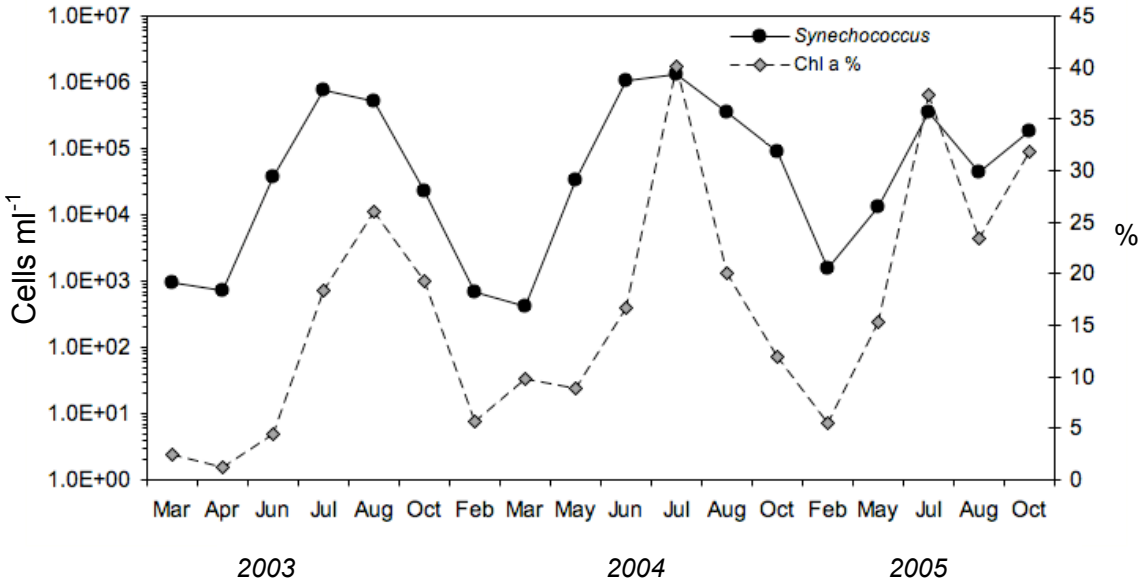


Fig. 6-7. Average *Synechococcus* abundance in the Chesapeake Bay and its contribution to total Chl *a* concentration. Each data point represents the mean of nine stations. No Chl *a* data are available after October 2005.

The composition of phycocyanin-enriched (PC) and phycoerythrin-enriched (PE) *Synechococcus* varied greatly from the northern to southern Bay (Table 6-1). There was pronounced seasonal as well as spatial variations in the distribution of PE vs. PC type *Synechococcus* in the Chesapeake Bay. During summer blooms, the PC type *Synechococcus* were dominant (> 75%) in the surface water in the northern Bay region, and the percentage of PE type cells gradually increased from the northern to southern Bay (Table 6-1).

Table 6-1. Percentage of PE type *Synechococcus* cells observed at three Chesapeake Bay stations from June 2004 to November 2006 (T: surface water; M: middle layer water; B: bottom water).

Station		2004				2005					2006			
		Jun	Jul	Aug	Oct	Feb	May	Jul	Aug	Oct	Mar	Jun	Jul	Nov
Stn.858	T	1.5	11.6	5.6	30.2	94.5	19.1	17.3	18.3	11.1	77.8	24.9	2.4	35.6
	M	2.0	20.2	12.1	46.3	87.9	21.4	27.8	48.0	24.6	59.5	25.5	6.7	55.5
	B	12.2	32.8	32.2	44.5	77.4	80.0	37.3	33.6	25.7	60.6	24.8	51.9	61.3
Stn. 804	T	32.1	38.3	38.1	51.4	92.8	7.5	64.0	88.3	61.2	98.4	88.4	21.6	70.6
	M	54.3	38.6	36.4	57.8	96.6	48.1	72.1	54.4	48.1	100	81.7	26.3	83.2
	B	59.2	38.6	40.9	62.7	86.7	63.2	95.6	74.7	69.0	100	76.8	54.6	83.8
Stn. 707	T	60.3	38.3	47.0	60.8	93.2	55.3	79.4	72.2	88.9	100	75.8	53.5	79.7
	M	51.2	43.9	N.D.	65.3	99.5	56.1	60.4	60.6	76.9	100	84.8	73.6	91.9
	B	56.1	43.9	50.8	73.6	98.5	82.1	98.0	76.5	88.9	100	74.2	74.1	94.1

N.D.: no data.

In wintertime, PE type *Synechococcus* were predominant throughout the Bay, and could make up ca.100% of the total picocyanobacterial community in the middle and southern Bay regions. The ratio of PC vs. PE type cells appeared to be correlated with the salinity gradient in summer months (Fig. 6-8). A higher proportion of PE

type *Synechococcus* was observed in 2005 and 2006 summers compared to 2004. The percentage of PE type *Synechococcus* appeared to increase from surface to bottom waters in summer (Table 6-1).

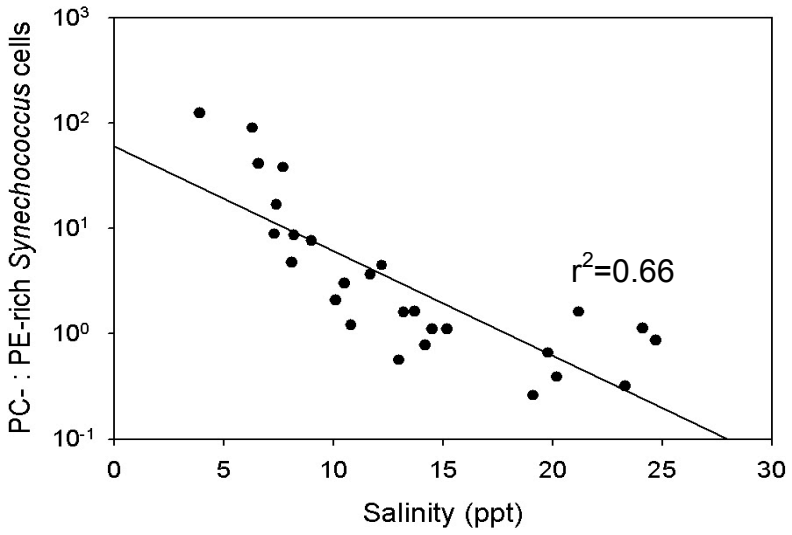


Fig. 6-8. Relationship between PC-type vs. PE-type *Synechococcus* and salinity in surface water during summer months from 2004 to 2006 (n=29).

Synechococcus phage titers and distributions

The titers of cyanophages infective to *Synechococcus* WH7803 ranged from a few to 6.2×10^5 infectious units ml^{-1} (Fig. 6-9 and Table 6-2). In parallel to the total *Synechococcus* counts, WH7803 phage titers were high in summer and low in winter, and exceeded the total *Synechococcus* counts in some cases. The middle Bay region (Stn. 804) had higher cyanophage titers with annual average of 6.9×10^4 ml^{-1}, compared to the northern Bay (2.6×10^4 ml^{-1}) and southern Bay (4.6×10^4 ml^{-1}). The overall mean infectious synechophage titers (including all the seasonal and spatial data) in the Bay was 4.7×10^4 ml^{-1}, which was about one order of magnitude less abundant compared with their host abundance (3.0×10^5 ml^{-1}).

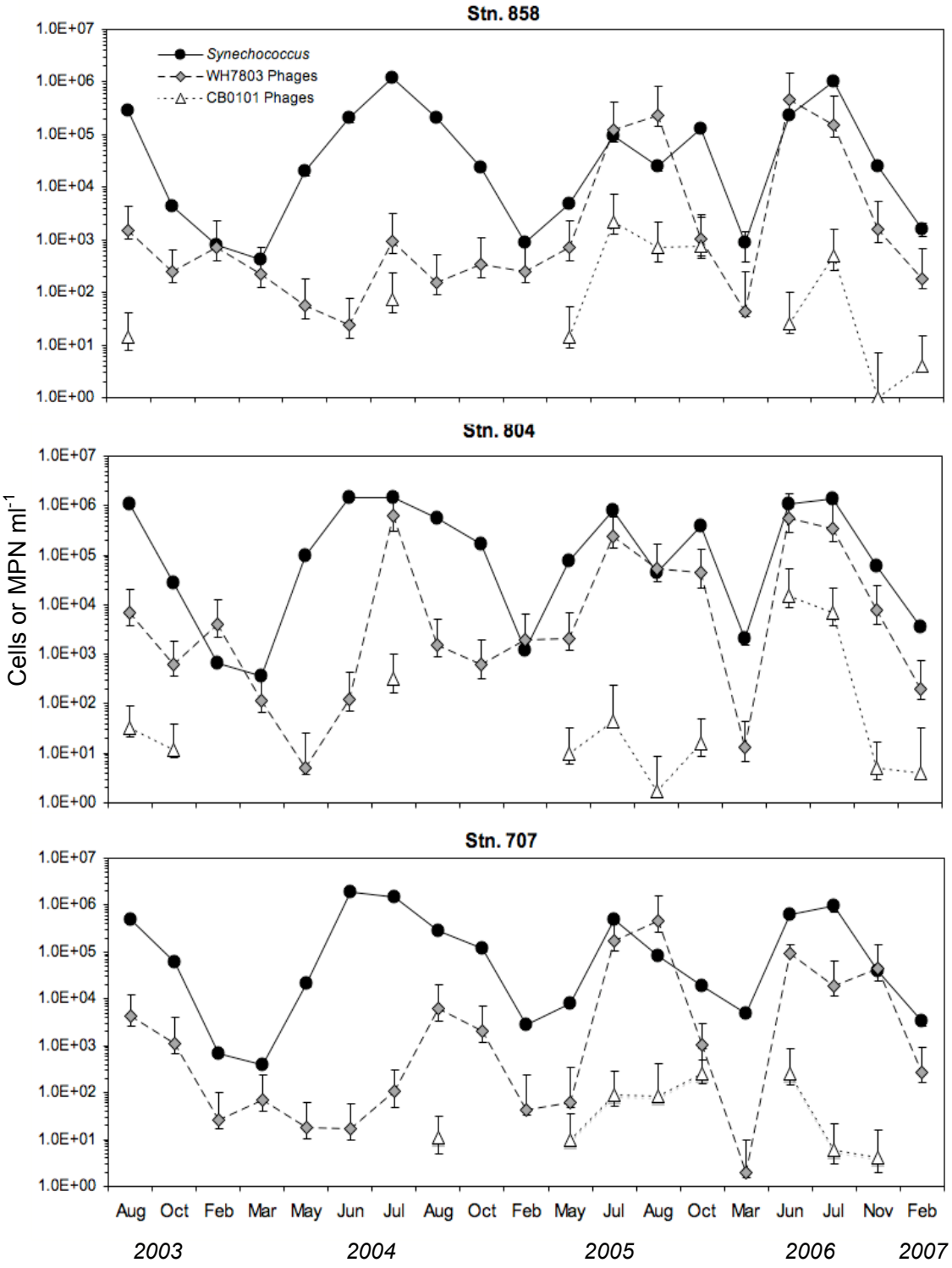


Fig. 6-9. Surface water *Synechococcus* cell density, WH7803 phage and CB0101 phage abundance at three Chesapeake Bay stations, from August 2003 to February 2007. Note that in February 2004, data from Stn. 845 was used to represent northern Bay.

The phage titers measured with *Synechococcus* strain CB0101, the strain isolated from Chesapeake Bay was remarkably lower than WH7803 phage titers, and were undetectable in nearly half of water samples (Fig. 6-9). In general, WH7803 phage titers are several hundred folds higher than CB0101 phage titers in summer (Table 6-2). Despite their low abundance, CB0101 phage titers showed a distribution pattern similar to WH7803 phage titers, high in summer and low in winter.

Together, WH7803 and CB0101 phage titers could contribute more than 2% of the total viral abundance in few cases, however, in general they were below 2% of total viral counts (Table 6-2).

Table 6-2. Summary of cyanophages infecting WH7803 and CB0101 obtained by MPN assay from summer samples at three Chesapeake Bay stations.

	Syn-cells ml^{-1}	WH7803 phage MPN (ml^{-1})	CB0101 phage MPN (ml^{-1})	WH7803 vs. CB0101 phage titer ratio	Most abundant WH7803 phage type[a]	Cyanophage% of total *VLPs*[b]	Lysed Syn%[c]
			July 2004				
Stn. 858	1.23E+06	9.25E+02	7.2E+01	1.28E+01	Podo	< 0.1	< 0.1
Stn. 804	1.44E+06	6.21E+05	3.1E+02	1.98E+03	Sipho	2.7	0.8
Stn. 707	1.46E+06	1.05E+02	N.D.	N.A.	Podo	< 0.1	< 0.1
Mean	1.38E+06	2.07E+05	1.93E+02	1.07E+03	N.A.	0.9	0.3
			August 2004				
Stn. 858	2.04E+05	1.53E+02	N.D.	N.A.	Sipho?	< 0.1	< 0.1
Stn. 804	5.62E+05	1.53E+03	N.D.	N.A.	Podo	< 0.1	< 0.1
Stn. 707	2.89E+05	3.68E+03	1.10E+01	3.35E+02	Podo	< 0.1	< 0.1
Mean	3.52E+05	1.79E+03	3.67E+00	4.87E+02	N.A.	< 0.1	< 0.1
			July 2005				
Stn. 858	9.63E+04	1.23E+05	2.19E+03	5.63E+01	Myo	1.5	2.3
Stn. 804	7.98E+05	2.34E+05	4.30E+01	5.44E+03	Sipho?	2.3	0.5
Stn. 707	4.84E+05	1.70E+05	9.14E+01	1.86E+03	Myo	1.5	0.6
Mean	4.59E+05	1.75E+05	7.74E+02	2.27E+02	N.A.	1.8	1.1
			August 2005				
Stn. 858	2.45E+04	2.34E+05	7.13E+02	3.28E+02	Myo	1.3	17.2
Stn. 804	4.55E+04	5.42E+04	2.00E+00	2.71E+04	Myo	0.2	2.2
Stn. 707	8.20E+04	4.62E+05	8.40E+01	5.50E+03	Podo	2.4	10.2
Mean	5.07E+04	2.50E+05	2.66E+02	9.39E+02	N.A.	1.3	9.9
			Jun 2006				
Stn. 858	2.37E+05	4.59E+05	2.60E+01	1.77E+04	Myo	0.3	3.5
Stn. 804	1.08E+06	5.42E+05	1.54E+04	3.52E+01	Myo	0.5	0.9
Stn. 707	6.16E+05	9.60E+04	2.53E+02	3.79E+02	Sipho?	0.2	0.3
Mean	6.44E+05	3.66E+05	5.23E+03	7.00E+01	N.A.	0.3	1.6
			July 2006				
Stn. 858	1.01E+06	1.55E+05	4.97E+02	3.12E+02	Myo	0.6	0.3
Stn. 804	1.39E+06	3.42E+05	7.03E+03	4.87E+01	Myo	0.5	0.4
Stn. 707	9.38E+05	1.92E+04	6.00E+00	3.20E+03	Sipho?	0.1	< 0.1
Mean	1.11E+06	1.72E+05	2.51E+03	6.87E+01	N.A.	0.4	0.35

a. Cyanomyoviruses and cyanopodoviruses were determined by PCR method (see Method section). Siphoviruses (with question marker) need further confirmation either by sequencing or TEM observation.

b. Refers sum of WH7803 and CB0101 phage MPN abundance to total VLPs counts.

c. Calculated based on WH7803 phage MPN abundance.

N.D.: Not detectable.

N.A.: Not applicable.

Impact of cyanophages on *Synechococcus* abundance

Estimated based on the WH7803 – phage system, a strong correlation between cyanophage abundance and host cell density was seen for 2006; however, such a relationship was not obvious in 2004 and 2005 (Fig. 6-10). High phage-host contact rates were often seen in summers than in wintertime (Fig. 6-11). During the summer *Synechococcus* blooms, each *Synechococcus* cells could encounter infectious cyanophage twice on a daily basis (i.e. 204% contact rate). In contrast, less than 0.1% of *Synechococcus* cells will collide with infectious cyanophages per day during the winter season (Fig. 6-11). The phage-host contact rates also varied dramatically on the spatial scale. For example, a range of 0.03 to 204% of contact rate was seen from the southern to middle Bay in July 2004. Considering all seasonal samples, the average phage-host contact rates were 32.2, 16.9 and 13.9% for the middle, northern and southern Bay, respectively.

The percentage of *Synechococcus* lysed by viral infection (infection rate) followed a trend similar to that of the contact rates (Fig. 6-11 and Table 6-2). Overall, infection rate varied from 0.7 to 1.4% from the southern to northern Bay, considering all seasonal samples. High lysis rates were often seen in the summer months. About 1-10% of *Synechococcus* cells could be lysed by cyanophage in the summer 2005, Approximately, 0.35-1.6% *Synechococcus* cells were lysed in the summer 2006, but less than 0.8% *Synechococcus* were lysed in the summer 2004.

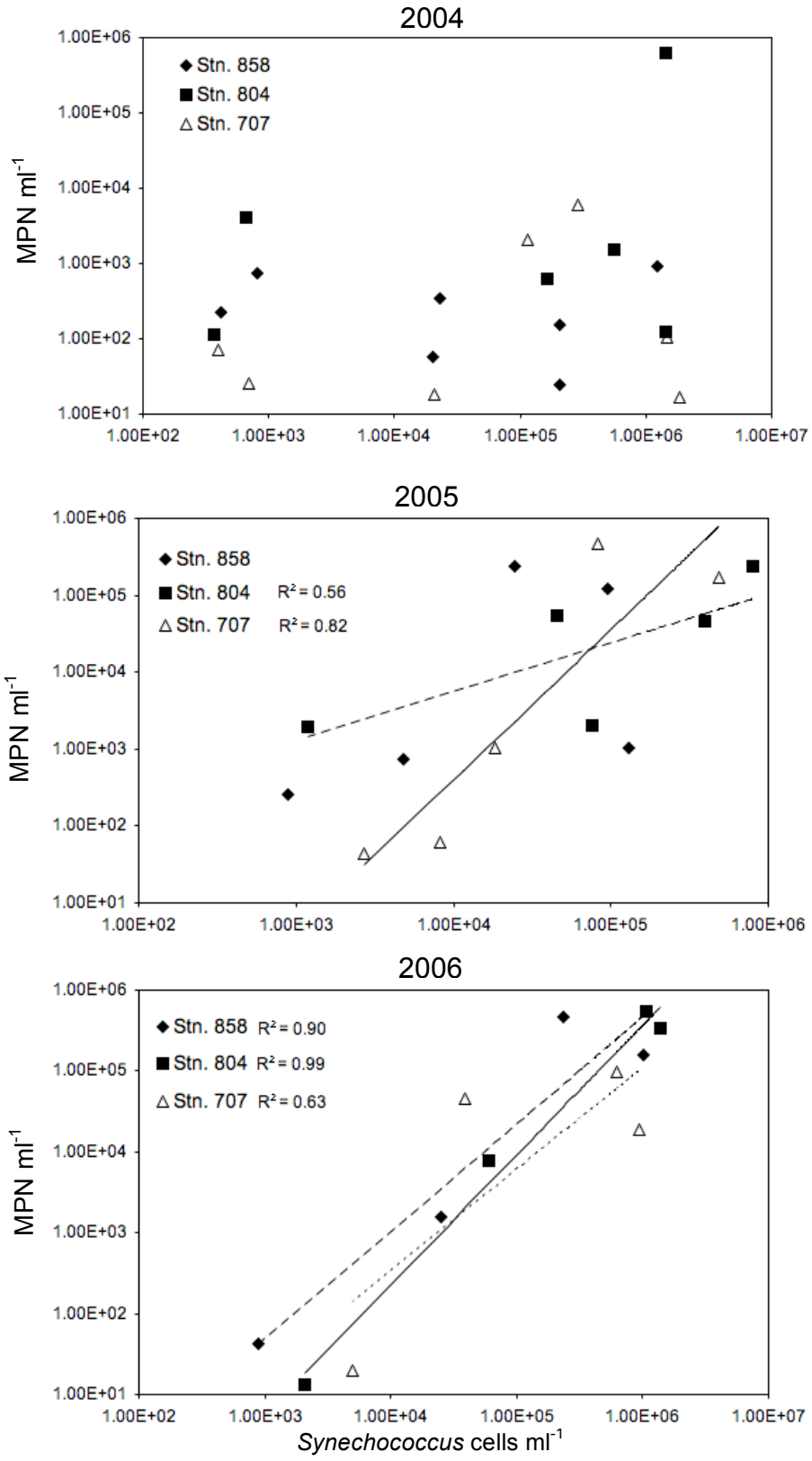


Fig. 6-10. Relationship between *Synechococcus* abundance and their phage titer (MPN ml $^{-1}$ measured with WH7803) in the Chesapeake Bay from 2004 to 2006. Note that the cell abundance and phage titers are log-transformed.

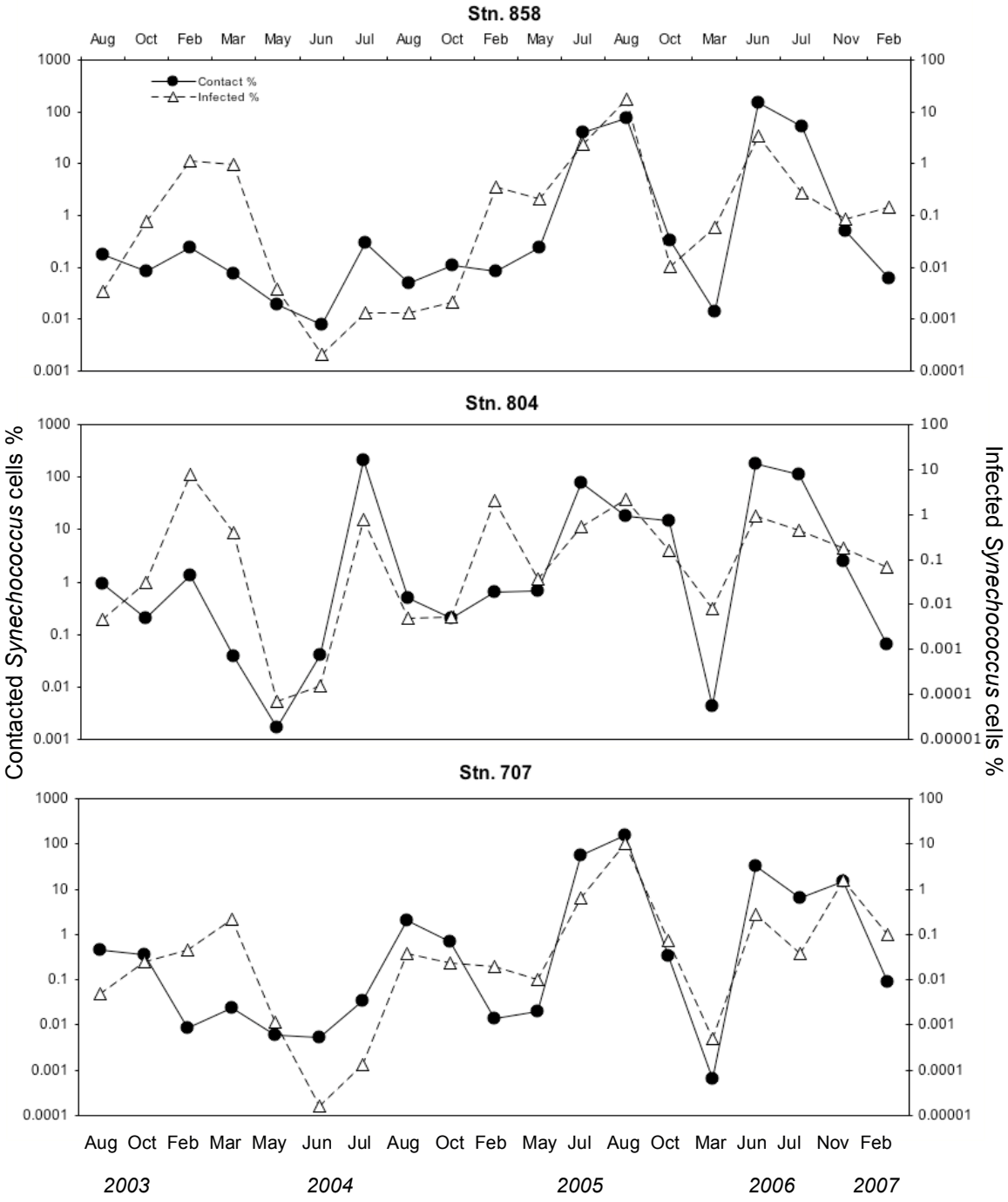


Fig. 6-11. Percentage of *Synechococcus* cells contacted and lysed by cyanophages per day at three Chesapeake Bay stations from August 2003 to February 2007. Percentages were estimated from the total *Synechococcus* counts and cyanophages infecting WH7803. Note that the data are log-transformed.

Most abundant *Synechococcus* phage types

Three different types of cyanophages (myo-, podo- and siphovirus) could be detected in high abundance (>10^5 ml^{-1}) in the summer samples (Table 6-2). The occurrence frequency of each phage type varied in different years. The most abundant *Synechococcus* WH7803 phage type was either podovirus or siphovirus in 2004 summers but no myoviruses could be detected by PCR with g20 gene primers. While cyanomyoviruses could be the dominant type in the summer of 2005 and 2006, as 67% of MPN end-point lysates were detected positively for cyanomyovirus g20 gene PCR amplification (Table 6-2). Overall, cyanomyoviruses account for 44% (eight out 18 samples) of the most abundant phage clones.

Discussion

This multi-year survey showed that the total *Synechococcus* counts and cyanophage titers co-vary and exhibit strong seasonal patterns in the Chesapeake Bay. A strong correlation between water temperature and *Synechococcus* abundance was observed in the Bay over the four-year period survey (Fig. 6-6). The temperature driven annual pattern of *Synechococcus* has been observed in temperate coastal waters (Waterbury and Valois 1993; Marston and Sallee 2003) and estuarine waters (Affronti and Marshall 1994; Agawin et al. 1998; Ning et al. 2000). Warm temperature perhaps favors the rapid growth of *Synechococcus* in the Bay. *Synechococcus* in the Chesapeake Bay often exceeded 10^6 cells ml^{-1} during summer months. The annual average of *Synechococcus* in the Bay is 3×10^5 ml^{-1}, which is 10 -100 fold more abundant than those reported for open oceans (typically 10^3-10^4 cells

ml^{-1}) (Li 1998) and North Atlantic coastal water (Waterbury and Valois 1993; Marston and Sallee 2003). The abundance of *Synechococcus* in Chesapeake Bay is comparable with those reported in San Francisco Bay (Ning et al. 2000) and sub-tropic Florida Bay estuarine ecosystems (Phlips et al. 1999; Murrell and Lores 2004).

No clear correlation between *Synechococcus* abundance and the concentrations of nutrients (nitrate plus nitrite, ammonia and phosphate) was observed in the Chesapeake Bay. The nutrient concentrations in the Bay are usually high during the winter-spring period, mainly due to the strong river run-off in that period (Malone 1992). It is known that *Synechococcus* can utilize various nitrogen sources for growth (Waterbury et al. 1986; Collier et al. 1999; Moore et al. 2002; Palenik et al. 2003). High nutrient level in the Bay and ability to utilize various type of nitrogen sources may explain the week relationship between *Synechococcus* abundance and the level of inorganic nutrients.

Although salinity has little impact on the spatial variation of total *Synechococcus* counts, the distribution of PC vs. PE type *Synechococcus* was affected by the salinity gradient. PC type *Synechococcus* are dominant in the northern Bay, while more PE type *Synechococcus* are more abundant from middle to southern Bay. A positive correlation was seen between the percentage of PE type *Synechococcus* and salinity along the north-south transect in the Bay, indicating that salinity could be an important factor influencing the distribution of PE and PC type of *Synechococcus* on the spatial scale. Inter-annual variation on percentage of PE *Synechococcus* also appeared to correlate with salinity changes. In the deluge year of 2004, the percentage

of PE cells in the summer surface water was ca. 38% in the middle Bay, but climbed to 64-88% during July and August in drought year of 2005.

Cyanophages infecting *Synechococcus* WH7803 could be detected in all seasons and different locations in the Chesapeake Bay, indicating that they are ubiquitous and persistent microbial component in the estuarine environments. The high abundance of *Synechococcus* phages (commonly > 10^5 ml^{-1}) detected during summertime suggested that they are important members in the viral assemblages in the Chesapeake Bay. Recent metagenomic analysis of Chesapeake Bay virioplankton also revealed a high proportion of cyanophage sequences in late summer (September 2002) sample (Bench et al. 2007). The dynamic variation in *Synechococcus* phage titers indicates they are active pathogens to their host populations. *Synechococcus* WH7803 is a sensitive strain to phage infection and has been commonly used for titering cyanophages in many studies (Suttle and Chan 1994; Lu et al. 2001; Marston and Sallee 2003; Sullivan et al. 2003; Mühling et al. 2005). Nonetheless, the phage titers obtained using WH7803 by no mean represent all the cyanophages of *Synechococcus*.

The tight co-variation between *Synechococcus* abundance and their phage titers in the Chesapeake Bay indicate that cyanophages could have the potential for regulating host populations. The average lysed *Synechococcus* cells per day in summers was 2.2 %, which is in good agreement with previous TEM observation (0.8 -2.8%) of natural *Synechococcus* populations (Proctor and Fuhrman 1990). The average *Synechococcus* infection rate (considering all seasons) by cyanophages was 0.96%, which is lower than that reported for Gulf of Mexico waters (Suttle and Chan

1994), but higher than that of Woods Hole harbor water (Waterbury and Valois 1993).

A great variation on percent infected *Synechococcus* cells was observed between the two summers of year 2004 and 2006. In the summer of deluge year (2004), the contact rate between hosts and infectious cyanophage was remarkably high (up to 204% day^{-1} in Stn. 804), due to the high abundance of host *Synechococcus* and their co-occurring cyanophages. Less than 1% of the *Synechococcus* was infected to produce the observed phage titers, which is consistent with what has been reported for the Woods Hole Harbor water (Waterbury and Valois 1993). In August 2005, 77% *Synechococcus* were contacted by cyanophage on a daily basis, and 17.2% of *Synechococcus* population were possibly lysed. At the same time, the *Synechococcus* cell abundance was ca. 2.5×10^4 cells ml^{-1}. The *Synechococcus* VBR (virus to bacteria ratio) was ca. 10, which was the highest one observed in this study. Also notably, the overall cyanophage infection rate appeared to be high throughout the Bay in August 2005, as 10.2 % and 2.2% *Synechococcus* could be infected in the lower and middle Bay, respectively.

Podoviruses and siphoviruses appeared to be predominant phage types during 2004 summer, while myoviruses prevailed in summer 2005. Cyanomyoviruses are known to have a broad host range, while podo- and siphoviruses are usually strain specific (Suttle and Chan 1993; Waterbury and Valois 1993; Chen and Lu 2002; Sullivan et al. 2003). The distinct feature between different types of cyanophages may account for the observed difference in infection rates between years. As

cyanomyoviruses are able to infect broader range of host cells, they may have greater impact on overall host abundance compared with host-specific cyanophages.

In conclusion, highly abundant *Synechococcus* and their phages were observed repeatedly in the summer waters in the Chesapeake Bay, indicating that they both are important microbial components in this eutrophic estuary. The seasonal co-variation between *Synechococcus* and their phage abundance suggests that they have active and dynamic interactions in the Bay. The cyanophages could infect a substantial portion of *Synechococcus* during summer blooms, indicating that they may play important roles on regulating host population, and subsequently influence the carbon fixation and nutrient flow in the estuarine ecosystem. In the Chesapeake Bay, the impacts of cyanophages on *Synechococcus* mortality varied under different conditions, higher in the drought years but lower in the deluge years. We speculated that the variance in environmental gradients might influence the host composition and consequently affect the viral impacts on host population.

Chapter 7: Inter-annual survey in the Chesapeake Bay II: Population succession of *Synechococcus* and cyanophages

Abstract

Our earlier study has demonstrated that cyanophage titers co-vary with host *Synechococcus* abundance in the Chesapeake Bay (Chapter 6). However, it is not known how the composition of *Synechococcus* and cyanophages changed in the Chesapeake Bay over time and space. In this study, seasonal succession of *Synechococcus* and cyanomyovirus (one of the three major cyanophage types) populations was investigated at three Bay regions over two-year cycles. The genetic diversity of *Synechococcus* and cyanomyovirus was determined by denaturing gradient gel electrophoresis analysis of the host *rbcL* gene and the phage g20 gene fragments respectively. Two primer sets, one for marine *Synechococcus* and one for cyanomyoviruses, were successfully used to amplify the target genes from bacterial and viral samples collected from Chesapeake Bay. Diverse and novel genotypes of *Synechococcus* and cyanomyoviruses were found, reflecting their high genetic complexity in the Chesapeake Bay. Sequencing results showed that these primer sets are suitable to co-monitor the composition of *Synechococcus* and cyanomyoviruses in the Bay. The seasonal variation of *Synechococcus* species composition was stronger than that of cyanomyovirus assemblage in the Bay. No significant correlation on the spatial and temporal variations between the host and phage populations was observed. We speculated that the broad host range of cyanomyoviruses might account for the

non-synchronized patterns of *Synechococcus* and their myoviruses in the estuarine ecosystem.

Introduction

The discovery of high abundance of viruses and high viral infection frequencies of bacteria in aquatic environments (Bergh et al. 1989; Proctor and Fuhrman 1990; Suttle et al. 1990a) has initiated the idea that viruses might influence species diversity and community compositions in nature (Fuhrman and Suttle 1993; Thingstad et al. 1993). Viral infection is a stochastic process (Murray and Jackson 1992) and depends on the abundance of viruses and suitable hosts (Moebus 1996; Wommack and Colwell 2000). The "phages-kill-the-winner" model based on the laboratory experiments revealed a reciprocal relationship between bacterial and co-occurring phage populations (Thingstad and Lignell 1997; Thingstad 2000). In this scenario, the co-existence of competing bacterial species is ensured by the presence of phages that kill the most abundant members of host cells, whereas the difference in substrate affinity between the competing bacterial cells with different growth rate determines the viral types and abundance. This theory has been partially tested in chemostatic bacterial cultures (Lenski 1988; Bohannan and Lenski 2000; Middelboe et al. 2001). A number of field studies appeared to support this hypothesis (Steward et al. 1996; Weinbauer and Hofle 1998b; Wommack et al. 1999b; Hewson et al. 2006). Early viral addition experiments were conducted to test this hypothesis in the natural waters (Suttle 1992; Peduzzi and Weinbauer 1993). However, in such studies, neither the host nor the virus diversity was monitored. In recent studies, efforts have been made to monitor the changes of bacterial

communities (Šimek et al. 2001; Fuhrman and Schwalbach 2003; Hewson and Fuhrman 2006; Bouvier and del Giorgio 2007) or both viral and bacterial communities (Schwalbach et al. 2004; Winter et al. 2004; Hewson and Fuhrman 2006; Weinbauer et al. 2007) upon adding the viral communities to seawater. It was found that virus might have significant influence on bacterial community structure, although such effects were not consistent between water samples. The limitation of these approaches is that only major groups of bacteria or viruses can be detected, therefore it may not provide enough resolution to detect the phage-host interaction occurred at the species or even strain level.

To co-monitor both viral and host groups at the species or strain level, it is necessary to develop gene markers specific for each groups. Finding conserved gene markers for host system is less problematic compared to phage system, mainly because viral evolution follows the mosaic model (Hendrix 1999) and no universal genes like 16S rRNA or 18S rRNA gene can be found among all viruses. Fortunately, several gene markers have been found to be conserved among certain groups of DNA viruses. For example, the DNA polymerase gene among the phyconaviruses (large dsDNA algal viruses)(Chen and Suttle 1996), the viral capsid assemblage gene (g20) among the myoviruses infecting *Synechococcus* (Fuller et al. 1998), and the DNA polymerase gene among the podoviruses infection picocyanobacteria (Chapter 4). Specific PCR primers have been developed and used to study the genetic diversity of these groups of viruses in aquatic environments (Chen and Suttle 1995; Fuller et al. 1998). High genetic diversity of cyanomyoviruses has been found in freshwater, coastal and oceanic waters, based on the g20 gene (Wilson et al. 2000; Zhong et al.

2002; Frederickson et al. 2003; Marston and Sallee 2003; Dorigo et al. 2004; Sandaa and Larsen 2006; Wilhelm et al. 2006)

Synechococcus and their phage are probably the most well studied host-virus system in the marine environment. They are abundant and ubiquitous, and exhibit strong seasonal co-variation in term of their abundance (Waterbury et al. 1979; Suttle 2000). The dynamic interaction between *Synechococcus* and their phages could serve as a model system for testing the "kill the winner" hypothesis in natural aquatic environments. Recently, the population structure of *Synechococcus* and their co-occurring cyanomyoviruses was co-monitored in the Red Sea, based on the *rpo*C1 gene and g20 genes, respectively (Mühling et al. 2005). The abundance and genetic diversity of *Synechococcus* and cyanomyovirus appeared to co-vary in the Red Sea, suggesting that cyanomyovirus could be responsible for shaping the host abundance and diversity in the oligotrophic water (Mühling et al. 2005).

Genetic diversity of *Synechococcus* in aquatic environments has been studied extensively based on several commonly used genetic markers, i.e. 16S rRNA (Giovannoni et al. 1990; Schmidt et al. 1991); *rpo C1* (Palenik 1994; Toledo and Palenik 1997), *rbcL* (Pichard et al. 1997), ITS (Rocap et al. 2002; Ahlgren and Rocap 2006) and etc. However, no PCR primers specific for marine *Synechococcus* were available in any of previous studies. In order to co-monitor both marine *Synechococcus* and their phages in nature, we designed a set of PCR primers specific for marine *Synechococcus* based on the *rbcL* gene sequences (Chapter 2).

In this study, we exploited our newly designed *rbcL* and g20 primers to monitor the population structure of marine *Synechococcus* and cyanomyoviruses,

respectively, over two-year period (from September 2002 to July 2004) in the Chesapeake Bay.

Materials and Methods

Water Samples collection

Water samples were collected as described in Chapter 6. In this study, Stn. 858, Stn. 804 and Stn. 707 were chosen to represent oligohaline, mesohaline and polyhaline regions in the Chesapeake Bay respectively. It was not possible to collect water sample from Stn. 858 in February 2004 cruise (due to the frozen condition), the water sample collected from Stn.845 (ca. 12 nautical miles south to Stn. 858) was used instead. For each station, 250 ml of surface water was filtered immediately upon water collection through 0.22 μm pore-size polycarbonate filters (47 mm diameter, Millipore). The filters were stored at –20°C for later DNA extraction. To concentrate viral assemblages from Chesapeake Bay water samples, 50-liter surface waters from each station were collected and processed as described in Chapter 5. The final viral concentrates (ca. 300 ml) were and stored at 4°C in the dark untill the DNA extraction.

DNA extraction from Chesapeake Bay bacterial and viral assemblages

DNA extraction followed the protocols as described in Chapter 2 and Chapter 4. Environmental DNA samples from September 2002 to July 2004 cruises (n=12) were used for investigating the population successions *Synechococcus* and cyanomyovirus in the Bay. The preliminary trials of DGGE analyses of *rbcL* gene also included bacterial DNA samples collected from Baltimore Inner Harbor in 2002 (Chapter 2).

PCR amplification

The partial *Synechococcus rbcL* gene from environmental samples was PCR-amplified using primer set MSF1 and MSR1 (Chapter 2). The primer MSF1 contained oligonucleotide sequence 5'-GGTCCACTGTGTGGTCCGAGG-3', the primer MSR1 has nucleotide sequence 5'-GTTCTCGTCGTCCTTGGTGAAGTC-3'. This primer set was designed specifically for marine *Synechococcus* spp. and has been successfully used to investigate the genetic diversity of *Synechococcus* in cultured isolates and environmental samples (Chapter 2).

The newly designed PCR primer set SMP-1F (5'-GTAGAATTTTCTACA TTGATGTTG -3') and SMP-2R (5' -TTCATWTCWTCCCAWTCTTC -3') was used to amplify partial cyanomyovirus *g20* gene. This primer set was designed based on the g20 gene sequences from 28 known cyanomyoviruses and 26 representative environmental sequences. The primer set has shown improved specificity for cyanomyoviruses, compared with the CPS1 and CPS8 primer set (Zhong et al. 2002). To obtain the PCR production for DGGE analysis, 40-nt GC-clamp (5'-CGCCCGC CGCGCCCCGCGCCCGGCCCGCCGCCCCCGCCCC-3') was attached to primers MS-1F and SMP-1F in The GC-clamped primers were only used in the re-conditioning PCR step (see below).

All PCR reactions were performed in duplicate tubes with 50-μl volume containing 1 × reaction buffer (Invitrogen, Carlsbad, CA) with 2.0 mM $MgSO_4$, 100 μM dNTPs, 10 pmole of each primer, 1 unit of Platinum HIFI *Taq* DNA polymerase (Invitrogen, Carlsbad, CA) and 50 ng DNA as templates. The PCR program included an initial denaturing at 94°C for 3 min, followed by 30 cycles of 94°C for 1 min,

annealing at 52°C for the g20 gene and 65°C for the *rbcL* gene for 30 sec, and extension at 72°C for 1 min. The final extension of PCR amplification was at 72°C for 10 min and all the PCR reactions were concluded at 4°C. To prepare the PCR samples for DGGE analysis, re-conditioning PCR was performed to minimize the hetero-duplex PCR artifacts (Thompson et al. 2002). The PCR products from the initial duplicate reactions were pooled and 2 μl of each was transferred into 48 μl of fresh PCR mixture containing GC-clamped PCR primers. The PCR was run for addition 15 cycles followed by a 20-min final extension step at 72°C.

DGGE and cluster analysis of banding pattern

DGGE was performed using a Dcode^{TM} Universal Mutation Detection System (Bio-Rad). PCR products (25 μl of each sample) obtained by re-conditioning PCR were separated on a 1.5 mm-thick vertical polyacrylamide (acrylamide : bisacrylamide ratio of 37.5:1) gel with a linear gradient of the denaturants urea and formamide from 50% to 65% with 8% polyacrylamide for *rbcL* amplicons; and from 40% to 70% with 6 % polyacrylamide (preliminary trial were from 40% to 50% with 8% polyacrylamide, see Fig. 7-3) for g20 amplicons. Electrophoresis was performed at 60°C in 0.5 × TAE buffer, and 75 V for 16 h. Nucleic acids were visualized by staining the gel with SYBR Gold for 15 min at room temperature.

DGGE gels were photographed using Kodak EDAS 290 electrophoresis documentation and analysis system (Eastman Kodak Company, New Haven, CT). The DGGE banding patterns (absence and presence of bands) obtained from September 2002 to July 2004 samples were analyzed by using the GelComparII software package (Applied Maths). For each DGGE profile, the dendrogram was

constructed from a binary matrix of similarity values. The pairwise distance matrix was calculated based on the absence or presence of bands, and was analyzed with unweighted pair group mean average (UPGMA) alogrithm and presented as a dendrogram. The DGGE profiles of preliminary trials of *rbcL* and g20 genes were not included in the analysis.

Sequencing and phylogenetic analysis

During the preliminary trials of DGGE separation of PCR-amplified *rbcL* and g20 gene fragments, the representative DGGE bands were excised from gels, re-amplified using non-GC clamped PCR primers. The PCR products were purified using Qiaquick PCR purification kit (Qiagen) before they were sequenced. Purified PCR products were sequenced using MSF1 for *rbcL* gene and SMP1F for g20 gene amplicons respectively. Sequence alignment and phylogenetic reconstruction were conducted using Mac Vector 7.2 program (GCG, Madison, WI). Tajima-Nei distance matrix analysis was used to calculate the distances for the aligned DNA sequences, and neighbor-joining method was used to construct phylogenetic tree. The phylogenetic trees were constructed by using neighbor-joining method based on ca. 350 bp sequence for *rbcL* gene and ca. 390 bp sequence for g20 gene. Bootstrap values for both trees were obtained from analysis of 1000 re-samplings of each data set.

Nonmetric multidimensional scaling (MDS) analysis

MDS was performed based on the distance matrix using SAS System. The relationship between samples based on *Synechococcus* DGGE patterns were illustrated in three-dimension MDS plots, with cell abundance data as the third

dimension (Z-axis). The relationship between samples as for cyanophage DGGE patterns were shown in two-dimension MDS plots. Samples with higher similarity were plotted closer, while samples with the lower similarity were separated further apart.

Results

Seasonal variation in *Synechococcus* abundance

From 12 seasonal samples collected from September 2002 to July 2004, three distinct peaks of *Synechococcus* abundance were observed in all three regions (north to south transect) in the Chesapeake Bay (Fig. 7-1). The maxima of *Synechococcus* abundance occurred in the summer months from Jun to August, with a maximum cell density of 1.9×10^6 cells ml^{-1} in Stn.707 in Jun 2004. On average, higher *Synechococcus* concentrations occurred in 2004 summer (1.3×10^6 cells ml^{-1}, n=6) than in 2003 summer (5.1×10^5 cells ml^{-1}, n=9). The lowest cell density was observed in March 2004 samples with *Synechococcus* abundance of 4.0×10^2 cells. Over three orders of magnitude in variations of *Synechococcus* abundance were evident crossing the north to south transect in the Bay (Fig. 7-1). This is in contrast to the spatial variation in *Synechococcus* cell density, which was within 10-fold difference.

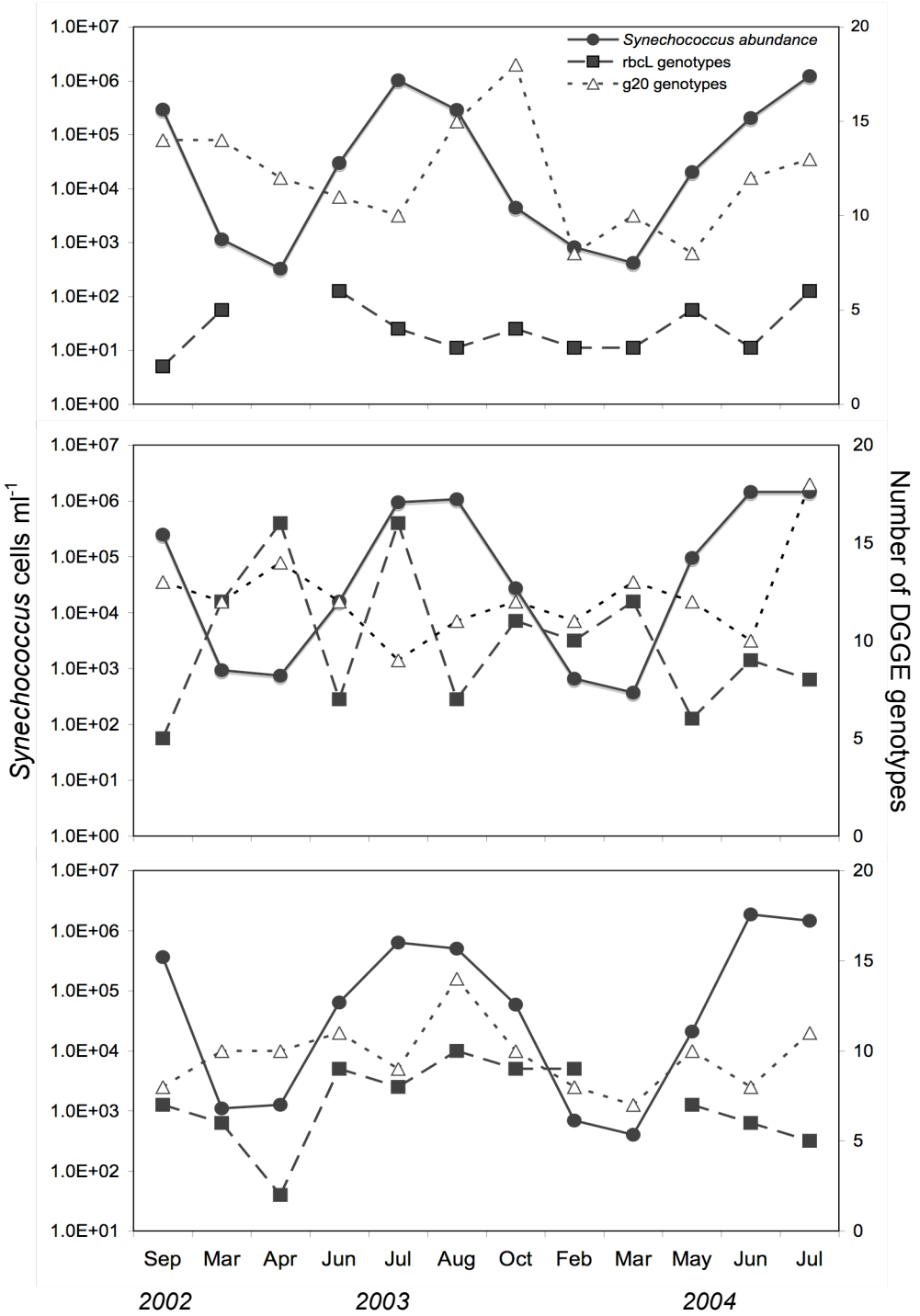


Fig. 7-1. *Synechococcus* abundance, variations in *Synechococcus* genotypes and cyanophage genotypes observed in Stn. 858 (northern Bay), Stn. 804 (middle Bay) and Stn. 707 (southern Bay), from September 2002 to July 2004. The genetic richness of *Synechococcus* and cyanomyovirus were determined from the number of bands in the DGGE gels.

Preliminary test of PCR primer sets

The *rbcL* primers were first tested against seven samples collected on different months from Baltimore Inner Harbor and five samples collected from different stations along the Bay (Fig. 7-2). All the samples yielded the expected PCR amplicons (data not shown). Twelve bands were excised and re-amplified, and eleven of them yielded good sequencing results. Sequence analysis showed that eight of 11 bands appeared to cluster with MC-B *Synechococcus*. Notably, band 1 and 10 (both in MC-B cluster) appeared to persist throughout all seasonal samples in the Baltimore Inner Harbor and northern Bay stations during September 2002. The sequences of band 3, 9 and 11 fell in the MC-A *Synechococcus* group. They were found in January sample in Baltimore Inner Harbor and September samples from middle to southern Bay stations (Fig. 7-2).

Ten viral samples collected from different months between 2002 and 2004 in the middle Bay were tested against the g20 primers (Fig. 7-3). All the 16 excised bands yielded good sequencing results, and twelve of them (75%) fell within the well-defined cyanomyovirus cluster (Fig. 7-3). Intriguingly, these 12 g20 sequences from the Chesapeake Bay appeared to be more closely related to each other compared to other known cyanomyovirus isolates. Among them, band 8, 9 and 10 were observed in all samples (Fig. 7-3). The sequences of band 2, 5, 12, and 13 did not cluster with known cyanomyoviruses and previously discovered Chesapeake Bay environmental clones (Chapter 5). Three of them (band 2, 5 and 12) formed a phylogenetic group distinct from previously recognized environmental clone clusters. It is noteworthy

that these three bands all occurred in winter with strong fluorescent signal on the DGGE gel (Fig. 7-3).

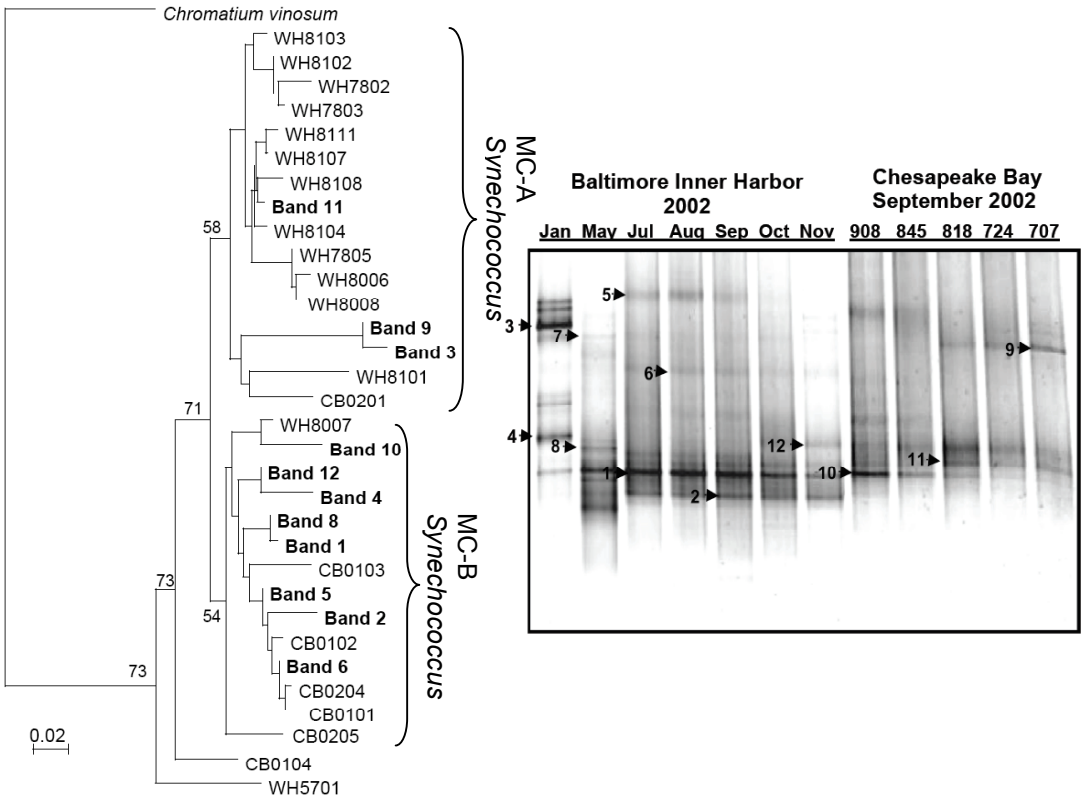


Fig. 7-2. Left: Phylogenetic analysis of *Synechococcus rbcL* gene sequences (ca. 350 bp) recovered from excised DGGE bands. The bootstrap values (>50) were shown on the major nodes. Right: DGGE profile of PCR-amplified *Synechococcus rbcL* gene fragments from Baltimore Inner Harbor samples (2002), and five stations from Chesapeake Bay (September 2002).

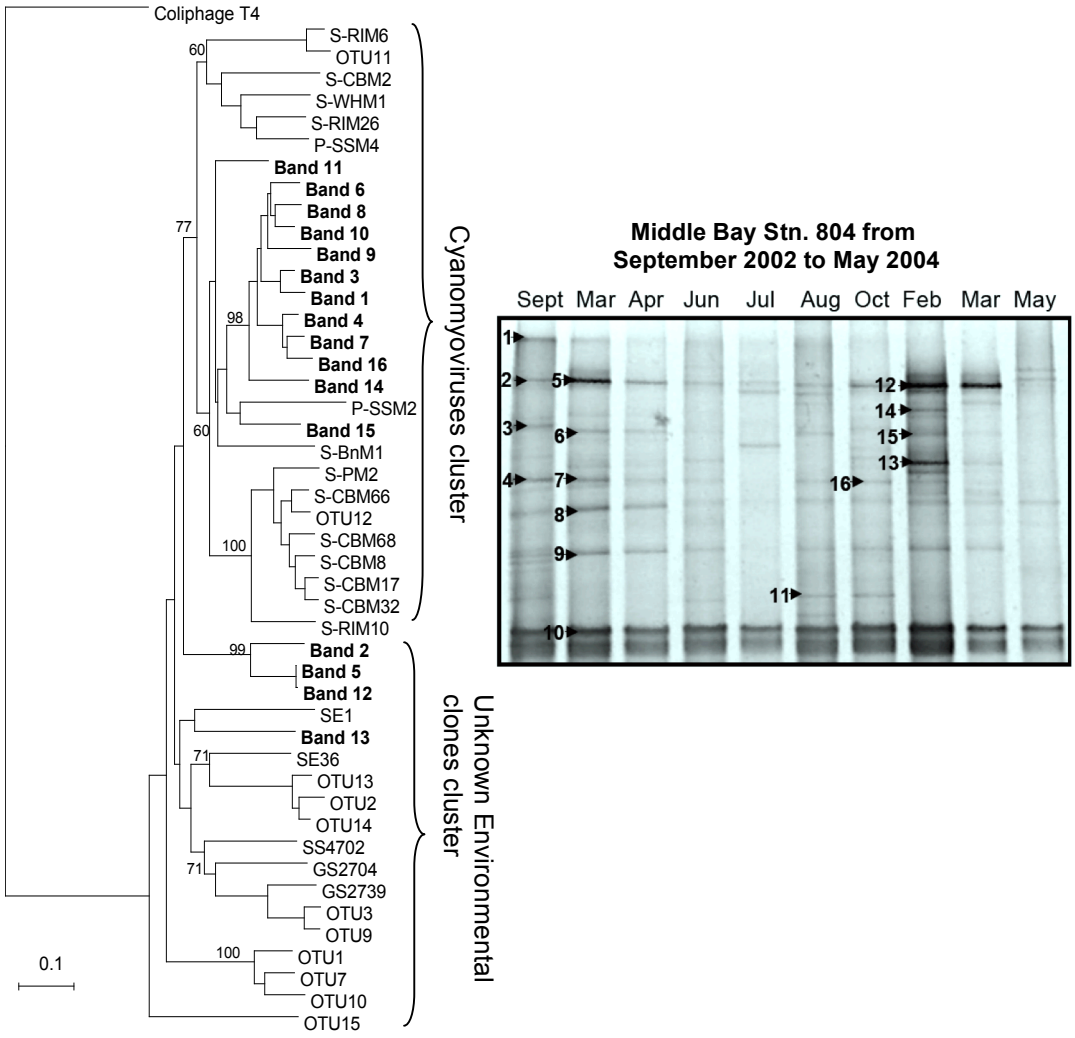


Fig. 7-3. Left: Phylogenetic analysis of cyanomyovirus g20 gene sequences (ca. 390 bp) from excised DGGE bands. The bootstrap values (>50) were shown on the major nodes. Right: DGGE profile of PCR-amplified g20 gene fragments at Stn. 804 in the Chesapeake Bay from September 2002 to May 2004.

Seasonal changes in *Synechococcus* and cyanophage genotypes in the Bay

Among 36 bacterial DNA samples collected from over two-year period, two samples (April 2003 from Stn. 858 and March 2004 from Stn. 707) failed on the *rbcL* gene amplification. In both cases, the *Synechococcus* cell density was below 10^3 cells ml^{-1} (Fig. 7-1). The g20 gene was amplified from all the viral DNA samples.

The complexity of *Synechococcus* population in the Chesapeake Bay varied greatly on both spatial and temporal scales. The observed *Synechococcus* genotypes in the three regions of the Bay ranged from two to 16 among all the samples analyzed (Table 7-1 and Fig. 7-1). The genetic richness of *Synechococcus* in the northern Bay (Stn. 858) was lowest among the three Bay regions, and exhibited the least temporal variation compared to the middle and southern Bay (Table 7-1 and Fig 7-1). In contrast, the middle Bay had the highest but the most variable genetic richness of *Synechococcus* population among the three Bay regions. There was no clear relationship between the abundance and diversity of *Synechococcus* in the Bay. For examples, the genetic diversity of *Synechococcus* was not necessarily highest during the summer blooms.

In contrast, the complexity of cyanomyoviruses in the Bay remained relatively stable over time and space. Between seven and 18 g20 genotypes could be detected in all samples, while the average detectable g20 genotypes varied between 10-12 among the three Bay regions (Table 7-1 and Figure 7-1). There was no clear correlation between genetic richness of cyanomyovirus (g20 genotypes) and *Synechococcus* (*rbcL* genotypes) (Fig. 7-1).

Table 7-1 Genetic richness of *Synechococcus* and cyanomyoviruses in the northern, middle and southern Chesapeake Bay. The number of genotypes was determined from differentiable DGGE bands defined by GelComparII software.

Location	*Synechococcus* genotypes (mean)	Cyanomyovirus genotypes (mean)
Northern Bay (Stn. 858)	2-6 (4)	8-18 (12)
Middle Bay (Stn. 804)	5-16 (10)	8-18 (12)
Southern Bay (Stn. 707)	2-10 (7)	7-14 (10)

Clustering analysis of *Synechococcus* and cyanophage populations

The DGGE analysis showed that the composition of *Synechococcus* population varied dramatically with seasons in all three stations (Fig. 7-4, left panels). The clustering analysis indicated that a high similarity between the populations of *Synechococcus* during warm seasons (from May to October), particularly in the summer months (Fig. 7-4, right panels). During the winter months (February and March), *Synechococcus* populations also shared certain similarity, and the similarity level appeared to increase from the northern to southern Bay (Fig. 7-4, right panels).

In contrast, dramatic seasonal variation was not seen for cyanomyoviruses. The composition of cyanomyovirus assemblages shared more than 52% similarity in all seasons and in the three Bay regions (Fig. 7-5, left panels).

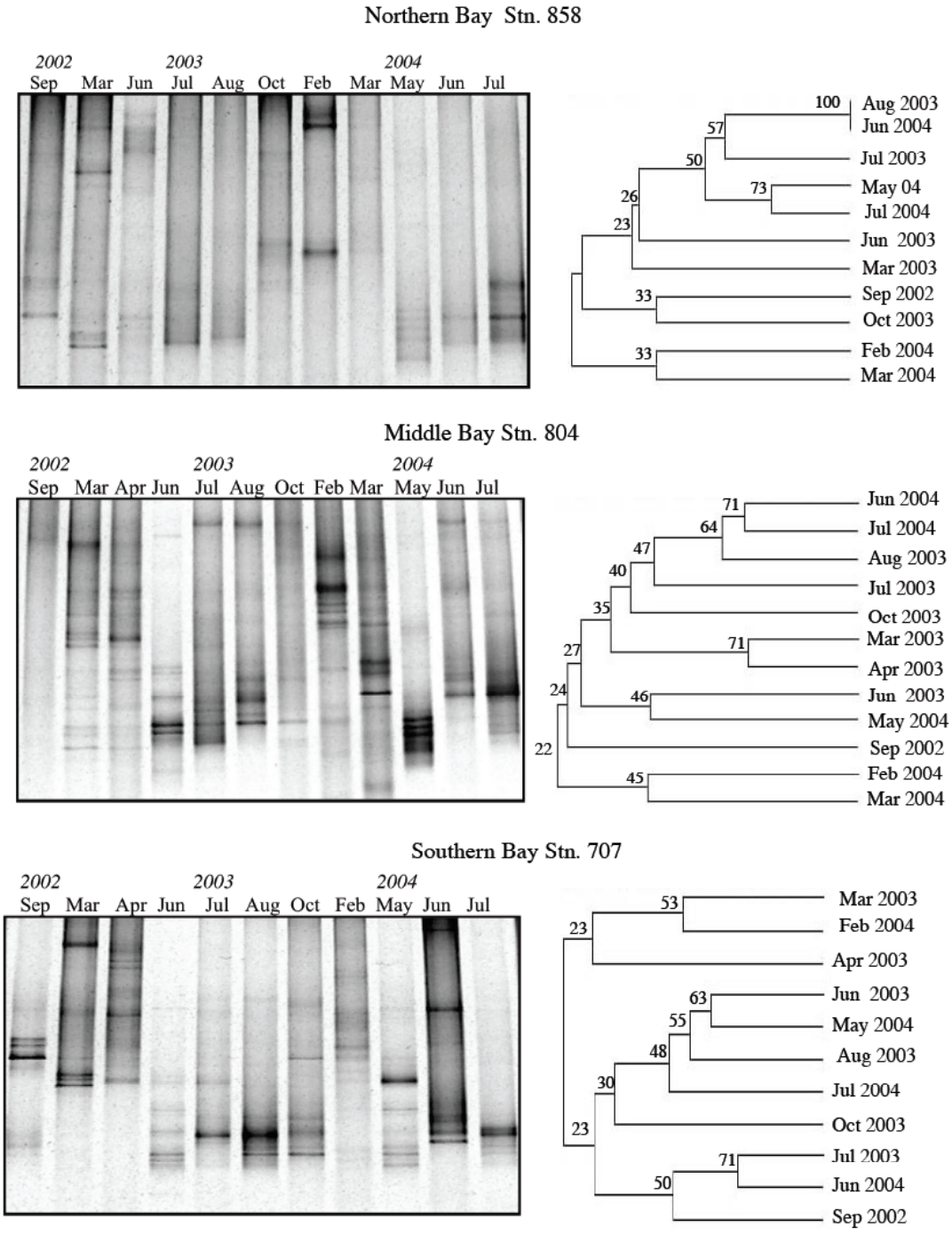


Fig. 7-4. Left panels: *Synechococcus* population variations as revealed by DGGE analysis of PCR-amplified rbcL gene fragments from three stations in the Chesapeake Bay. Right panels: Clustering analysis of corresponding *rbcL*-DGGE profiles. For each DGGE profile, dendrogram was obtained based on similarity matrix of *Synechococcus rbcL* gene DGGE fingerprints from September 2002 to July 2004. Similarities >20% between banding patterns are shown in % at each branch division.

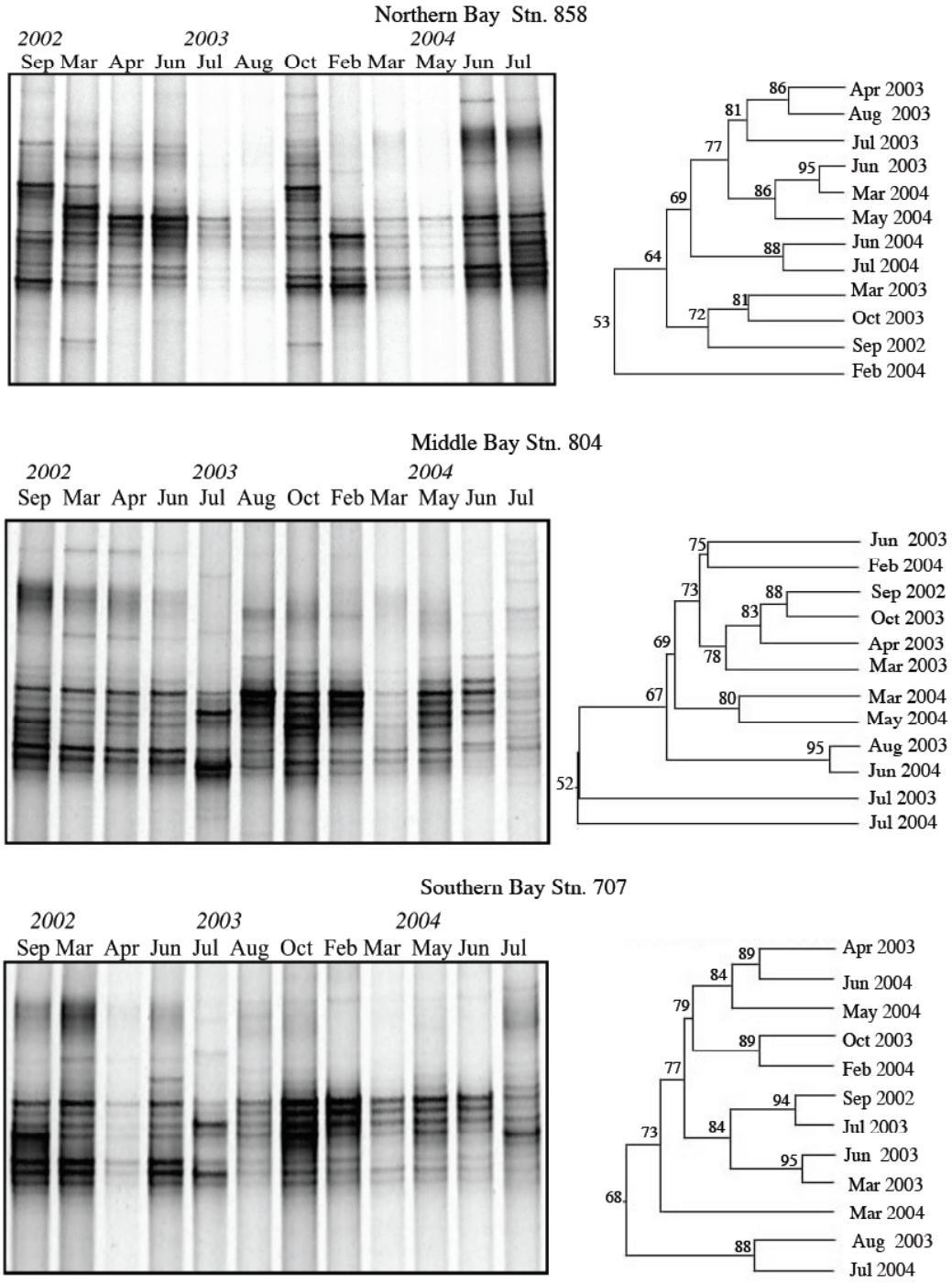


Fig. 7-5. Left panels: cyanomyovirus population variations as revealed by DGGE analysis of PCR-amplified g20 gene fragments from three stations in the Chesapeake Bay. Right panels: Clustering analysis of corresponding g20-DGGE profile. Dendrogram was drawn based on similarity matrix of g20 gene DGGE banding patterns from September 2002 to July 2004. Similarities >50% between banding patterns are shown in % at each branch division.

MDS analysis of *Synechococcus* and cyanophage genotype variations

The MDS analysis was consistent with the clustering analysis. *Synechococcus* population in the summer months was more similar to each other, particularly in the northern and middle Bay regions (Fig. 7-6). However, it was difficult to discern the clear seasonal patterns for cyanomyovirus populations in all three regions in the Chesapeake Bay (Fig. 7-7).

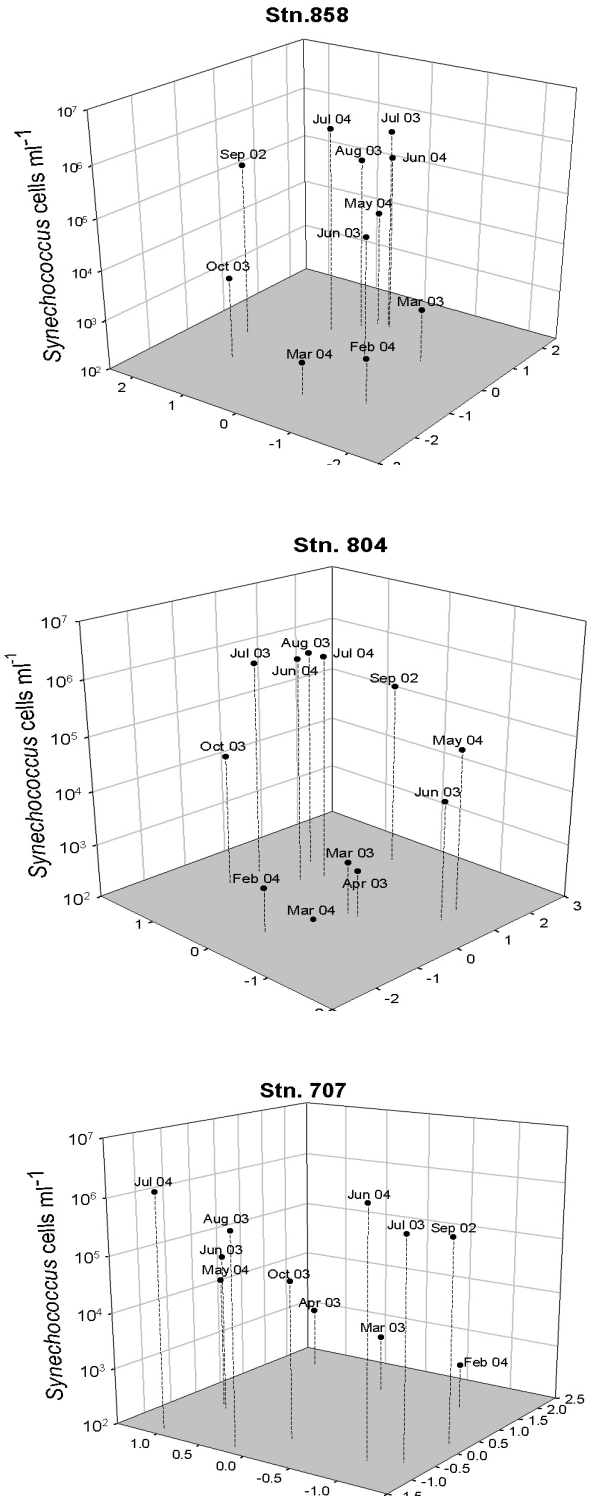


Fig. 7-6. MDS analysis of *Synechococcus rbcL*-DGGE patterns from three Chesapeake Bay stations. The abundance of *Synechococcus* was integrated in Z-axis for each station.

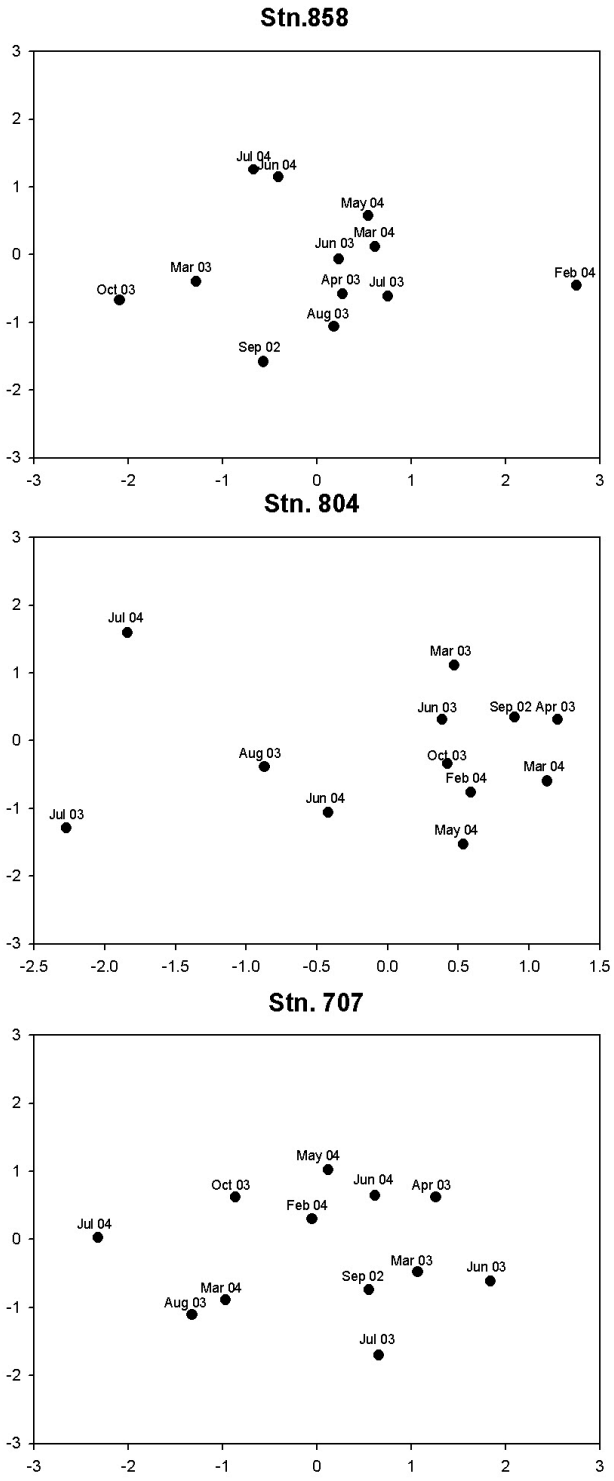


Fig. 7-7. MDS analysis of cyanomyovirus g20-DGGE patterns from three stations in the Chesapeake Bay.

Discussion

Our preliminary trials demonstrated that the spatial and temporal variations of marine *Synechococcus* and cyanomyovirus populations could be co-monitored using the DGGE analysis with specific marker genes. All the sequenced bands from the *rbcL* DGGE gel clustered with marine *Synechococcus*, suggesting the high specificity of the *rbcL* gene primers. Among the 11 bands analyzed, eight bands were closely related to the MC-B *Synechococcus* isolated from Chesapeake Bay (Fig. 7-2). The number of *rbcL* genotypes resolved by DGGE analysis was more than two times higher than that obtained with RFLP analysis (Chapter 2). Therefore, the newly designed marine *Synechococcus rbcL* PCR primers are highly specific and suitable for exploring the genetic diversity of marine *Synechococcus* in the natural environment. All the DGGE bands of the g20 amplicons were related to the viral capsid gene. The new g20 PCR primer set demonstrated improved specificity for cyanomyoviruses compared to the original CPS1/CPS8 primer set (Zhong et al. 2001). Seventy-five percent (12 out of 16 bands) of the g20 sequences recovered from the DGGE gel clustered with known cyanomyovirus isolates (Fig. 7-3), while only 19 % of g20 (40 out of 207 clones) sequences amplified with CPS1/CPS8 primers were affiliated with known cyanomyoviruses (Zhong et al. 2002). In addition, PCR-DGGE analysis of g20 genes revealed at least 18 genotypes in the Bay, which are higher than the genotypes estimated based on the T-RFLP method we developed earlier (Chapter 5). Therefore, DGGE analysis appears to be a sensitive tool to investigate the natural diversity of cyanomyoviruses in aquatic environments. Because the g20 gene primers only target cyanomyoviruses, a subset of cyanophages, the results obtained in this

study could not be extended to other cyanophage groups such as cyanopodoviruses and cyanosiphoviruses. In addition, one should be careful when interpreting the g20 genotypes. Not all the g20 genotypes detected by DGGE are myoviruses infecting *Synechococcus*.

Genetic diversity of *Synechococcus* and cyanophages

As revealed by DGGE analysis, as least 16 different *Synechococcus* genotypes and 18 distinct cyanomyovirus genotypes could be found in the same water samples, indicating the presence of diverse *Synechococcus* and cyanophages in the Bay. The genetic diversity of cyanomyovirus the Chesapeake Bay is comparable to that reported in Red Sea (Mühling et al. 2005) and British Columbia coastal water (Frederickson et al. 2003); higher than that found in the Atlantic Ocean (Wilson et al. 2000) but lower compared to Lake Bourget (Dorigo et al. 2004). Intriguingly, many novel genotypes of *Synechococcus* and cyanomyovirus were found in the Bay as revealed by sequence analysis, as they were not closely related to any of the cultured representatives (Fig. 7-2 and Fig. 7-3). Some of these genotypes could be persistent and dominant components in the Chesapeake Bay, indicating that an unexplored novel *Synechococcus*-phage system could exist in the Bay. This observation extended the findings from the culture-dependent study of *Synechococcus* (Chapter 2) and their phages (Chapter 4).

Seasonal succession of *Synechococcus* and cyanomyovirus populations

The concurrent seasonal succession was not found between *Synechococcus* and their cyanomyoviruses in the Chesapeake Bay estuary. Less variation in g20 genetic richness found in the Bay indicates that the cyanomyovirus populations are relatively

stable regardless the change of host population composition. This result is different from what has been found in Red Sea (Mühling et al. 2005), and also different from our previous observations in Baltimore Inner Harbor (Chapter 5), indicating the impact of cyanomyovirus on host *Synechococcus* populations is different from that observed in the oligotrophic open oceans. We tend to believe that such a discrepancy is mainly due to the difference in primer specificity. Mühling et al. (2005) used the CPS4/ CPS5 primer set (Wilson et al. 1999) to amplify a ca. 118-bp g20 fragment in Red Sea. This primer set was designed based on the conserved regions of g20 gene and only four known myovirus sequences (one from T4 and three from cyanomyoviruses) were available at the time (Fuller et al. 1998). The specificity of this primer set has never been tested for natural viral assemblages, as such short sequences could not provide enough information to differentiate cyanomyovirus and other phage genotypes. Our early study applied the CPS1/CPS8 primer set to investigate annual change in cyanomyovirus diversity in Baltimore Inner Harbor (Chapter 5). However, as mentioned above, the recovery of true cyanomyovirus sequences with CPS1/CPS8 primers (19%) is much lower than our newly designed primers (75%). The specificity of the CPS1/CPS8 primer set has been questioned in recent studies, as many g20 sequences obtained by using this primer set may not originate from cyanomyoviruses (Short and Suttle 2005; Wilhelm et al. 2006). To co-monitor both host and phage populations, it is critical to have PCR primers that will target both groups accurately. The new g20 primers used in this study referred much more known cyanomyovirus sequences (see method). Therefore, it is plausible that the new primer set provides more accurate analysis of cyanomyovirus population

composition. Poor specificity of PCR primers used in previous studies may detect many myoviruses which infect other bacteria, and therefore mask the true signals of cyanomyovirus. In this case, the seasonal succession of cyanomyoviruses could not be interpreted precisely. Further studies are needed to evaluate the data resulted from the different primer sets.

The observed poor synchronicity between *Synechococcus* and cyanomyoviruses are more likely due to the polyvalence of cyanomyoviruses. "Kill-the-winner" theory (Thingstad and Lignell 1997) assumed that viruses specifically infect the certain hosts. This assumption is generally true at the species or genus level in most cases (Ackermann and DuBow 1987a), although some viruses have a broad range of hosts (Riemann and Middelboe 2002). As for the three known cyanophage types that infect *Synechococcus*, cyanomyovirus have been known to have a broad host range (Waterbury and Valois 1993; Suttle 2000; Lu et al. 2001). Some cyanomyoviruses are not only capable of cross-infecting between MC-A and MC-B *Synechococcus* (Chapter 4), but also between *Synechococcus* and *Prochlorococcus* (Sullivan et al. 2003). Because different types of cyanomyoviruses could potentially infect multiple *Synechococcus* spp., the changes of host population composition might pose little impact on cyanomyovirus population succession and vice versa. As such, the "kill-the-winner" model would be more suitable for the phages with a narrow host range, but not applicable for broad-host-range cyanomyoviruses. Abundance and phage types are important factor to assess the impact of cyanophages on host population structures. It was found that during summer *Synechococcus* blooms in 2004, host-specific podoviruses and siphoviruses rather than myoviruses

appeared to be the most abundant types in the Bay (Chapter 6). Such an observation favors the speculation that cyanomyovirus might not be the major player for regulating the host *Synechococcus* populations in the Chesapeake Bay at that time. Otherwise, a co-variation or concurrent patterns between *Synechococcus* and cyanomyovirus populations should be expected.

In summary, despite the strong seasonal variation in *Synechococcus* population structure, the cyanomyovirus composition exhibited few changes on temporal and spatial scales. The polyvalence of cyanomyoviruses may be responsible for the observed stable population structures over time and space. In order to gain a better understanding of the impact of cyanophages on host populations, future investigation of other cyanophage groups (podovirus and siphovirus) are necessary.

Chapter 8: Summary

My dissertation was devoted to study *Synechococcus* and their viruses in the Chesapeake Bay. In the past six years, I have taken a great deal of effort to isolate and characterize *Synechococcus* and cyanophages living in the Chesapeake Bay. The Microbial Observatory for Virioplankton Ecology (MOVE) project in the Chesapeake Bay (from September 2002 to February 2007) has provided unique opportunity to investigate the abundance, composition, distribution and interactions of *Synechococcus* and their viruses in the largest estuarine ecosystem in USA. My dissertation research unveils unique features of *Synechococcus* and their viruses living in a large estuarine ecosystem. Meanwhile, many new interesting questions have arisen from my study and deserved further investigation.

Major findings

I. Unique group of *Synechococcus* (MC-B) was found in the Chesapeake Bay estuary.

Based on the phylogenetic analysis of *rbcL* and ITS gene sequences, vast majority of Chesapeake Bay *Synechococcus* isolates were classified into marine cluster B (MC-B) *Synechococcus* (Chapter 2 and Chapter 3). Prior to my work, the MC-B cluster only contained one strain (WH8007). My work confirmed the presence of diverse *Synechococcus* in the MC-B group. The finding of diverse and abundant MC-B type *Synechococcus* contributes significantly to the taxonomy of marine *Synechococcus*. Clone library analysis of ITS gene sequences recovered from the Chesapeake Bay water samples showed that *Synechococcus* are diverse in the Bay and distinct genotypes exist in different regions of the Bay (Chapter 3). This finding

supports the notion of "microdiversity" found in picocyanobacteria. Both phycocyanin- and phycoerythrin-rich *Synechococcus* are present in the MC-B cluster, suggesting containing only phycocyanin is no longer a stable taxonomic feature for MC-B *Synechococcus.* Chesapeake Bay *Synechococcus* strains can grow in a much broader range of salinity. This capability probably grants the MC-B *Synechococcus* competition advantages over MC-A strains in the Chesapeake Bay estuary where a remarkable salinity gradient is present. These findings also imply that cyanobacteria may develop versatile capability to adapt to various habitats, and become successful microorganisms on the planet Earth.

II. Cyanophages infecting MC-B *Synechococcus* spp.

Many cyanophages infecting MC-A or oceanic *Synechococcus* have been isolated and characterized. However, phages infecting MC-B or estuarine *Synechococcus* remained largely unexplored before my effort. In this study, seven phages isolated from four different MC-B strains were characterized (Chapter 4). Highly host specific podoviruses and siphoviruses are commonly found for MC-B *Synechococcus*. This is in contrast to the prevalence of myoviruses for MC-A or oceanic *Synechococcus.* Resistance to polyvalent myovirus infection appeared to be a common feature for MC-B *Synechococcus* strains. This finding suggests that genetically diverse picocyanobacteria are susceptible to different types of virus. Prior to my work, little is known about the latent period and burst size for cyanophages. I found that podoviruses infecting *Synechococcus* have significant short latent period compared with myoviruses and siphoviruses. Cyanopodovirus appear to be a super-virulent phage type among cyanophages infecting *Synechococcus*. The burst size of

cyanophages appears to be related to their genome size. Cyanophages appear to utilize all available host DNA content to maximize their progeny phage production. All known cyanopodoviruses contain the DNA polymerase gene. Phylogenetic analysis showed that *pol* gene was highly conserved among cyanopodoviruses infecting marine *Prochlorococcus*, MC-A and MC-B *Synechococcus*. Unique sequences of photosynthetic gene (*psbA*) were identified in two MC-B *Synechococcus* cyanopodoviruses. They cluster with some mysterious environmental *psbA* sequences, which were speculated to come from cyanopodoviruses. My cultivation effort suggested that those unidentified *psbA* gene sequences found in natural seawaters are very likely from cyanopodoviruses.

III. *Synechococcus* and their phages are abundant, ubiquitous and important microbial components in the Chesapeake Bay.

The spatial and temporal distributions of *Synechococcus* abundance and cyanophage titers were investigated bi-monthly over four consecutive years (from September 2002 to February 2007). This is the first detailed ecological survey of picocyanobacterial and their viruses in a large estuarine ecosystem. Both *Synechococcus* and their phage abundance exhibited strong seasonal patterns, and were highly correlated with water temperature (Chapter 6). Inter-annual variation of the host and viral abundance was also evident, and appeared to be influenced by seasonal changes. Massive occurrence of *Synechococcus* or "*Synechococcus* bloom" was observed throughout the Bay in every summer. *Synechococcus* abundance often exceeded 10^6 ml^{-1}, and account for nearly one third of total bacterial counts in summer. *Synechococcus* can contribute more than 50% of total phytoplankton

chlorophyll *a* and primary production in the Bay during summertime. These findings indicate that *Synechococcus* are major players in carbon fixation, energy flow and nutrient cycling during summertime in the Chesapeake Bay ecosystem. My study also provided the evidence that the pigment composition of cyanobacteria (PE vs. PC type) varies dramatically from the upper to lower bay, likely due to the chromatic adaptation. This finding suggests that distribution of different type of *Synechococcus* in the Chesapeake Bay is subjected to environmental impacts and predictable with time and space. Infectious cyanophage titers were found highly abundant ($> 5 \times 10^5$ MPN ml^{-1}) and co-varied strongly with *Synechococcus* abundance over time and space (Chapter 6), suggesting that active infection and lyses take place between *Synechococcus* and their viruses. Infectious cyanophage titers can exceed their host abundance and may cause significant portion of *Synechococcus* mortality during summertime. This finding further supports the idea that viruses can play important roles in controlling host biomass. Released nutrients and trace elements from *Synechococcus* by viral lysis can be subsequently utilized by phytoplankton and bacteria and therefore enhance the energy flow and nutrient cycling in the Chesapeake Bay ecosystem.

IV. The "killing the winner" phenomenon was not seen between *Synechococcus* and their myoviruses.

A great deal of my research effort was undertaken to develop specific gene makers that are suitable to detect both *Synechococcus* and cyanophage populations in nature, with the goal to test the "killing the winner" hypothesis (Chapter 2-5, and 7). I have successfully developed a set of PCR primers specific for monitoring the genetic

diversity of marine *Synechococcus*. I also significantly improved the PCR primer specificity for cyanomyoviruses. DGGE analyses with these two primer sets were used to monitor the changes of *Synechococcus* and cyanomyovirus populations in different seasons and locations. The composition of *Synechococcus* species changed dramatically in different seasons, but the composition of cyanomyoviruses was relatively stable. If the "killing the winner" mechanism takes place between *Synechococcus* and cyanomyoviruses, one should be able to see a dramatic change on cyanomyovirus composition at the same time. However, no obvious coupling was found between the population dynamics of *Synechococcus* and cyanomyoviruses, suggesting that the "killing the winner" mechanism may not apply to the interaction between *Synechococcus* and cyanomyovirus. The poor co-variation between *Synechococcus* and cyanomyovirus population succession could be related to polyvalence of cyanomyoviruses. It would be very interesting to see how *Synechococcus* interact with strain-specific cyanopodoviruses, and this can be done by monitoring changes of cyanopodovirus population using the DNA *pol* gene as a marker molecule in the near future.

Significance

My dissertation focuses on the ecological interaction between cyanobacteria and cyanophage in the estuarine ecosystem. I identified a new group of cyanobacteria living the Chesapeake Bay estuary, and demonstrated their ecological adaptation to this unique ecosystem. Some of my findings challenged the traditional concepts on cyanobacterial taxonomy. Phenotypic features are no long reliable characters for

categorizing cyanobacteria. Diversity of cyanobacteria in nature would be much underappreciated without molecular characterization. It is possible that Chesapeake Bay cyanobacteria are more diverse than cyanobacteria living in the ocean. Compared with open ocean system, picocyanobacteria in the estuarine ecosystem are largely unexplored yet. Much more works should be encouraged to explore the diversity of cyanobacteria in various estuaries. My study also provided the evidence that the pigment composition of cyanobacteria (PE vs. PC type) varies dramatically from the upper to lower bay, likely due to the chromatic adaptation. The differential distribution of PE vs. PC type cyanobacteria can reflect the light quality in natural aquatic environment. Picocyanobacterial abundance in the Bay appears to increase in the past decade, and this can be related to eutrophication and increased surface water temperature. Picocyanobacteria appear to be very sensitive to certain environmental parameters like temperature, light and nutrients, and can serve as a biomarker for the changing environments such as global climate warming. The linkages between cyanobacteria and changing environmental conditions can also provide useful information for effective environmental managements.

During summer months, there are estimated ca. 2×10^{21} *Synechococcus* cells in the Chesapeake Bay and they can fix up to 5×10^{9} g carbon in the Bay. Picocyanobacteria can be important nutrient and energy source as well as essential microbial components in the estuarine ecosystem.

The most intriguing finding in my study is the highly specific host-phage relationship between cyanobacteria and cyanophage in the Chesapeake Bay. This is in distinct contrast to what has been learned in the open ocean. It certainly adds a new

chapter to the ecology of cyanophage. At this point, it is still not clear why estuarine (and perhaps freshwater) cyanobacteria favor more specific host-virus relationship, while polyvalence is more common for oceanic cyanobacteria. We still do not know if such a distinct virus-host strategy between oceanic and estuarine waters is the consequence of virus-host co-evolution or ecological adaptation of cyanobacteria. My finding partially emphasizes the difficulty of using viruses as biological agent to control the natural cyanobacterial blooms due to the complex interactions between cyanophages and their host. This idea was proposed many decades ago, however, no achievement has been made yet (Suttle 2000).

My work suggests that virus-host interactions can be very different in a eutrophic estuarine ecosystem than in oligotrophic oceans. Cyanobacterial and their virus abundance decreases from eutrophic estuary to oligotrophic oceans. Therefore, the contact frequency between host and virus also decreases. Higher contact frequency may increase the potential for viral infection and consequently result in increased selective pressure for cyanobacteria to become resistant to viral infection. If so, impacts of viral infection on cyanobacterial community are expected to be different depending on the composition of hosts. As such, viral effects on microbial community may be more complex than what we have appreciated and it cannot be generalized to apply for different ecosystems.

My dissertation has laid the groundwork on the ecology, biology and evolution of cyanobacteria and cyanophage in the Chesapeake Bay. Novel culture systems for estuarine *Synechococcus* and their phages have been established and maintained in the Center of Marine Biotechnology, and will be useful to other

research communities. The knowledge obtained from Chesapeake Bay *Synechococcus*-phage system has provided and will continue to provide new insights into the phage-host interaction in the estuarine ecosystem. Moreover, the large data inventory generated via this long-term survey will be soon available on the MOVE project website (http://www.virusecology.org/MOVE/Home.html), which will continue to benefit the future studies in the Chesapeake Bay.

Future perspectives

If I had a chance to continue working on this project, I would like to direct my research as follows:

Genome sequence of a few MC-B Synechococcus strains. Although many *Synechococcus* genomes have been sequenced, the vast majority focuses on oceanic strains. No genome representing MC-B or estuarine *Synechococcus* has been sequenced. Recently, the genome sequences of coastal and oceanic *Synechococcus* strains have provided new insights into the niche adaptation of marine *Synechococcus*. Further studies are necessary to address the question why MC-B *Synechococcus* are thriving in the estuarine ecosystem but not in open oceans. I believe that genome sequence of a MC-B strain will deepen our understanding on biology and ecology of *Synechococcus* living in the estuarine environment.

Genome sequence of cyanosiphoviruses. Currently, several genomes of cyanomyoviruses and cyanopodoviruses have been sequenced, but no sequence of cyanosiphoviruses is available. Siphoviruses are often associated with lysogenic infection. The presence of lysogeny in *Synechococcus* is still a debated issue. Genome

sequence of MC-B siphoviruses may provide additional information related to *Synechococcus* lysogeny. It will also allow us to explore if the *psbA* gene is present in cyanosiphoviruses.

Genetic diversity of cyanopodoviruses in the marine environments. All the known cyanopodoviruses contain the DNA polymerase gene. I have designed PCR primers that has been used successfully to amplify the *pol* gene for cyanopodoviruses. This primer set will soon be used to explore the genetic diversity of cyanopodovirus, and co-variation between cyanopodoviruses and *Synechococcus* in the Chesapeake Bay.

Genetic diversity of psbA gene in Chesapeake Bay viral communities. I have designed a set of PCR primers that specifically target the *psbA* gene for cyanobacteria and cyanophages. This primer set has been tested in various labs and proved to be the most suitable primer set to explore the genetic diversity of *psbA* gene in marine environments (Wang and Chen, unpublished data; Beja, personal communication). Preliminary DGGE test showed promising results. Chesapeake Bay appears to contain many *psbA* sequences different from those in the open ocean. Further investigation should be undertaken to understand the diversity of *psbA* gene in the Bay.

Appendix A

Media

I. SN Medium for growing MC-B *Synechococcus* spp. (modified from Waterbury et al. 1986)

1. Prepare 1 liter of autoclaved seawater ultrafiltrate (virus-free) in Pyrex bottle; adjust the salinity to 15 ppt by adding DI H_2O. After cooling, aseptically add the following:

Quantity	Compound	Stock Solution
2.5 mL	$NaNO_3$	300 g/L dH_2O
2.6 mL	K_2HPO_4 (anhydrous)	6.1 g/L dH_2O
5.6 mL	$Na_2EDTA \cdot 2H_2O$	1 g/L dH_2O
2.6 mL	Na_2CO_3	4 g/L dH_2O
1 mL	H_2SeO_3	1.29 mg/L dH_2O
1 mL	Cyano trace metal solution	(see recipe below)

2. Cyano trace metal solution recipe:

Quantity	Compound
6.25 g	Citric Acid $\cdot H_2O$
6 g	Ferric ammonium citrate
1.4 g	$MnCl_2 \cdot 4H_2O$
0.39 g	$Na_2MoO_4 \cdot 2H_2O$
0.025 g	$Co(NO_3)_2 \cdot 6H_2O$
0.222 g	$ZnSO_4 \cdot 7H_2O$

Dissolve each metal compound individually in 100 ml DI H_2O. Combine the six solutions and mix; bring volume up to 1 liter with DI H_2O. Filter-sterilize and store in refridge. Discard the solution if precipitate occurs.

II. SM medium for storage of cyanophages

Mix the stock solutions of NaCl (5 M), Tris HCl (1 M, pH=7.4-7.6), $MgSO_4$ (1 M) and 2% gelatin (2 g of gelatin powder in 100 ml ddH_2O, autoclaved) to the final concentration of :

NaCl	100 mM
Tris HCl	50 mM
$MgSO_4$	10 mM
Gelatin	0.1%

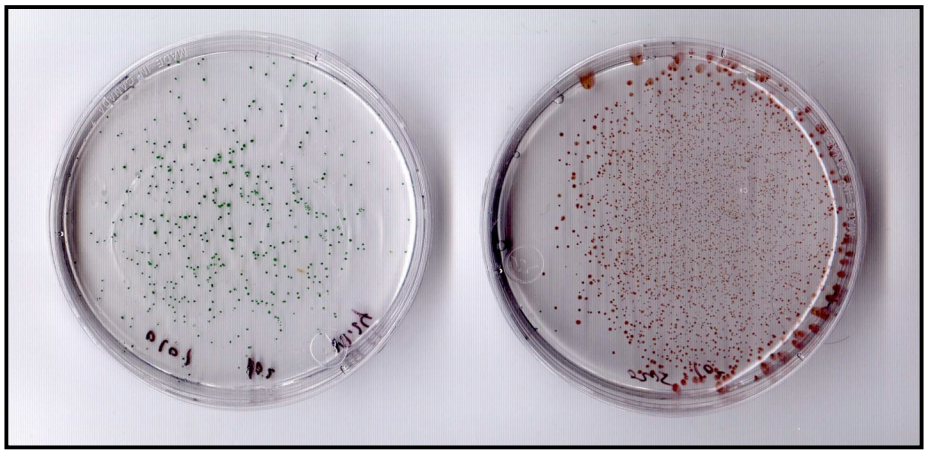
Autoclave and store in refridge.

Appendix B

Protocols for isolation and purification of Chesapeake Bay *Synechococcus* and their phages

I. Isolation of *Synechococcus* spp.

1. Enrich Chesapeake Bay water samples (either unfiltered or 3-μm pre-filtered) with SN nutrients in 50 ml culture flasks and under illumination ca. 20 μE m^{-2} s^{-1} at room temperature (ca. 23-28°C) for two weeks. Skip this procedure if *Synechococcus* abundance in water samples is high (> 10^4 cells ml^{-1}).
2. Inoculate 10 to 100 μl of enrichments or natural water samples with 5 ml of 0.5 % LMP agarose (Invitrogen) melted in SN medium (15 ppt) and held at 35°C.
3. Pour the mixtures evenly onto 1.0 % solid agar plates prepared using SN nutrient and LMP agarose.
4. Incubate plates under ca. 5-10 μE m^{-2} s^{-1} illumination for 24 h and then transfer the plates to an incubator with constant illumination of 20-30 μE m^{-2} s^{-1} at 26°C.
5. Monitor plates daily for up to one month. Pick up green or reddish colonies using Pasteur pipettes and then resuspend the colonies in culture vessels containing 1 ml of SN medium.
6. Repeat procedures 2-5 at least twice to obtain the purified clonal *Synechococcus* isolates.



CB 0101 CB 0205

Fig. B-1. Plate demonstration of top agar overlay method to isolate Chesapeake Bay *Synechococcus* strains.

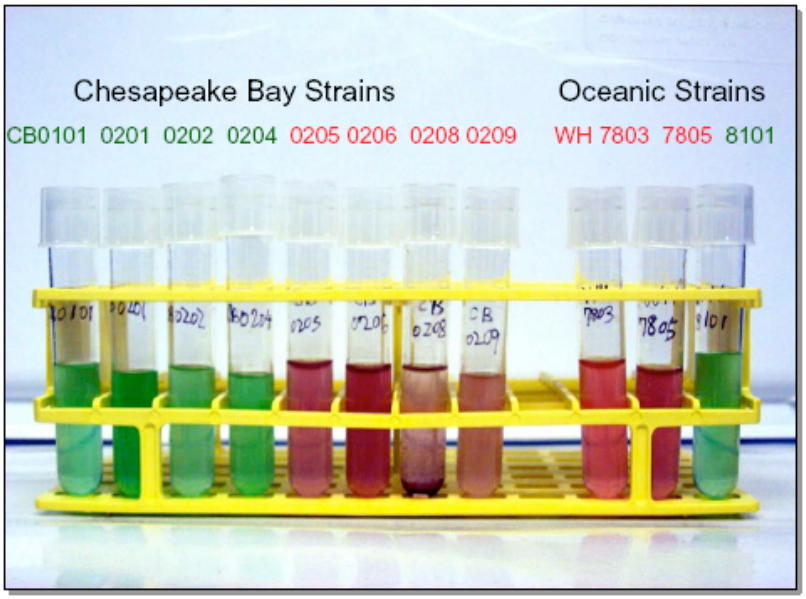


Fig. B-2. Cultures of Chesapeake Bay *Synechococcus* strains.

II. Isolation of cyanophages

1. Inoculate 100 μl of serial dilutions (with SM medium) of viral concentrates (or 0.22-μm pre-filtered seawater samples) to 1 ml of concentrated exponentially growing *Synechococcus* cells (ca. 10^9 cells) and mix well.
2. Add host-virus mixture into 5 ml of 0.5% LMP agarose (Invitrogen) melted in SN medium and held at 35°C.
3. Vortex the suspensions thoroughly and pour evenly onto 0.6% solid SN medium agar plates. Incubate the plate under ca. 5 μE m^{-2} s^{-1} illumination for 24 h and then transfer the plates to an incubator with a constant illumination of ca. 20 μE m^{-2} s^{-1} at 26°C.
4. Monitor plates daily for up to three weeks. When the plaques are observed, pick up the plaques of interest by using a Pasteur pipette and then resuspend them in 2 ml of SM medium. The suspensions should be stored at 4°C in the dark.
5. Repeat the above procedures at least twice to obtain the purified clonal cyanophage isolates.

CB0101 CB0208

Fig. B-3. Plate demonstration of top agar overlay method to isolate viruses infecting Chesapeake Bay *Synechococcus* strains.

III. Cyanophage amplification and purification

1. Inoculate clonal cyanophage lysate (obtained by plaque assay) to 1 liter of exponentially growing host culture.
2. Upon clearance of the host cells (1 to 6 days), add 10 ml of chloroform to the lysates. Remove the cell debris in the lysates by centrifugation at 10,000 rpm in a Backman J2-21 centrifuge. The supernatants can be stored at 4°C till further use.
3. Add RNase A (2 μg ml^{-1} final conc.) and DNase I (2 μg ml^{-1} final conc.) into lysates to digest released host DNA for 1 h at room temperature.

4. Add NaCl to a final concentration of 1M in the lysates and incubate the lysates on ice for at least 30 mins.
5. Centrifuge lysates 10,000 rpm in a Beckman JA14 rotor for 30 mins.
6. Pool the supernatants and filter through 0.45-μm membrane filter (Type HA, Millipore) and further concentrate the lysates to 60 ml using a 50 kDa cut-off Minimate TFF ultrafiltration system (Pall Corp. Ann Arbor MI.).
7. Added polyethlene glycol (PEG 8,000, 100 g L^{-1} final conc.) and incubate lysates overnight at 4°C.
8. Centrifuge the phage particles at 15,000 rpm in a Beckman JA-21 rotor for 1 h at 4°C.
9. Resuspend the viral pellet with 6 ml SM medium and incubate overnight at 4°C.
10. Add CsCl to phage suspension to a final concentration of 0.5 g ml^{-1} and ultracentrifuge at 60,000 rpm (200,000 g) using a T-8100 rotor (Sorvall Discovery 100S centrifuge) for 24 h.
11. Extract the visible virus band with a needle syringe (gauge 22) and dialyze in Slide-A-Lyzer 30K MWCO dialysis cassettes (Pierce) twice in SM medium overnight at 4°C.
12. CsCl purified cyanophage suspensions can be stored at 4°C until further uses.

Bibliography

Abedon ST, Hyman P, Thomas C (2003) Experimental examination of bacteriophage latent-period evolution as a response to bacterial availability. Appl Environ Microbiol 69: 7499-7506.

Ackermann HW (2007) 5500 Phages examined in the electron microscope. Arch Virol 152: 227-243.

Ackermann HW, DuBow MS (1987a) Viruses of prokaryotes: General properties of bacteriophages. CRC Press, Inc, Boca Raton, Florida.

Ackermann HW, DuBow MS (1987b) Viruses of prokaryotes: Natural groups of bacteriophages. CRC Press, Inc, Boca Raton, Florida.

Affronti LF, Jr , Marshall HG (1993) Diel abundance and productivity patterns of autotrophic picoplankton in the lower Chesapeake Bay. J Plankton Res 15: 1-8.

Affronti LF, Jr, Marshall HG (1994) Using frequency of dividing cells in estimating autotrophic picoplankton growth and productivity in the Chesapeake Bay. Hydrobiologia 284: 193-203.

Agawin NSR, Duarte CM, Agustí S (1998) Growth and abundance of *Synechococcus* sp. in a Mediterranean Bay: seasonality and relationship with temperature. Mar Ecol Prog Ser 170: 45-53.

Ahlgren NA, Rocap G (2006) Culture isolation and culture-independent clone libraries reveal new marine *Synechococcus* ecotypes with distinctive light and N physiologies. Appl Environ Microbiol 72: 7193-7204.

Alfreider A, Vogt C, Hoffmann D, Babel W (2003) Diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from groundwater and aquifer microorganisms. Microb Ecol 45: 317-328.

Avaniss-Aghajani E, Jones K, Chapman D, Brunk C (1994) A molecular technique for identification of bacteria using small subunit ribosomal RNA sequences. Biotechniques 17: 144-149.

Badger MR, Hanson D, Price GD (2002) Evolution and diversity of CO_2 concentrating mechanisms in cyanobacteria. Funct Plant Biol 29: 161-173.

Becker S, Singh AK, Postius C, Böger P, Ernst A (2004) Genetic diversity and distribution of periphytic *Synechococcus* spp. in biofilms and picoplankton of Lake Constance. FEMS Microbiol Ecol 49: 181-190.

Becker SM, Fahrbach P, Ernst A (2002) Quantitative tracing, by Taq nuclease assays, of a *Synechococcus* ecotype in a highly diversified natural population. Appl Environ Microbiol 68: 4486-4494.

Bench SR, Hanson TE, Williamson KE, Ghosh D, Radosovich M et al. (2007) Metagenomic characterization of Chesapeake Bay virioplankton. Appl Environ Microbiol: AEM.00938-00907 (in press).

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

Bohannan BJM, Lenski RE (2000) Linking genetic change to community evolution: insights from studies of bacteria and bacteriophages. Ecol Lett 3: 362-377.

Bouvier TC, del Giorgio PA (2002) Compositional changes in free-living bacterial communities along a salinity gradient in two temperate estuaries. Limnol Oceanogr 47: 453-470.

Bouvier TC, del Giorgio PA (2007) Key role of selective viral-induced mortality in determining marine bacterial community composition. Environ Microbiol 9: 287-297.

Brahamsha B (1996a) An abundant cell-surface polypeptide is required for swimming by the nonflagellated marine cyanobacterium *Synechococcus*. Proc Natl Acad Sci USA 93: 6504-6509.

Brahamsha B (1996b) A genetic manipulation system for oceanic cyanobacteria of the genus *Synechococcus*. Appl Environ Microbiol 62: 1747-1751.

Bratbak G, Thingstad F, Heldal M (1994) Viruses and the microbial loop. Microb Ecol 28: 209-221.

Brown CM, Lawrence JE, Campbell DA (2006) Are phytoplankton population density maxima predictable through analysis of host and viral genomic DNA content? J Mar Biol Ass UK 86: 491-498.

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

Campbell L, Carpenter EJ, Iacono VJ (1983) Identification and enumeration of marine chroococcoid cyanobacteria by immunofluorescence. Appl Environ Microbiol 46: 553-559.

Ceyssens P-J, Lavigne R, Mattheus W, Chibeu A, Hertveldt K et al. (2006) Genomic analysis of *Pseudomonas aeruginosa* phages LKD16 and LKA1: Establishment of the ϕKMV subgroup within the T7 supergroup. J Bacteriol 188: 6924-6931.

Chen F, Suttle CA (1995) Amplification of DNA polymerase gene fragments from viruses infecting microalgae. Appl Environ Microbiol 61: 1274-1278.

Chen F, Suttle CA (1996) Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. Virology 219: 170-178.

Chen F, Lu JR (2002) Genomic sequence and evolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. Appl Environ Microbiol 68: 2589-2594.

Chen F, Suttle CA, Short SM (1996) Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. Appl Environ Microbiol 62: 2869-2874.

Chen F, Lu JR, Binder B, Hodson RE (2001) Application of digital image analysis and flow cytometry to enumerate marine viruses stained with SYBR Gold. Appl Environ Microbiol 67: 539-545.

Chen F, Wang K, Bachoon DS, Lu J, Lau S et al. (2004) Phylogenetic diversity of *Synechococcus* in the Chesapeake Bay revealed by ribulose-1,5-bisphosphate carboxylase-oxygenase (RuBisCO) large subunit gene (*rbcL*) sequences. Aquat Microb Ecol 36: 153-164.

Chen F, Wang K, Kan J, Suzuki MT, Wommack KE (2006a) Diverse and unique picocyanobacteria in Chesapeake Bay, revealed by 16S-23S rRNA internal transcribed spacer sequences. Appl Environ Microbiol 72: 2239-2243.

Chen F, Wang K, Stewart J, Belas R (2006b) Induction of multiple prophages from a marine bacterium: a genomic approach. Appl Environ Microbiol 72: 4995-5001.

Chisholm SW (1992) Phytoplankton size. pp. 213-237. *In* Falkowski PG, Woodhead AD (ed.), Primary Productivity and Biogeochemical Cycles in the Sea. Plenum Press, New York.

Chisholm SW, Olson RJ, Zettler ER, Goericke R, Waterbury JB et al. (1988) A novel free-living prochlorophyte abundant in the oceanic euphotic zone. Nature 334: 340-343.

Christaki U, Courties C, Karayanni H, Giannakourou A, Maravelias C et al. (2002) Dynamic characteristics of *Prochlorococcus* and *Synechococcus* consumption by bacterivorous nanoflagellates. Microb Ecol 43: 341-352.

Chiura, HX (1997) Generalized gene transfer by virus-like particles from marine bacteria. Aquat Microb Ecol 13: 75-83.

Clegg MT (1993) Chloroplast gene sequences and the study of plant evolution. Proc Natl Acad Sci USA 90: 363-367.

Clokie MRJ, Mann NH (2006) Marine cyanophages and light. Environ Microbiol 8: 2074-2082.

Clokie MRJ, Shan J, Bailey S, Jia Y, Krisch HM et al. (2006) Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. Environ Microbiol 8: 827-835.

Collier JL, Brahamsha B, Palenik B (1999) The marine cyanobacterium *Synechococcus* sp. WH7805 requires urease (urea amidohydrolase, EC 3.5.1.5) to utilize urea as a nitrogen source: molecular-genetic and biochemical analysis of the enzyme. Microbiology 145: 447-459.

Crosbie ND, Pöckl M, Weisse T (2003) Dispersal and phylogenetic diversity of nonmarine picocyanobacteria, Inferred from 16S rRNA Gene and cpcBA-intergenic spacer sequence analyses. Appl Environ Microbiol 69: 5716-5721.

DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ et al. (2006) Community genomics among stratified microbial assemblages in the ocean's interior. Science 311: 496-503.

Dolan JR, Šimek K (1999) Diel periodicity in *Synechococcus* populations and grazing by heterotrophic nanoflagellates: Analysis of food vacuole contents. Limnol Oceanogr 44: 1565-1570.

Dorigo U, Jacquet S, Humbert JF (2004) Cyanophage diversity, inferred from g20 gene analyses, in the largest natural lake in France, Lake Bourget. Appl Environ Microbiol 70: 1017-1022.

Ducklow HW, Shiah FK (1993) Estuarine bacterial production. pp.261-284. *In*: Ford T (ed.), Aquatic Microbiology: An ecological approach. Blackwell, London.

Ducklow HW, Schultz G, Raymond P, Bauer J, and Shiah FK (1999) Bacterial dynamics in large and small estuaries. Proceedings of the 8th International Symposium on Microbial Ecology, Halifax, Canada.

Dufresne A, Salanoubat M, Partensky F, Artiguenave F, Axmann IM et al. (2003) Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome. Proc Natl Acad Sci USA 100: 10020-10025.

Elsaied H, Naganuma T (2001) Phylogenetic diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from deep-sea microorganisms. Appl Environ Microbiol 67: 1751-1765.

Ernst A, Becker S, Wollenzien UIA, Postius C (2003) Ecosystem-dependent adaptive radiations of picocyanobacteria inferred from 16S rRNA and ITS-1 sequence analysis. Microbiology 149: 217-228.

Felsenstein J (1993) PHYLIP: phylogeny Inference Package (version 3.5). University of Washington, Seattle.

Ferris MJ, Palenik B (1998) Niche adaptation in ocean cyanobacteria. Nature 396: 226-228.

Ferris MJ, Kuhl M, Wieland A, Ward DM (2003) Cyanobacterial ecotypes in different optical microenvironments of a 68˚C hot spring mat community revealed by 16-23S rRNA internal transcribed spacer region variation. Appl Environ Microbiol 69: 2893-2898.

Frederickson CM, Short SM, Suttle CA (2003) The physical environment affects cyanophage communities in British Columbia inlets. Microb Ecol 46: 348-357.

Fuhrman JA (1992) Bacterioplankton roles in cycling of organic matter: The microbial food web. pp. 361-383. *In* Falkowski PG, Woodhead AD (ed.), Primary Productivity and Biogeochemical Cycles in the Sea. Plenum Press, New York.

Fuhrman JA (1999) Marine viruses and their biogeochemical and ecological effects. Nature 399: 541-547.

Fuhrman JA, Suttle CA (1993) Viruses in marine planktonic systems. Oceanography 6: 51-63.

Fuhrman JA, Noble RT (1995) Viruses and protists cause similar bacterial mortality in coastal seawater. Limnol Oceanogr 40: 1236-1242.

Fuhrman JA, Campbell L (1998) Marine ecology: Microbial microdiversity. Nature 393: 410-411.

Fuhrman JA, Schwalbach M (2003) Viral influence on aquatic bacterial communities. Biol Bull 204: 192-195.

Fuhrman JA, Wilcox RM, Noble RT, Law NC (1993) Viruses in marine food webs. Trends Microb Ecol: 295-298.

Fuller NJ, Wilson WH, Joint IR, Mann NH (1998) Occurrence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. Appl Environ Microbiol 64: 2051-2060.

Fuller NJ, Marie MD, Partensky F, Vaulot D, Post AF et al. (2003) Clade-specific 16S ribosomal DNA oligonucleotides reveal the predominance of a single marine *Synechococcus* clade throughout a stratified water column in the Red Sea. Appl Environ Microbiol 69: 2430-2443.

Garza DR, Suttle CA (1998) The effect of cyanophages on the mortality of *Synechococcus* spp. and selection for UV resistant viral communities. Microb Ecol 36: 281.

Giovannoni SJ, Britschgi TB, Moyer CL, Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. Nature 345: 60-63.

Goericke R, Welschmeyer NA (1993) The marine prochlorophyte *Prochlorococcus* contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. Deep Sea Res 40: 2283-2294.

Gregory MF, Watson F, Tabita R (1997) Microbial ribulose 1,5-bisphosphate carboxylase: a molecule for phylogenetic and enzymeological investigation. FEMS Microbiol Lett 146: 13-22.

Hambly E, Tetart F, Desplats C, Wilson WH, Krisch HM et al. (2001) A conserved genetic module that encodes the major virion components in both the coliphage T4 and the marine cyanophage S-PM2. Proc Natl Acad Sci USA 98: 11411-11416.

Hardies SC, Comeau AM, Serwer P, Suttle CA (2003) The complete sequence of marine bacteriophage VpV262 infecting *Vibrio parahaemolyticus* indicates that an ancestral component of a T7 viral supergroup is widespread in the marine environment. Virology 310: 359-371.

Heidelberg, JF, Heidelberg KB, Colwell RR (2002) Seasonality of Chesapeake Bay bacterioplankton species. Appl Environ Microbiol 68: 5488-5497.

Hendrix RW (1999) Evolution: the long evolutionary reach of viruses. Curr Biol 9: 914-917.

Hendrix RW, Lawrence JG, Hatfull GF, Casjens S (2000) The origins and ongoing evolution of viruses. Trends Microbiol 8: 504-508.

Hendrix RW, Smith MCM, Burns RN, Ford ME, Hatfull GF (1999) Evolutionary relationships among diverse bacteriophages and prophages: All the world's a phage. Proc Natl Acad Sci USA 96: 2192-2197.

Hennes KP, Simon M (1995) Significance of bacteriophages for controlling bacterioplankton growth in a mesotrophic lake. Appl Environ Microbiol 61: 333-340.

Herdman M, Castenholz RE, Iteman I, Waterbury JR, Rippka R (2001) Subsection I (Formerly Chroococcales Wettstein 1924, emend. Rippka, Deruelles, Waterbury, Herdman and Stanier1979). pp. 493-514. *In* Boone DR,

Castenholz RW, Garrity GM (ed.) Bergey's Manual of Systematic Bacteriology, 2nd ed., Springer Publishers, New York.

Hewson I, Fuhrman JA (2006) Viral impacts upon marine bacterioplankton assemblage structure. J Mar Biol Ass UK 86: 577-589.

Hewson I, Fuhrman JA (2007) Characterization of lysogens in bacterioplankton assemblages of the Southern California borderland. Microbial Ecol 53: 631-638.

Hewson I, Winget DM, Williamson KE, Fuhrman JA, Wommack KE (2006) Viral and bacterial assemblage covariance in oligotrophic waters of the West Florida Shelf (Gulf of Mexico). J Mar Biol Ass UK 86: 591-603.

Honda D, Yokota A, Sugiyama J (1999) Detection of seven major evolutionary lineages in cyanobacteria based on the 16S rRNA gene sequence analysis with new sequences of five marine *Synechococcus* strains. J Mol Evol 48: 723-739.

Horrigan SG, Montoya JP, Nevins JL, McCarthy JJ, Ducklow HW et al. (1990) Nitrogenous nutrient transformations in the spring and fall inthe Chesapeake Bay. Estuar Coast Shelf Sci 30:369-391.

Iriarte A (1993) Size-fractionated chlorophyll *a* biomass and picophytoplankton cell density along a longitudinal axis of a temperate estuary (Southampton Water). J Plankton Res 15: 485-500.

Iriarte A, Purdie DA (1994) Size distribution of chlorophyll a biomass and primary production in a temperate estuary (Southampton Water): the contribution of photosynthetic picoplankton. Mar Ecol Prog Ser 115: 283-297.

Iteman I, Rippka R, Marsac NT, Herdman M (2000) Comparison of conserved structural and regulatory domains within divergent 16S rRNA-23S rRNA spacer sequences of cyanobacteria. Microbiology 146: 1275-1286.

Janse I, Meima M, Edwin W, Karkinaal A, Zwart G (2003) High-resolution differentiation of cyanobacteria by using rRNA-internal transcribed spacer denaturing gradient gel electrophoresis. Appl Environ Microbiol 69: 6634-6643.

Jiang SC, Kellogg CA, Paul JH (1998) Characterization of marine temperate phage-host systems isolated from Mamala Bay, Oahu, Hawaii. Appl Environ Microbiol 64: 535-542.

Jonas RB, Tuttle JH, Stoner DL, Ducklow HW (1988) Dual-label radioisotope method for simultaneously measuring bacterial production and metabolism in natural waters. Appl Environ Microbiol 54: 791-798.

Jonas RB, Tuttle JH (1990) Bacterioplankton and organic carbon dynamics in the lower mesohaline Chesapeake Bay. Appl Environ Microbiol 56: 747-757.

Johnson PW, Sieburth JM (1979) Chroococcoid cyanobacteria in the sea: A ubiquitous and diverse phototrophic biomass. Limnol Oceanogr 24: 928-935.

Kan J, Crump BC, Wang K, Chen F (2006) Bacterioplankton community in Chesapeake Bay: Predictable or random assemblages. Limnol Oceanogr 51: 2157-2169.

Kan J, Suzuki M, Wang K, Evans SE, Chen F (2007) High temporal but low spatial heterogeneity of bacterioplankton in the Chesapeake Bay. Appl. Environ. Microbiol. 73: 6776-6789.

Kitts CL (2001) Terminal restriction fragment patterns: A tool for comparing microbial communities and assessing community dynamics. Curr Issues Intestinal Microbiol 2: 17-25.

Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. Briefings in Bioinformatics 5: 150-163.

Laloui W, Palinska KA, Rippka R, Partensky F, Marsac NT et al. (2002) Genotyping of axenic and non-axenic isolates of the genus *Prochlorococcus* and the OMF-'*Synechococcus*' clade by size, sequence analysis or RFLP of the Internal Transcribed Spacer of the ribosomal operon. Microbiology 148: 453-465.

Lenski RE (1988) Dynamics of interactions between bacteria and virulent bacteriophage. Adv Microb Ecol 10: 1-44.

Li WKW (1994) Primary production of prochlorophytes, cyanobacteria, and eukaryotic ultraphytoplankton: measurements from flow cytometric sorting. Limnol Oceanogr 39: 169-175.

Li WKW (1995) Composition of ultraphytoplankton in the Central North Atlantic. Mar Ecol Prog Ser 122: 1-8.

Li WKW (1998) Annual average abundance of heterotrophic bacteria and *Synechococcus* in surface waters. Limnol Oceanogr 43: 1745-1753.

Li WKW, Subba Rao DV, Harrison WG, Smith JC, Cullen JJ et al. (1983) Autotrophic picoplankton in the tropical ocean. Science 219: 292-295.

Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW (2005) Photosynthesis genes in marine viruses yield proteins during host infection. Nature 438: 86-89.

Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F et al. (2004) Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. Proc Natl Acad Sci USA 101: 11013-11018.

Lindell D, Jaffe JD, Coleman ML, Futschik ME, Axmann IM et al. (2007) Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. Nature 449: 83-86.

Liu HB, Nolla HA, Campbell L (1997a) *Prochlorococcus* growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. Aquat Microb Ecol 12: 39-47.

Liu WT, Marsh TL, Cheng H, Forney LJ (1997b) Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl Environ Microbiol 63: 4516.

Liu X, Shi M, Kong S, Gao Y, An C (2007) Cyanophage Pf-WMP4, a T7-like phage infecting the freshwater cyanobacterium *Phormidium foveolarum*: Complete genome sequence and DNA translocation. Virology 366: 28-39.

Louis VR, Russek-Cohen E, Choopun N, Rivera ING, Gangle B et al. (2003) Predictability of *Vibrio cholerae* in Chesapeake Bay. Appl Environ Microbiol 69: 2773–2785.

Lu J, Chen F, Hodson RE (2001) Distribution, isolation, host specificity, and diversity of cyanophages infecting marine *Synechococcus* spp. in the Georgia river estuaries. Appl Environ Microbiol 67: 3285-3290.

Malone TC (1992) Effects of water column processes on dissolved oxygen, nutrients, phytoplankton and zooplankton. *In* Smith DE, Leffler M, Makiernan G, (ed.) Oxygen Dynamics in the Chesapeake Bay. Maryland Sea Grant, College Park.

Malone TC, Ducklow HW, Peele ER, Pike SE (1991) Picoplankton carbon flux in Chesapeake Bay. Mar Ecol Prog Ser 78: 11-22.

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor.

Mann NH (2003) Phages of the marine cyanobacterial picophytoplankton. FEMS Microbiol Rev 27: 17-34.

Mann NH, Cook A, Millard A, Bailey S, Clokie M (2003) Bacterial photosynthesis genes in a virus. Nature 424: 741.

Mann NH, Clokie MRJ, Millard A, Cook A, Wilson WH et al. (2005) The genome of S-PM2, a "photosynthetic" T4-type bacteriophage that infects marine *Synechococcus* strains. Appl Environ Microbiol 187: 3188-3200.

Marsh TL (1999) Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplicons. Curr Opin Microbiol 2: 323-327.

Marston MF, Sallee JL (2003) Genetic diversity and temporal variation in the cyanophage community infecting marine *Synechococcus* species in Rhode Island's coastal waters. Appl Environ Microbiol 69: 4639-4647.

Martin E, Benson R (1988) Phages of cyanobacteria. pp. 607-645. *In* Calendar R (ed.) The bacteriophages. Plenum Publisher, New York.

McDaniel L, Paul JH (2005) Effect of nutrient addition and environmental factors on prophage induction in natural populations of marine *Synechococcus* species. Appl Environ Microbiol 71: 842-850.

McDaniel L, Houchin LA, Williamson SJ, Paul JH (2002) Plankton blooms: Lysogeny in marine *Synechococcus*. Nature 415: 496.

McFadden BA, Tabita FR (1974) D-ribulose 1,5-bisphosphate carboxylase and the evolution of autotrophy. Bio Systems 6: 93-112.

Middelboe M, Hagström A, Blackburn N, Sinn B, Fischer U et al. (2001) Effects of bacteriophages on the population dynamics of four strains of pelagic marine bacteria. Microb Ecol 42: 395-406.

Millard A, Clokie MR, Shub DA, Mann NH (2004) Genetic organization of the *psbAD* region in phages infecting marine *Synechococcus* strains. Proc Natl Acad Sci USA 101: 11007–11012.

Miziorko HM, Lorimer. GH (1983) Ribulose 1,5-bisphosphate carboxylase-oxygenase. Annu Rev Biochem 52: 507-535.

Moebus K (1996) Marine bacteriophage reproduction under nutrient-limited growth of host bacteria .II. Investigations with phage-host system [H3:H3/1]. Mar Ecol Prog Ser 144: 13-22.

Molineux IJ (2006) The T7 group. pp. 277-301. *In* Calendar R (ed.) The bacteriophages. 2nd ed. Oxford University Press, Oxford.

Moore LR, Rocap G, Chisholm. SW (1998) Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. Nature 393: 464-467.

Moore LR, Post AF, Rocap G, Chisholm SW (2002) Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. Limnol Oceanogr 47: 989-996.

Mühling M, Fuller NJ, Millard A, Somerfield PJ, Marie D et al. (2005) Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. Environ Microbiol 7: 499-508.

Murray AG, Jackson GA (1992) Viral dynamics: a model of the effects of size, shape, motion and abundance of single-celled planktonic organisms and other particles. Mar Ecol Prog Ser 89: 103-116.

Murrell MC, Lores EM (2004) Phytoplankton and zooplankton seasonal dynamics in a subtropical estuary: importance of cyanobacteria. J Plankton Res 26: 371-382.

Muyzer G (1999) Genetic fingerprinting of microbial communities: present status and future perspective. *In* Bell CR, Brylinsky M, Johnson-Green P (ed.), Microbial Biosystems: New frontiers. Proceedings of the 8th International Symposium on Microbial Ecology. Atlantic Canada Society for Microbial Ecology, Halifax, Canada.

Muyzer G, Waal EC, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59: 695-700.

Ning X, Cloern JE, Cole BE (2000) Spatial and temporal variability of picocyanobacteria *Synechococcus* sp. in San Francisco Bay. Limnol Oceanogr 45: 695-702.

Nixon S, Ammerman J, Atkinson L, Berounsky V, Billen G et al. (1996) The fate of nitrogen and phosphorus at the land-sea margin of the North Atlantic Ocean. Biogeochemistry 35: 141-180.

Noble PA, Bidle KD, Fletcher M (1997) Natural microbial community compositions compared by a back-propagating neural network and cluster analysis of 5S rRNA. Appl Environ Microbiol 63:1762-1770.

Officer CB, Biggs RB, Taft JL, Cronin LE, Tyler MA et al. (1984) Chesapeake Bay anoxia: Origin, development, and significance. Science 223: 22-27.

Olson RJ, Chisholm SW, Zettler ER, Armbrust EV (1988) Analysis of *Synechococcus* pigment types in the sea using single and dual beam flow cytometry. Deep Sea Res 35: 425-440.

Olson RJ, Chisholm SW, Zettler ER, Altabet MA, Dusenberry JA (1990) Spatial and temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. Deep Sea Res A, Oceanogr Res 37: 1033-1051.

Ortmann AC, Lawrence JE, Suttle CA (2002) Lysogeny and lytic viral production during a bloom of the cyanobacterium *Synechococcus* spp. Microb Ecol 43: 225-231.

Osborn AM, Moore ERB, Timmis KN (2000) An evaluation of terminal-restriction fragment length polymorphism (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ Microbiol 2: 39-50.

Pace ML (1988) Bacterial mortality and the fate of bacterial production. Hydrobiologia 159: 41-49.

Palenik B (1994) Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. Appl Environ Microbiol 60: 3212-3219.

Palenik B (2001) Chromatic adaptation in marine *Synechococcus* strains. Appl Environ Microbiol 67: 991-994.

Palenik B, Haselkorn R (1992) Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. Nature 355: 265-267.

Palenik B, Brahamsha B, Larimer FW, Land M, Hauser L et al. (2003) The genome of a motile marine *Synechococcus*. Nature 424: 1037-1042.

Palenik B, Ren Q, Dupont CL, Myers GS, Heidelberg JF et al. (2006) Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. Proc Natl Acad Sci USA 103: 13555-13559.

Pan LA, Zhang J, Zhang LH (2007) Picophytoplankton, nanophytoplankton, heterotrohpic bacteria and viruses in the Changjiang Estuary and adjacent coastal waters. J Plankton Res 29: 187-197.

Partensky F, Hess WR, Vaulot D (1999a) Prochlorococcus, a marine photosynthetic prokaryote of global significance. Microb Mol Biol Rev 63: 106-127.

Partensky F, Blanchot J, Vaulot D (1999b) Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: A review. Bulletin de l'Institut Oceanographique (Monaco) 19: 457-475.

Paul JH (1999) Microbial gene transfer: an ecological perspective. J Mol Microb Biotech 1: 45-50.

Paul JH, Sullivan MB (2005) Marine phage genomics: what have we learned? Curr Opin Biotech 16: 299-307.

Peduzzi P, Weinbauer MG (1993) Effect of concentrating the virus-rich 2-200-nm size fraction of seawater on the formation of algal flocs (marine snow). Limnol Oceanogr 38: 1562-1565.

Phlips EJ, Badylak S, Lynch TC (1999) Blooms of the picoplanktonic cyanobacterium *Synechococcus* in Florida Bay, a subtropical inner-shelf lagoon. Limnol Oceanogr 44: 1166-1175.

Pichard SL, Campbell L, Paul. JH (1997) Diversity of the rubulose bisphosphate carboxylase/oxygenase form I gene (*rbcL*) in natural phytoplankton communities. Appl Environ Microbiol 63: 3600-3606.

Platt T, Rao DVS, Irwin B (1983) Photosynthesis of picoplankton in the oligotrophic ocean. Nature 301: 702-704.

Pope WH, Weigele PR, Chang J, Pedulla ML, Ford ME et al. (2007) Genome sequence, structural proteins, and capsid organization of the cyanophage Syn5: A "horned" bacteriophage of marine *Synechococcus*. J Mol Biol 368: 966-981.

Proctor LM, Fuhrman JA (1990) Viral mortality of marine bacteria and cyanobacteria. Nature 343: 60-62.

Rao, VB, Thaker V, Black LW (1992) A phage T4 in vitro packaging system for cloning long DNA molecules. Gene 113: 25-33.

Ray RT, Haas LW, Sieracki ME (1989) Autotrophic picoplankton dynamics in a Chesapeake Bay sub-estuary. Mar Ecol Prog Ser 52: 273-285.

Riemann L, Middelboe M (2002) Viral lysis of marine bacterioplankton: Implications for organic matter cycling and bacterial clonal composition. Ophelia 56: 57-68.

Robertson BR, Tezuka N, Watanabe MM (2001) Phylogenetic analyses of *Synechococcus* strains (cyanobacteria) using sequences of 16S rDNA and part of the phycocyanin operon reveal multiple evolutionary lines and reflect phycobilin content. Int J Syst Evol Microbiol 51: 861-871.

Rocap G, Distel DL, Waterbury JB, Chisholm SW (2002) Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S- 23S ribosomal DNA internal transcribed spacer sequences. Appl Environ Microbiol 68: 1180-1191.

Rocap G, Larimer FW, Jane L, et al. (2003) Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. Nature 424: 1042-1047.

Rohwer F, Segall A, Steward G, Seguritan V, Breitbart M et al. (2000) The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with nonmarine phages. Limnol Oceanogr 45: 408-418.

Rudi K, Skulberg OM, Jakobsen KS (1998) Evolution of cyanobacteria by exchange of genetic material among phyletically related strains. J Bacteriol 180: 3453-3461.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S et al. (2007) The Sorcerer II global ocean sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biol 5: e77.

Safferman RS, Morris ME (1963) Algal virus: Isolation. Science 140: 679-680.

Safferman RS, Cannon RE, Desjardins PR, Gromov BV, Haselkorn R et al. (1983) Classification and nomenclature of viruses of cyanobacteria. Intervirology 19: 61-66.

Sandaa R-A, Larsen A (2006) Seasonal variations in virus-host populations in Norwegian coastal waters: Focusing on the cyanophage community infecting marine *Synechococcus* spp. Appl Environ Microbiol 72: 4610-4618.

Saye DJ, Miller RV (1989) The aquatic environment: consideration of horizontal gene transmission in a diversified habitat. pp. 223-259. *In* Levy SB, Miller RV (ed.) Gene Transfer in the Environment. McGraw-Hill, New York, N.Y.

Scanlan DJ, West NJ (2002) Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. FEMS Microbiol Ecol 40: 1-12.

Schmidt TM, DeLong EF, Pace NR (1991) Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J Bacteriol 173: 4371-4378.

Scholl D, Merril C (2005) The genome of bacteriophage K1F, a T7-like phage that has acquired the ability to replicate on K1 strains of *Escherichia coli.* J Bacteriol 187: 8499-8503.

Scholl D, Kieleczawa J, Kemp P, Rush J, Richardson CC et al. (2004) Genomic analysis of bacteriophages SP6 and K1-5, an estranged subgroup of the T7 supergroup. J Mol Biol 335: 1151-1171.

Schopf JW (2000) The Fossil Record: Tracing the Roots of the Cyanobacterial Lineage. pp. 13-35. *In* Whitton BA, Potts M, (ed.) The Ecology of Cyanobacteria: Their Diversity in Time and Space. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Schwalbach MS, Hewson I, Fuhrman JA (2004) Viral effects on bacterial community compostion in marine plankton microcosms. Aquat Microb Ecol 34: 117-127.

Seshadri R, Kravitz SA, Smarr L, Gilna P, Frazier M (2007) CAMERA: A community resource for metagenomics. PLoS Biol 5: e75.

Shapiro LP, Haugen EM (1988) Seasonal distribution and temperature tolerance of *Synechococcus* in Boothbay Harbor, Maine. Estuarine Coastal Shelf Sci 26: 517-525.

Sharon I, Tzahor S, Williamson S, Shmoish M, Man-Aharonovich D et al. (2007) Viral photosynthetic reaction centre genes and transcripts in the marine environment. ISME J advance online publication, 9 August 2007; doi:10.1038/ismej.2007.67.

Shiah FK, Ducklow HW (1994) Temperature regulation of heterotrophic bacterioplankton abundance, production, and specific growth rate. Limnol Oceanogr 39:1243-1258.

Shimada A, Kanai S, Maruyama T (1995) Partial sequence of ribulose-1,5-bisphosphate carboxylase/oxygenase and the physiology of *Prochloron* and *Prochlorococcus* (Prochlorales). J Mol Evol 40: 671-677.

Short CM, Suttle CA (2005) Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. Appl Environ Microbiol 71: 480-486.

Šimek K, Pernthaler J, Weinbauer MG, Hornak K, Dolan JR et al. (2001) Changes in bacterial community composition and dynamics and viral

mortality rates associated with enhanced flagellate grazing in a mesoeutrophic reservoir. Appl Environ Microbiol 67: 2723-2733.

Smith EM, Kemp WM (2003) Planktonic and bacterial respiration along an estuarine gradient: response to carbon and nutrient enrichment. Aquat Microb Ecol 30: 251-261.

Steward GF, Smith DC, Azam F (1996) Abundance and production of bacteria and viruses in the Bering and Chukchi seas. Mar Ecol Prog Ser 131: 287-300.

Sullivan MB, Waterbury JB, Chisholm SW (2003) Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus.* Nature 424: 1047-1051.

Sullivan MB, Coleman ML, Weigele P, Rohwer F, Chisholm SW (2005) Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. PLoS Biol 3: e144.

Sullivan MB, Lindell D, Lee JA, Thompson LR, Bielawski JP et al. (2006) Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. PLoS Biol 4: e234.

Suttle CA, Chan AM, Cottrell MT (1990) Infection of phytoplankton by viruses and reduction of primary productivity. Nature 347: 467-469.

Suttle CA (1992) Inhibition of photosynthesis in phytoplankton by the submicron size fraction concentrated from seawater. Mar Ecol Prog Ser 87: 105-112.

Suttle CA, Chan AM (1993) Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-reactivity and growth characteristics. Mar Ecol Prog Ser 92: 99-109.

Suttle CA, Chan AM (1994) Dynamics and distribution of cyanophages and their effect on marine *Synechococcus* spp. Appl Environ Microbiol 60: 3167-3174.

Suttle CA (1994) The significance of viruses to mortality in aquatic microbial communities. Microb Ecol 28: 237-243.

Suttle CA (2000) Cyanophages and their role in the ecology of cyanobacteria, pp. 563-589. *In* Whitton BA, Potts M, (ed.) The Ecology of Cyanobacteria: Their Diversity in Time and Space. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Suttle CA (2005a) The viriosphere: the greatest biological diversity on Earth and driver of global processes. Environ Microbiol 7: 481-482.

Suttle CA (2005b) Viruses in the sea. Nature 437: 356-361.

Suttle CA (2007) Marine viruses-major players in the global ecosystem. Nat Rev Micro 5: 801-812.

Suzuki MT, Beja O, Taylor LT, Delong EF (2001) Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. Environ Microbiol 3: 323-331.

Tamura K, Kumar S (2002) Evolutionary distance estimation under heterogeneous substitution pattern among lineages. Mol Biol Evol 19: 1727-1736.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596-1599.

Thingstad TF (2000) Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. Limnol Oceanogr 45: 1320-1328.

Thingstad TF, Lignell R (1997) Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. Aquat Microb Ecol 13: 19-27.

Thingstad TF, Heldal M, Bratbak G, Dundas I (1993) Are viruses important partners in pelagic food webs? Trends Ecol Evol 8: 209-213.

Thompson JR, Marcelino LA, Polz MF (2002) Heteroduplexes in mixed-template amplification: formation, consequence and elimination by "reconditioning PCR". Nucleic Acids Res 30: 2083-2088.

Ting CS, Rocap G, King J, Chisholm SW (2002) Cyanobacterial photosynthesis in the oceans: the origins and significance of divergent light-harvesting strategies. Trends Microbiol 10: 134-142.

Toledo G, Palenik B (1997) *Synechococcus* diversity in the California Current as seen by RNA polymerase (*rpoC1*) gene sequences of isolated strains. Appl Environ Microbiol 63: 4298-4303.

Toledo G, Palenik B, Brahamsha B (1999) Swimming marine *Synechococcus* strains with widely different photosynthetic pigment ratio form a monophyletic group. Appl Environ Microbiol 65: 5247-5251.

Urbach E, Robertson DL, Chisholm SW (1992) Multiple evolutionary origins of Prochlorophytes within the cyanobacterial radiation. Nature 355: 267-270.

Urbach E, Scanlan DJ, Distel DL, Waterbury JB, Chisholm SW (1998) Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). J Mol Evol 46: 188-201.

Wang K, Chen F (2004) Genetic diversity and population dynamics of cyanophage communities in the Chesapeake Bay. Aquat Microb Ecol 34: 105-116.

Waterbury JB, Willey JM (1988) Isolation and growth of marine planktonic cyanobacteria. Methods Enzymol 167: 100-105.

Waterbury JB, Rippka R (1989) The order Chroococcocales. pp. 1728-1746. *In* Krieg NR, Holt JB, (ed.) Bergey's Manual of Systematic Bacteriology.Williams & Wilkins, Baltimore.

Waterbury JB, Valois FW (1993) Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. Appl Environ Microbiol 59: 3393-3399.

Waterbury JB, Watson SW, Guillard RRL, Brand LE (1979) Widerspread occurrence of a unicellular, marine, planktonic cyanobacterium. Nature 277: 293-294.

Waterbury JB, Watson SW, Valois FW, Franks DG (1986) Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. pp. 71-120. *In* Platt T, Li WKW (ed.) Photosynthetic Picoplankton. Can. Bull. Fish. Aquat. Sci., Toronto.

Waterbury JB, Willey JM, Franks DG, Valois FW, Watson SW (1985) A cyanobacterium capable of swimming motility. Science 230: 74-76.

Watson GMF, Tabita. FR (1996) Regulation, unique gene organization, and unusual primary structure of carbon fixation genes from a marine phycoerythrin-containing cyanobacterium. Plant Mol Biol 32: 1103-1115.

Watson GMF, Tabita. FR (1997) Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation. FEMS Microbiol Lett 146: 13-22.

Wawrik B, Paul. JH, Campbell L, Griffin D, Houchin L et al. (2003) Vertical structure of the phytoplankton community associated with a coastal plume in the Gulf of Mexico. Mar Ecol Prog Ser 251: 87-101.

Weigele PR, Pope WH, Pedulla ML, Houtz JM, Smith AL et al. (2007) Genomic and structural analysis of Syn9, a cyanophage infecting marine *Prochlorococcus* and *Synechococcus*. Environ Microbiol 9: 1675-1695.

Weinbauer MG, Peduzzi P (1995) Effect of virus-rich high molecular weight concentrates of seawater on the dynamics of dissolved amino acids and carbohydrates. Mar Ecol Prog Ser 127: 245-253.

Weinbauer MG, Hofle MG (1998a) Significance of viral lysis and flagellate grazing as factors controlling bacterioplankton production in a eutrophic lake. Appl Environ Microbiol 64: 431-438.

Weinbauer MG, Hofle MG (1998b) Distribution and life strategies of two bacterial populations in a eutrophic lake. Appl Environ Microbiol 64: 3776-3783.

Weinbauer MG, Rassoulzadegan F (2004) Are viruses driving microbial diversification and diversity? Environ Microbiol 6: 1-11.

Weinbauer MG, Hornak K, Jezbera J, Nedoma J, Dolan JR et al. (2007) Synergistic and antagonistic effects of viral lysis and protistan grazing on bacterial biomass, production and diversity. Environ Microbiol 9: 777-788.

Whitman WB, Campbell DA, Wiebe WJ (1998) Prokaryotes: the unseen majority. Proc Natl Acad Sci USA 95: 6578-6583.

Whitton BA, Potts M (2000) The Ecology of Cyanobacteria: Their Diversity in Time and Space. Kluwer Academic Publishers, Dordrecht, Netherlands.

Wilhelm SW, Suttle CA (1999) Viruses and nutrient cycles in the sea: viruses play critical roles in the structure and function of aquatic food webs. Biosci 49: 781-788.

Wilhelm SW, Carberry MJ, Eldridge ML, Poorvin L, Saxton MA et al. (2006) Marine and freshwater cyanophages in a Laurentian Great Lake: evidence

from infectivity assays and molecular analyses of g20 genes. Appl Environ Microbiol 72: 4957-4963.

Wilson WH, Carr NG, Mann NH (1996) The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium *Synechococcus* sp. WH7803. J Phycol 32: 506-516.

Wilson WH, Joint IR, Carr NG, Mann NH (1993) Isolation and molecular characterization of five marine cyanophages propagated on *Synechococcus* sp. strain WH7803. Appl Environ Microbiol 59: 3736-3742.

Wilson WH, Nicholas JF, Joint IR, Mann NH (1999) Analysis of cyanophage diversity and population structure in a south-north transect of the Atlantic ocean. Bulletin de l'Institut Oceanographique (Monaco) 19: 209-216.

Wilson WH, Nicholas JF, Joint IR, Mann NH (2000) Analysis of cyanophage diversity in the marine environment using denaturing gradient gel electrophoresis, pp. 565-570. *In* Bell CR, Brylinsky M, Johnson-Green P (ed.), Microbial Biosystems: New Frontier. Proceedings of the 8th International Symposium on Microbial Ecology Halifax, Canada.

Winter C, Smit A, Herndl GJ, Weinbauer MG (2004) Impact of virioplankton on archaeal and bacterial community richness as assessed in seawater batch cultures. Appl Environ Microbiol 70: 804-813.

Wommack KE, Colwell RR (2000) Virioplankton: viruses in aquatic ecosystems. Microb Mol Biol Rev 64: 69-114.

Wommack KE, Ravel J, Hill RT, Colwell RR (1999a) Hybridization analysis of Chesapeake Bay virioplankton. Appl Environ Microbiol 65: 241-250.

Wommack KE, Ravel J, Hill RT, Colwell RR (1999b) Population dynamics of Chesapeake Bay virioplankton: Total community analysis using pulsed field gel electrophoresis. Appl Environ Microbiol 65: 231-240.

Wommack KE, Hill RT, Kessel M, Russek-Cohen E, Colwell RR (1992) Distribution of viruses in the Chesapeake Bay. Appl Environ Microbiol 58: 2965-2970.

Wood AM, Townsend D (1990) DNA polymorphism within the WH7803 serogroup of marine *Synechococcus* spp. (Cyanobacteria). J Phycol 26: 576-585.

Wood AM, Horan PK, Muirhead K, Phinney DA, Yentsch CM et al. (1985) Discrimination between types of pigments in marine *Synechococcus* spp. by scanning spectroscopy, epifluorescence microscopy, and flow cytometry. Limnol Oceanogr 30: 1303-1315.

You L, Suthers PF, Yin J (2002) Effects of *Escherichia coli* physiology on growth of phage T7 in vivo and in silico. J Bacteriol 184: 1888-1894.

Zeidner G, Bielawski JP, Shmoish M, Scanlan DJ, Sabehi G et al. (2005) Potential photosynthesis gene recombination between *Prochlorococcus* and *Synechococcus* via viral intermediates. Environ Microbiol 7: 1505-1513.

Zhong Y, Chen F, Wilhelm SW, Poorvin L, Hodson RE (2002) Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. Appl Environ Microbiol 68: 1576-1584.

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L et al. (2007) Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. Environ Microbiol: (published online) 27 September 2007; doi: 10.1111/j.1462-2920.2007.01440.x