Population Dynamics of Cyanomyovirus in a Tropical Eutrophic Reservoir

BEE HUI YEO¹, AND KARINA YEW-HOONG GIN²*

¹School of Civil and Environmental Engineering, Nanyang Technological University, Singapore; and ²Department of Civil and Environmental Engineering, National University of Singapore, Singapore

(Received April 22, 2014—Accepted October 16, 2014—Published online December 27, 2014)

Samples from three stations in Kranji Reservoir, Singapore (n = 21) were collected and analyzed for cyanomyovirus abundance and diversity. A total of 73 different g20 (viral capsid assembly protein genes) amino acid sequences were obtained from this study. A phylogenetic analysis revealed that the 73 segments were distributed in six major clusters (α to ζ), with four unique subclusters, which were identified as KRM-I, KRM-II, KRM-III, and KRM-IV. The cyanomyovirus community in Kranji Reservoir exhibited a large degree of diversity; the clones obtained in this study showed similarities to those from many different environments, including oceans, lakes, bays, and paddy floodwater, as well as clones from paddy field soils. However, the sequences in this study were generally found to be more closely related to the g20 sequences of freshwaters and brackish waters than those from marine environments. The rarefaction curves and Chao 1 indices from this study showed that the diversity of the cyanomyovirus community was greater during the Inter-monsoon periods than the Southwest and Northeast Monsoons. A few seasonal changes in the taxa were observed: (i) Cluster ζ was absent during the Southwest Monsoon, and (ii) most of the samples fell into Group 3 in the PCA score plot during the Northeast Monsoon, and the fraction of Cluster ε increased significantly.

Key words: cyanophages, reservoir, cyanobacteria, phylogenetic diversity

The discovery of an abundance of viruses in aquatic systems since the 1990s has re-initiated investigations into the ecological roles of viruses (6, 35). This re-initiated research has changed traditional and conceptual understanding of the function and regulation of aquatic ecosystems (15) with the realization of viruses as important players in aquatic ecosystems. Culture-independent approaches and molecular techniques have been chosen to better understand the diversity and distribution of viruses in the aquatic environment. Culture-based approaches require the maintenance and growth of a host organism (22) for viruses to infect. However, 95% of the bacteria reported to date cannot be cultivated (4). In contrast, molecular techniques allow researchers to directly characterize complex viral consortia (22).

The gene encoding the vertex portal protein (g20) in the Myoviridae family (the term cyanomyoviruses will be used hereafter for the sake of convenience) has been exploited to examine the diversity of natural cyanophage communities (14). Genetically similar g20 clones have been detected at distant sites with different nutrient status, temperatures, and salinities (11, 21). For example, clones from the Gulf of Mexico (marine environment, depth of 110 m) share 99% identity with g20 clones from Lake Constance (freshwater environment, depth of 3 m) (21). These findings suggest that some g20 clones are widely distributed with no apparent geometrical segregation (11) or that horizontal gene exchanges occur among cyanophage communities (21). However, Sullivan et al. (23) suggested that g20 clones were distributed according to geographic segregation. For example, the population structure in an estuary and open ocean differed from each other and clonal diversity changed from the surface to the deep chlorophyll maximum (DCM) layer, as shown by Zhong et al. (39). Thus, similar g20 clones can be obtained from various environments, but g20 diversity distribution patterns (such as phylogenetic trees) from different environments differ and unique clades can be obtained from individual locations.

The temporal variation of cyanomyoviruses has been demonstrated from various environments e.g. Lake Bourget, Chesapeake Bay, the Red Sea, Norwegian coastal water, paddy flood water, and paddy field soils (5, 16, 19, 31–33). Mühling et al. (16) demonstrated that cyanomyovirus population diversity was greater during the spring and winter than the summer and autumn, and co-varied with Synechococcus diversity. Thus, previous studies showed both spatial and temporal variations in g20 clone distribution.

Singapore is a country with a tropical rainforest climate that is divided into four periods according to the average prevailing wind direction: Northeast Monsoon (NE) (December to early March), Inter Monsoon (late March to May), Southwest Monsoon (SW) (June to September), and second Inter Monsoon (October to November). Rainfall in Singapore generally begins to increase in October (monthly mean: 158.8 mm) and peaks in December (monthly mean: 329.5 mm) during the Northeast Monsoon (17). Kranji Reservoir is part of the Kranji catchment, which consists of different land-uses including residential areas, agriculture, forests, and reserve areas (10). Chua et al. (3) demonstrated that surface runoff loadings from the Kranji Catchment positively correlated with peak flow. Gin et al. (9) also suggested that heavier precipitation during the Northeast Monsoon affected the environmental parameters by dilution. Therefore, reservoir waters in Singapore face temporal changes due to the monsoon seasons. Kranji Reservoir has been categorized as a lake between a eutrophic and hypereutrophic status (26). Diverse cyanobacteria populations have previously been detected (Aug 2008 to Feb 2009), and species included

* Corresponding author. E-mail: ceeginyh@nus.edu.sg; Tel: 65–6516–8104.
Cyanomyoviruses in a tropical eutrophic reservoir

Microcystis, Anabaena, and Synechococcus (27).

The objective of this study was to obtain a clearer understanding of the diversity and dynamics of cyanomyoviruses in Kranji Reservoir. Cyanomyovirus diversity was investigated using a PCR-cloning-sequencing approach with PCR primers targeting the g20 gene, which is myoviral-specific. In this study, we reported, for the first time, the distribution of g20 clones in tropical reservoir water. We also investigate the effects of monsoon seasons on g20 clone distribution and determined differences in the genetic makeup of cyanomyovirus populations in a Singapore reservoir and other aquatic environments.

Materials and Methods

Sampling site and field sampling

Kranji Reservoir was formed by damming the mouth of the Kranji estuary. The reservoir has a water surface area of approximately 2.8 km² and is mainly supplied by three tributaries in the southern part of the reservoir: Tengah River, Kangkar River, and Peng Siang River. Water flow from Pang Sua River was diverted into Kranji Reservoir at the end of 2005 as the fourth tributary at the southern River. Water flow from Pang Sua River was diverted into Kranji Reservoir: Tengah River, Kangkar River, and Peng Siang River. The objective of this study was to obtain a clearer understanding of the diversity and dynamics of cyanomyoviruses in Kranji Reservoir.

Environmental parameters

Water quality parameters were measured at each station and date according to standard methods. 

Concentration of viral communities and DNA extraction

A total of 12 L of surface water was collected with autoclaved, sample-rinsed carboys and transported to the laboratory immediately. Briefly, 12 L of sample was gently filtered through a 50 or 20-μm pore size with low binding protein membrane filters (Pall Corporation) to remove particulate matter, phytoplankton, and zooplankton. The filtrate was subsequently filtered through a 0.2-μm pore size membrane (Pall Corporation) to remove most of the bacteria. After undergoing pre-filtration, samples were stored in a 4°C cool room. The viruses in the filtrate of approximately 10 L of sample were concentrated to a final volume ranging from 25 to 500 mL using tangential flow filtration (TFF) with a 30 kDa membrane cassette (Pall Corporation). Secondary concentration was performed using an Amicon Ultra centrifugal filter device (Millipore).

Polymerase chain reaction (PCR)

PCR amplification was performed with CPS1/CPS8 primers (39) to detect the g20 gene. Amplification reactions were carried out in 50-μL volumes containing 1× Taq polymerase buffer (Promega), 4 mM MgCl₂, 0.2 mM of a deoxynucleosidetriphosphate (dNTP) mix, 0.5 μM of each primer, 2.5 units Taq polymerase with hot start (Promega), and 5 μL template DNA. PCR amplification was carried out with a Mastercycler® pro thermal cycler from Eppendorf and performed with the following cycle profile: initial denaturation at 94°C for 3 min, 36 cycles of denaturation at 94°C for 45 s, annealing at 36°C for 45 s, extension at 72°C for 1 min, and a final extension at 72°C for 5 min.

Polymerase chain reaction (PCR)

PCR amplification was performed with CPS1/CPS8 primers (39) to detect the g20 gene. Amplification reactions were carried out in 50-μL volumes containing 1× Taq polymerase buffer (Promega), 4 mM MgCl₂, 0.2 mM of a deoxynucleosidetriphosphate (dNTP) mix, 0.5 μM of each primer, 2.5 units Taq polymerase with hot start (Promega), and 5 μL template DNA. PCR amplification was carried out with a Mastercycler® pro thermal cycler from Eppendorf and performed with the following cycle profile: initial denaturation at 94°C for 3 min, 36 cycles of denaturation at 94°C for 45 s, annealing at 36°C for 45 s, extension at 72°C for 1 min, and a final extension at 72°C for 5 min.

Final PCR products were run on a 1.5% agarose gel in 1X Tris-Borate-EDTA stained with ethidium bromide for 1 h at 90 volts. The gel was visualized under UV trans-illumination and photographed to confirm the target amplicon. The correct base pair for the CPS1/CPS8 amplicon was ca. 592 bp.

Fig. 1. Map of sampling locations (Station A, Station B, and Station C) in Kranji Reservoir, Singapore.
Quantitative real-time PCR (qPCR)

A pair of g20 gene primers (CPS1 and CPS2) designed by Fuller et al. (7) specifically amplifies a gene encoding a capsid assembly protein (g20) belonging to cyanophages in the *Myoviridae* family. Matteson et al. (13) recently verified the specificity of CPS1/CPS2 primers (g20 primers used in this qPCR study) using DE-METAST-BLAST. Their findings confirmed that these primers could only amplify cyanomyovirus amplicons based on sequences in the Global Ocean Survey data set.

The standard used to determine the gene copy number of the g20 gene was based on the cloned plasmids using the specific sequence for g20, and the standard curve was obtained from the calculated plasmid concentration at various dilutions. The quantification of cyanomyoviruses that encode the g20 gene (CPS1/CPS2) was performed with 5 µL of DNA template, 0.5 µM of both primers, 10 µL of 2x reaction master mix (FastStart Universal SYBR Green Master Mix [ROX], Roche), and sterilized distilled water to a final volume of 20 µL. The thermal cycle and measurement of SYBR Green fluorescence binding were performed in a Light-Cycler qPCR using the following program: pre-incubation at 95°C for 15 min; amplification for 45 cycles consisting of denaturing at 95°C for 15 s, annealing at 58°C for 20 s, and extension at 72°C for 30 s; melting curve analysis (1 cycle amplification) heated to 95°C, annealing at 65°C for 60 s, and extension at 95°C. Only 6 samples from each station (total n=18) were enumerated for g20 gene concentration, i.e., with the exception of the December 2008 sample (this was due to a change in the qPCR machine leading to inconsistent results).

The qPCR products were run on a 3% agarose gel in 1x Tris-borate-EDTA (TBE buffer) stained with ethidium bromide at 100 volts for 1 h. The gels were photographed under UV transillumination with a camera (Kodak DC 290). The correct g20 qPCR amplicon was placed at ca. 165 bp.

Cloning

A pair of family-specific gene primers (CPS1 and CPS8) was used in this study. These primers were developed by Zhong et al. (39). These primers have been used to study the diversity of cyanomyoviruses infectious to *Synechococcus* and *Prochlorococcus* (23). However, not all cyanomyoviruses infectious to *Synechococcus* and *Prochlorococcus* were able to be detected by these primers (23).

The PCR amplicons from CPS1/CPS8 amplification were purified using the Wizard® SV Gel and PCR clean up system (Promega, Corporation, Madison, WI). Purified PCR products were subsequently cloned into the pGEM-T Vector System II (Promega) and the construct was transformed into JM109 competent cells according to the manufacturer’s instructions. At least eight clones were randomly picked up from each sample and the correct target fragments were confirmed using PCR. At least four successful transformants were magnified overnight in LB broth with ampicillin. Plasmid DNA was extracted from the competent cells using the Wizard® SV Miniprep DNA Purification System (Promega). The extracted plasmid DNA was then sent for sequencing. All sequencing services were provided by AITbiotech Company.

Phylogenetic analysis

All nucleotide sequences were edited and aligned with the BioEdit sequence alignment editor (http://www.mbio.ncsu.edu/bioedit/bioedit.html). After editing, sequences were compared with cyanophage prototypes and to other reference sequences available in the GenBank database using BLAST software. Subsequently, all sequences from this study together with some previously obtained cyanomyovirus isolate g20 gene sequences and uncultured g20 sequences from other environments were aligned with ClustalW Multiple Alignment and saved in the FAS format. A phylogenetic tree was constructed with MEGA 5 (25) using the aligned sequences. The phylogenetic tree was drawn using the neighbor-joining method with p-distance and 1,000 bootstrap replications.

The Chao-1 index and rarefaction curve were obtained for monsoons using PAST software (http://folk.uio.no/ohammer/past).

Statistical analysis

A Principal Component Analysis (PCA) was carried out using the PASW 18 software package (SPSS). The PCA analysis was used to compare samples collected from different sampling locations and different sampling periods. Six environmental variables (i.e., TN, TP, Chl, Turbidity, Secchi depth, and TSS) were chosen in this PCA analysis. These variables were previously reported to be significantly affected by cyanobacteria concentrations in Kranji Reservoir (27).

Nucleotide sequence accession numbers

The g20 sequences from this study were deposited in the GenBank database. The GenBank accession numbers for the sequences are KC485882–KC485966.

Results

Environmental characteristics

Tables S1 to S3 show the environmental variables monitored during this study. Table 1 shows the summary of environmental variables at Kranji Reservoir. Significantly higher concentrations of total nitrogen (TN) and total phosphorus (TP) were observed at Station C than at Stations A and B. The concentration of chlorophyll-a markedly increased at Station C between October 2008 (111 µg L⁻¹) and November 2008 (1,585 µg L⁻¹), and then remained between 76 µg L⁻¹ and

### Table 1. Environmental variables characterizing Stations A, B, and C from August 2008 to February 2009

| Variables                  | Station A          | Station B          | Station C          |
|----------------------------|--------------------|--------------------|--------------------|
|                            | Min. (month) | Max. (month) | Mean | Min. (month) | Max. (month) | Mean | Min. (month) | Max. (month) | Mean |
| Chlorophyll-a (µg L⁻¹)     | 41.9 (Oct)       | 100.1 (Feb)      | 72.46 | 40.2 (Nov)  | 109.6 (Sept) | 71.89 | 10.1 (Sept)  | 1584.8 (Nov) | 301.39 |
| Total nitrogen, TN (mg N L⁻¹) | 0.64 (Aug) | 2.50 (Sept)   | 1.19  | 0.61 (Aug)  | 3.67 (Sept)  | 1.23  | 1.01 (Aug)  | 2.91 (Dec)   | 2.09  |
| Total Phosphorus, TP (mg P L⁻¹) | 0.062 (Aug) | 0.101 (Feb) | 0.078 | 0.027 (Jan) | 0.110 (Feb) | 0.071 | 0.062 (Aug) | 0.211 (Feb) | 0.128 |
| TN/TP                      | 10.34 (Aug)    | 35.81 (Sept)    | 15.91 | 8.40 (Nov)  | 33.89 (Sept) | 17.32 | 10.29 (Oct) | 37.50 (Sept) | 18.64 |
| Calcium, Ca²⁺ (mg L⁻¹)     | 16.88 (Dec)    | 22.77 (Oct)     | 18.76 | 17.89 (Dec) | 21.94 (Oct)  | 19.43 | 15.56 (Nov) | 24.41 (Aug) | 21.2  |
| Magnesium, Mg²⁺ (mg L⁻¹)   | 1.88 (Nov)     | 2.77 (Feb)      | 2.41  | 1.98 (Nov)  | 2.76 (Feb)   | 2.41  | 1.04 (Nov)  | 3.65 (Aug)   | 2.56  |
| Dissolved oxygen, DO (mg L⁻¹) | 4.48 (Oct) | 7.65 (Jan)   | 6.07  | 4.09 (Nov)  | 7.12 (Sept)  | 5.83  | 2.97 (Aug)  | 8.68 (Oct)   | 5.85  |
| Temperature (°C)           | 26.7 (Jan)     | 30 (Sept)       | 28.68 | 26.5 (Jan)  | 30.7 (Sept)  | 28.84 | 26.2 (Nov/Jan) | 29.1 (Oct) | 27.83 |
| Conductivity (µs cm⁻¹)     | 185.35 (Oct)   | 241.95 (Sept)   | 201.6 | 183.6 (Jan) | 239.65 (Feb) | 202.08 | 106.5 (Nov) | 302.85 (Aug) | 209.09 |
| Turbidity (NTU)            | 25.05 (Aug)    | 51.6 (Sept)     | 35.55 | 13.25 (Nov) | 86.5 (Sept)  | 32.66 | 21 (Aug)    | 224.5 (Nov)  | 75.49 |
| pH                         | 7.58 (Oct)     | 8.8 (Feb)       | 8.13  | 7.34 (Nov)  | 9.01 (Sept)  | 8.15  | 7.5 (Nov)   | 9.7 (Feb)    | 8.34  |
| Total dissolved solid, TDS (g L⁻¹) | 0.103 (Nov) | 0.134 (Feb) | 0.114 | 0.104 (Nov) | 0.134 (Feb) | 0.114 | 0.063 (Nov) | 0.184 (Aug) | 0.120 |
| Total suspended solid, TSS (mg L⁻¹) | 17 (Oct) | 25.2 (Jan) | 19.8  | 10.8 (Dec)  | 28 (Sept)    | 20.69 | 10.4 (Nov)  | 210 (Nov)    | 51.62 |
| Secchi depth (cm)          | 43 (Aug)       | 67.5 (Oct)      | 52.29 | 39.5 (Aug)  | 69 (Dec)     | 57.21 | 19 (Nov)    | 70 (Aug)     | 43.36 |
Cyanomyoviruses in a tropical eutrophic reservoir

Cyanomyoviruses in a tropical eutrophic reservoir

other sampling sites, ranging from 10^2 to 10^5 gene copies mL^-1 (5.0 × 10^2 to 7.2 × 10^3 gene copies mL^-1 in Kranji Reservoir was closer to Norwegian coastal water (13). Our study showed that the g20 gene density at Station C was below the detection limit, but showed the presence of the g20 gene. A total of six sampling sites from three different locations. The amino acid sequence analyses revealed five pairs of identical amino acid sequences. Only one of the five pairs was from the same sample, while the rest were from different sampling dates, locations, or both. Sequence analyses of a total of 86 different g20 segments revealed that within the examined gene fragment, the most variable region consisted of an insertion/deletion site, which varied between sequences by up to four amino acids. The lengths of the g20 fragment varied among clones (between the primers): 546 bp for 64 clones (74.5%), 549 bp for 7 clones (8%), 552 bp for 14 clones (16.3%), and 554 bp for 1 clone. The identity within the clones in this study at the amino acid level ranged from 50% to 99%. The normal g20 fragment length was 546 bp. G20 fragments with 549 bp have also been identified from g20 clones from paddy floodwater and paddy field soils (31, 32), while those with 552 bp were reported only from paddy soil samples.

The BLAST search within NCBI showed that all our sequences at the amino acid level were under a putative conserved domain called Phage_T4_Gp20_Superfamily. Table S4 shows a summary of the closest relatives at the amino acid level, revealing that this study had the highest identities, varying from 60% to 96% with the g20 clones obtained from other environments (as listed in Table S4). The amino acid sequence analysis revealed no putative cyanomyovirus sequences. Therefore, these two clones were removed from the phylogenetic analysis and were not considered as g20 fragments. Therefore, g20 recovery efficiency was 97.7% (86/88).

The nucleotide sequence analysis revealed seven pairs of identical sequences and one set with three identical sequences at the nucleotide level (100%). All pairs of identical nucleotide sequences originated from the same sample, such as KRA1008M1 and KRA1008M2, which were 100% identical in their sequences and were from Station A on 21 October 2008. The 3 clones with identical nucleotide sequences were obtained on the same sampling date (16 January 2009), but from three different locations. The amino acid sequence analysis revealed five pairs of identical amino acid sequences. Only one of the five pairs was from the same sample, while the rest were from different sampling dates, locations, or both. Sequence analyses of a total of 86 different g20 segments revealed that within the examined gene fragment, the most variable region consisted of an insertion/deletion site, which varied between sequences by up to four amino acids. The lengths of the g20 fragment varied among clones (between the primers): 546 bp for 64 clones (74.5%), 549 bp for 7 clones (8%), 552 bp for 14 clones (16.3%), and 554 bp for 1 clone. The identity within the clones in this study at the amino acid level ranged from 50% to 99%. The normal g20 fragment length was 546 bp. G20 fragments with 549 bp have also been identified from g20 clones from paddy floodwater and paddy field soils (31, 32), while those with 552 bp were reported only from paddy soil samples.

The BLAST search within NCBI showed that all our sequences at the amino acid level were under a putative conserved domain called Phage_T4_Gp20_Superfamily. Table S4 shows a summary of the closest relatives at the amino acid level, revealing that this study had the highest identities, varying from 60% to 96% with the g20 clones obtained from other environments (as listed in Table S4). The clone names were labeled based on their location, sampling date, family (Myoviridae), and sample number. For example,
KRA0808M1 indicated clone sample 1 of the g20 fragment (M-Myoviridae) from Station A collected on August 2008. KRB and KRC represented Stations B and C respectively.

Table S4 shows that the clones from our study mostly had the highest similarities with the clones from fresh waters (e.g., Dianchi Lake, Cultus Lake, Lake Bourget, paddy field floodwater, and paddy field soil) and brackish waters (e.g., Sandusky Bay, Chesapeake Bay, Skidaway Estuarine, and ballast water). Only four clones (<5%) from this study had the highest similarities with marine environments (i.e., Atlantic Meridional Transect, DCM of Sargasso Sea and Beaufort Sea, Arctic Ocean). The highest identity clones compared favorably with the study by Wilhelm et al., (36), especially clones from Station B: 65.4% of the Station B clones had the highest similarities with the ballast water and Sandusky Bay clones obtained by Wilhelm et al. (2006). There were only two clones from Station B that showed the highest similarities with paddy field clones. However, Station A had 13 clones (48%) and Station C had 10 clones (38%) with the highest similarities with paddy field clones.

**Phylogenetic analyses**

In this study, after merging clones with identical amino acid sequences, 73 different g20 segments were obtained and used for phylogenetic analyses. Phylogenetic analyses were carried out based on alignments consisting of these 73 Kranji Reservoir g20 segments and 80 sequences from previous studies. The addition of the g20 segments from other environmental studies was to ensure that the representative sequences included most of the g20 phylogenetic lineage identified to date. The phylogenetic tree (Fig. 4) revealed that the g20 clones obtained from Kranji Reservoir were distributed into six major clusters (α to ζ). This assignment corresponded to previously reported g20 clones from marine water, freshwater, ballast water, paddy floodwater, and paddy field soil environments. The phylogenetic tree clustering of the g20 clones was consistent with previous findings, except that Cluster ζ was newly assigned. Clusters α to ε were clustered

**Fig. 4.** Neighbor-joining phylogenetic tree of the g20 gene amino acid sequence showing relationships between g20 amino acid sequences from Kranji Reservoir and other environmental sequences. Black circles indicate clones obtained in this study. Numbers in the parentheses are the accession numbers of amino acid sequences in the NCBI web site. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches (25). Bootstrap numbers less than 50 are not shown in the phylogenetic tree. The clone names are labeled according to the format WXYZ, where W is the sampling location, X is the sampling date, Y is the phage family (Myoviridae), and Z is the sample number.
Cyanomyoviruses in a tropical eutrophic reservoir

according to Wang et al. (31). One sequence, KRB1108M4, stood apart from these six clusters.

The following observations were made with reference to the phylogenetic tree (Fig. 4). Cluster α, with a bootstrap of 70%, contained clones from lakes (30, 37), paddy floodwater (31), Sandusky Bay (36), and this study (7 out of 86 clones). Cluster α exclusively included clones from Kranji Reservoir and clones from freshwater environments only. This was supported by Wang et al. (31), who also found that cluster α was specific to g20 genes from freshwater environments including Sandusky Bay particulates and floodwater from paddy fields.

Cluster β was the largest cluster and strongly supported (96%). It was observed for all sampling dates and, hence, could be the dominant species in Kranji Reservoir. It contained clones from freshwater environments (5, 31, 32, 36, 37), brackish waters (33), marine environments (21), and this study (41 clones, 48%). In this study, two unique subclusters were identified within Cluster β: KRM-I and KRM-II. Kranji Reservoir clones (14 clones, 16.5%) formed a weakly supported (47%) independent sister clade in subcluster KRM-I. This clade with another two sister clades as assigned by Wang et al. (31). PFS-I and PFS-II, established the strong bootstrap (100%) subcluster KRM-I in Cluster β. KRM-II was a strongly supported (100%) subcluster that contained seven clones from this study.

Cluster δ, also known as cluster CSP, was first designated by Short and Suttle (21). Cluster CSP contained g20 sequences from cyanophage isolates which infect marine Synechococcus and Prochlorococcus (21, 31). Cluster δ in this phylogenetic tree consisted of clones from marine waters (11, 23, 24, 39), freshwaters (36), and paddy floodwaters (31). This cluster was strongly supported (99%) and contained nine clones (11%) obtained from this study (all clones from August 2008 samples).

Cluster Y was the smallest cluster, but with strong bootstrap support (100%). Cluster Y only contained a clone from this study and two clones (PFW-CM11 and PFW-CM12) from paddy floodwaters (31). PFW-CM11 and PFW-CM12 formed a unique subcluster, PFW-IV, in Wang et al. (31)’s study.

Cluster ε was a weakly bootstrap-supported (40%) cluster that accounted for 25% of the 85 clones from this study. Cluster ε consisted of clones from the deep chlorophyll maximum (DCM) of the Sargasso Sea, Gulf Stream (39), Atlantic Ocean (11), Sandusky Bay (36), and lakes (5). Another subcluster within Cluster ε, consisting only of Kranji Reservoir clones, has been identified and named as KRM-III. KRM-III was a strongly supported (99%) subcluster consisting of eight clones (9.4%) from this study.

Cluster ζ was a newly identified cluster in this study. Cluster ζ with very weak bootstrap support (15%) contained clones from the DCM of the Sargasso Sea (39), Arctic cyanobacterial mat (21), paddy floodwater (31), paddy field soils (32), bay water (36), lake water (21), and clones from this study (12 clones, 14%). A subcluster within Cluster ζ consisting only of Kranji Reservoir clones was identified and named KRM-IV. KRM-IV was a strongly supported (96%) subcluster.

Clone KRB1108M4 was unique and not grouped to any cluster. It was most closely related to the DCM clone from the Sargasso Sea clone (SS4716).

Spatial changes in the cyanomyovirus community

In this study, the cluster distribution of cyanomyovirus communities was shown across the three locations (Station A, Station B, and Station C) based on the clones produced. The spatial cluster distribution in this study (Table 2) did not reveal differences in the overall patterns of the composition of cyanomyoviral assemblages for these three stations in Kranji Reservoir.

However, a certain degree of variation was observed in the proportion of clones for each cluster: (1) the proportion of cyanomyoviruses that belonged to Cluster ε were higher at Station B (7 out of 15 clones), (2) Station C had the highest proportion of Cluster α clones (4 out of 7), and (3) Station C had the lowest proportion of Cluster ε clones (2 out of 15).

Temporal changes in the cyanomyovirus community

The temporal cluster distribution of the myoviral cyanophage population is shown in Table 3. Between three to four g20 clusters were detected in samples collected from August 2008 to February 2009, except for November 2008 and January 2009.

The November 2008 sample from Station C was the only sample that showed no amplification when CPS1/CPS8 primers were used in PCR, and the sample concentration was under the detection limit when CPS1/CPS2 primers were used in qPCR. At this particular sampling site and date, the cyanobacterial bloom (with obvious scum) had the highest chlorophyll-a concentration (1,584 µg L⁻¹) and was an extreme case in the PCA. The November 2008 samples (from Stations A and B) exhibited greater genetic diversity than the other 6 months. Only 9 clones were obtained from the November 2008 samples, but these clones were distributed to 5 different clusters and six genotypes, and KRB1108M4 was absent in all clusters. In contrast, only two cyanophage genotypes (clusters β and ε) were detected in January 2009 when 11 clones had been sequenced.

The sequences belonging to Cluster δ only occurred in August 2008 and only one clone was obtained in November 2008 belonging to cluster Y. Cluster α was absent in January 2009 while cluster ε was absent in August 2008. The cluster α genotype was present for six months, but only consisted of seven clones, except for February 2009, during which

Table 2. Spatial cluster distributions of cyanomyoviruses

| Sampling point/Cluster | Number of clones (% of cluster from each station) | Total clones for each station |
|-----------------------|--------------------------------------------------|------------------------------|
|                       | β       | α      | γ       | δ       | ζ       | ε       |                      |
| Station A             | 17 (51.5) | 2 (6.1) | 1 (3.0) | 3 (9.1) | 4 (12.1)| 6 (18.2)| 33                     |
| Station B             | 12 (44.4) | 1 (3.7) | 0       | 3 (11.1)| 3 (11.1)| 7 (25.9)| 27 (with one outcluster) |
| Station C             | 12 (46.2) | 4 (15.4)| 0       | 3 (11.5)| 5 (19.2)| 2 (7.7)| 26                     |
two clones were detected. Cluster β was the only genotype present for all sampling dates; however, spatially, Cluster β was absent at Station B in August and November 2008. Cluster β was the dominant fraction in the September, October, and December 2008 and February 2009 clones.

The score plot of samples is shown in Fig. 5, in which samples were grouped into four groups. Groups 1 and 2 consisted of samples with high numbers of g20 sequences falling into cluster β and ζ (more closely related to g20 clones from paddy floodwater and paddy field soil). Group 3 consisted of samples with g20 sequences more closely related to g20 clones from lakes and bays. Group 3 samples were mainly from the NE monsoon. Group 4 samples had the most diverse clones, with clones closely related to different environments, including bays, paddy field soil, ballast water, paddy floodwater, and marine water.

Table 4 shows the fraction and distribution of clusters for the three different monsoons, i.e. SW Monsoon, Inter Monsoon, and NE Monsoon. The SW Monsoon in this study was represented by six samples with a total of 26 clones. The Inter Monsoon consisted of six samples over two months, with 23 clones produced from 5 samples (note that the Station C November 2008 sample did not give any amplification). A total of 36 clones obtained from 9 samples over three months represented the NE Monsoon.

Cluster β was the major cluster for the three different monsoon seasons. Cluster ζ only appeared from October 2008 onwards, and was absent in the SW Monsoon Season (August and September 2008), in which a total of 26 clones was investigated. Cluster ζ was more dominant during the Inter Monsoon period than the NE Monsoon. The fraction of Cluster ε increased significantly in the NE Monsoon. Twelve clones from the NE Monsoon fell into this cluster, while only two clones and one clone came from the SW Monsoon and Inter Monsoon, respectively. Subcluster KRM-III in Cluster ε only consisted of clones from the NE Monsoon.

Fig. 6 shows the rarefaction curves for three monsoon seasons. The SW and NE monsoon rarefaction curves reached an asymptote. However, the Inter Monsoon rarefaction curve did not reach an asymptote, thus suggesting that the diversity of g20 sequences was greater for the Inter-monsoon. The rarefaction curve results were supported by the Chao 1 index. The Chao 1 index is a non-parametric richness estimator that can be used to determine total richness (20). Chao 1 indices (mean) for the SW, NE, and Inter-monsoons were 4, 5, and 6.4 respectively. Clones for the SW and NE Monsoons were distributed to four clusters (taxa), as shown in Table 4. The clones obtained from this study were distributed into five clusters for the Inter-monsoon, with a Chao 1 index of 6.4, suggesting that greater diversity was present during the Inter Monsoon.

Discussion

Despite the important ecological roles played by cyanophages, data regarding the occurrence of cyanophages in tropical aquatic environments in Southeast Asian is lacking.
Cyanomyoviruses in a tropical eutrophic reservoir

The prevalence of cyanophages in a tropical eutrophic reservoir was examined in the present study. Local environmental g20 clones were obtained and compared with isolates from other environments using a phylogenetic analysis. Temporal and spatial changes in cluster distribution were obtained and analyzed according to changes in environmental parameters.

Phylogenetic diversity of g20 gene sequences in Kranji Reservoir

The cyanophage community in Kranji Reservoir revealed a large degree of diversity. Clones from this study showed similarities to those from different environments, including oceans, lakes, bays, paddy floodwater, and paddy field soil. A previous study on phytoplankton structures in Singapore reported that different Synechococcus communities were detected in the Singapore Strait and Johor Strait (8), and the Johor Strait was more eutrophic and had higher nutrient levels than the Singapore Strait. Different Synechococcus strains (a total of 9 Synechococcus clones) were present in Kranji Reservoir, including Synechococcus belonging to clones collected from a hydroelectric power plant reservoir, lakes, reservoir, and Chesapeake Bay (27). The richness of the genotypes of cyanobacteria in Kranji Reservoir is suggested to contribute to the large degree of diversity in this tropical cyanophage community.

Population dynamics of cyanomyoviruses

CPS1/CPS8 primers have successfully identified Synechococcus and Prochlorococcus phages by detecting the presence of the g20 gene (23, 39). However, no study to date has shown that these primers can detect Microcystis and Anabaena phages. The CPS1/CPS8 amplicons from this study are believed to originate from the g20 genes of cyanophages, which only infect Prochlorococcus or Synechococcus.

The temporal variations in cyanomyovirus cluster distribution exhibited significant temporal variations. Samples from August 2008 in cluster δ were closely related to known marine cyanomyovirus isolates. In this study, only 9 out of the 86 g20 clones examined belonged to cluster δ, indicating that marine type Synechococcus and Prochlorococcus were generally not predominant in Kranji Reservoir, except for August 2008. Table 3 shows that the Cluster β fraction markedly increased from September 2008, with the development of a unique sister clade in subcluster KRM-I (Fig. 4). Clones from this sister clade were more closely related to the clones in PFS-I. Wang et al. (32) previously suggested that Cluster β was the main cluster of the g20 gene from paddy field soils with PFS-I as a unique group. These changes in cluster distribution suggested that the Synechococcus community from Kranji Reservoir was more closely related to the Synechococcus community from paddy field soil. The PCA score plot (Fig. 5) also showed that the August 2008 and September 2008 samples were clustered into different groups. This temporal variation may shed some light on the role of cyanophages as indicators of shifts in the host community.

Phages and host interaction

The concept of “killing the winner” was proposed by Thingstad and Lignell (28) through a theoretical model. “Killing the winner” is an expression used to describe the process by which lytic viruses lyse the most successful population (fastest growing population) and allow for the co-existence of less competitive populations, thereby sustaining bacterial diversity (13, 34). The g20 gene concentration for Stations A and B showed a maximum in October (Fig. 3) and the proportions of the dominant cluster were 100% and 80%, respectively (Table S4). In November 2008, Stations A and B showed a significant decrease in the abundance of g20 genes and the fraction of Cluster β. Nine clones from Stations A and B were simultaneously classified into five different clusters, revealing a more diverse cyanomyovirus community. The PCA score plot showed few differences in the water conditions in these two months.

The “killing the winner” phenomenon may have occurred between October and November at these locations. In October, phages present in Kranji Reservoir that were able to infect the vulnerable dominant host could replicate effectively at a high contact rate, resulting in high concentrations, but low numbers of cyanomyovirus genotypes. The dominant species was subsequently lysed and a new Synechococcus community was formed. Mühling et al. (16) showed that cyanophage infection played an important role in the succession of Synechococcus genotypes. The cyanomyovirus diversity results obtained from this study demonstrated that cyanophage infection can potentially control the dominant species of cyanobacteria in Kranji Reservoir.

It was previously unclear whether cyanophages responded to changes in Synechococcus assemblages or if phage infection controlled the Synechococcus assemblage? Throughout this study, we found that fluctuations in water quality affected the growth of the Synechococcus community (during August 2008 and September 2008) and, subsequently, the genotype of cyanophages. With the emergence of dominant genotypes, cyanophages could be an important factor controlling the diversity of Synechococcus populations by “killing the winner” (during October 2008 and November 2008). As the host community diversifies, this, in turn, leads to an increase in the diversity of cyanophages. Thus, cyanophage diversity could potentially be used as an indicator of a shift in the strains of particular species of cyanobacteria.
References

1. APHA. 1995. Standard Methods for the Examination of Water and Wastewater-19th Edition. American Public Health Association, American Water Works Association and Water Environment Federation, Washington, D.C.

2. APHA. 1998. Standard methods for the Examination of Water and Wastewater, 20th Edition. American Public Health Association, American Water Works Association, and the Water Environment Federation, Washington, D.C.

3. Chua, L.H.C., E.Y.M. Lo, E.B. Shuy, and S.B.K. Tan. 2009. Nutrients and suspended solids in dry weather and storm flows from a tropical catchment with various proportions of rural and urban land use. J. Environ. Manage. 90:3635–3642.

4. Clokie, M.R.J., A.D. Millard, A.V. Letarov, and S. Heaphy. 2011. Cyanobacterial structural protein fene g20. Appl. Environ. Microbiol. 68:1576–1584.

5. Dorigo, U., S. Jacquet, and J.F. Humbert. 2004. Cyanophage diversity, and virus-host interactions in a Lake Bourget. Appl. Environ. Microbiol. 70:1017–1022.

6. Fuhrman, J.A. 1999. Marine viruses and their biogeochemical and ecological effects. Nature 399:541–548.

7. Fuller, N.J., W.H. Wilson, I.R. Joint, and N.H. Mann. 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.

8. Gin, K.-H., M. Holmes, S. Zhang, and X. Lin. 2006. Phycoplankton structure in the tropical port waters of Singapore, p. 347–375. In E. Wolanski (ed.), The Environment in Asia Pacific Harbours. Springer Netherlands.

9. Gin, K.Y.-H., X. Lin, and S. Zhang. 2000. Dynamics and size structure of phycoplankton in the coastal waters of Singapore. J. Plankton Res. 22:1465–1484.

10. Granger, E.C. 2010. Water quality modeling in Kranji catchment. Massachusetts Institute of Technology, Massachusetts.

11. Jameson, E., N.H. Mann, I. Joint, C. Sambles, and M. Mühling. 2011. The diversity of cyanomyovirus populations along a North-South Atlantic Ocean transect. Isme J. 5:1713–1721.

12. Matteson, A.R., S.N. Loar, R.A. Bourbonniere, and S.W. Wilhelm. 2011. Molecular enumeration of an ecologically important cyanophage in a Laurentian Great Lake. Appl. Environ. Microbiol. 77:6772–6779.

13. Matteson, A.R., J.M. Rowe, A.J. Ponsero, T.M. Pimentel, P.W. Boyd, and S.W. Wilhelm. 2013. High abundances of cyanomyoviruses in marine ecosystems demonstrate ecological relevance. FEMS Microbiol. Ecol. 84:223–234.

14. McDaniel, L.D., M. delaRosa, and J.H. Paul. 2006. Temperate and lytic cyanophages from the Gulf of Mexico. J. Mar. Bio. Assoc. U.K. 86:517–527.

15. Middelboe, M., S. Jacquet, and M. Weinbauer. 2008. Viruses in freshwater ecosystems: an introduction to the exploration of viruses in new aquatic habitats. Freshw. Biol. 53:1069–1075.

16. Mühling, M., N.J. Fuller, A. Millard, et al. 2005. Genetic diversity of marine Synechococcus and co-occurring cyanophage communities: evidence for viral control of phycoplankton. Environ. Microbiol. 7:499–508.

17. NEA. 2009. Weather Wise Singapore. National Environment Agency, Singapore.

18. Norušis, M.J. 2008. Chapter 17 Factor Analysis, p. 389–425, SPSS 16.0 statistical procedures companion. Prentice Hall.

19. Sandaa, R.A., and A. Larsen. 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.

20. Santos, M.A.S., and A.R. Soares. 2011. Computational tools for identification of microRNAs in deep sequencing data sets. In H. Lopes (ed.), Computational biology and applied bioinformatics. InTech.

21. Short, C.M., and C.A. Suttle. 2005. Nearly identical bacteriophage structural gene sequences are widely distributed in both marine and freshwater environments. Appl. Environ. Microbiol. 71:480–486.

22. Short, S.M., F. Chen, and S. Wilhelm. 2010. The construction and analysis of marker gene libraries, p. 82–91. In S.W. Wilhelm, M.G. Weinbauer, and C.A. Suttle (eds.), Manual of Aquatic Viral Ecology. American Society of Limnology and Oceanography, Inc., Texas.

23. Sullivan, M.B., M.L. Coleman, V. Quinnivan, et al. 2008. Portal protein diversity and phage ecology. Environ. Microbiol. 10:2810–2823.

24. Sullivan, M.B., K.H. Huang, J.C. Ignacio-Espinoza, et al. 2010. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. Environ. Microbiol. 12:3035–3056.

25. Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28:2731–2739.

26. Te, S.H., and K.Y.-H. Gin. 2011. The dynamics of cyanobacteria and microcystin production in a tropical reservoir of Singapore. Harmful Algae 10:319–329.

27. Te, S.H. 2012. Dynamics of cyanobacteria and cyanotoxin in a tropical reservoir. Ph.D Thesis. Nanyang Technological University, Singapore.

28. Thingstad, T., and R. Lignell. 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. Aquat. Microb. Ecol. 13:19–27.

29. USEPA. 1993. Determination of nitrate-nitrite nitrogen by automated colorimetry. U.S. ENVIRONMENTAL PROTECTION AGENCY, Washington, D.C.

30. Varela, V., E. Zaralli, K. Vareli, and L. Sainis. Cyanophage diversity in Lake Pamvotis. Unpublished.

31. Wang, G., J. Murase, S. Asakawa, and M. Kimura. 2010. Unique viral capsid assembly protein gene (g20) of cyanophages in the floodwater of a Japanese paddy field. Biol. Fert. Soils. 46:93–102.

32. Wang, G.H., S. Asakawa, and M. Kimura. 2010. Unique viral capsid assembly protein gene (g20) in FEMS Microbiol. Ecol. 76:352–359.

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

34. Weinbauer, M.G., and F. Rassoulzadegan. 2004. Are viruses driving microbial diversification and diversity? Environ. Microbiol. 6:1–11.

35. Wilhelm, S.W., and C.A. Suttle. 1999. Viruses and Nutrient Cycles in the Sea—Viruses play critical roles in the structure and function of aquatic food webs. Bioscience 49:781–788.

36. Wilhelm, S.W., M.J. Carberry, M.L. Eldridge, L. Poorvin, M.A. Saxton, and M.A. Doblin. 2006. Marine and freshwater cyanophages in a Laurentian Great Lake: Evidence from infectivity assays and molecular analyses of g20 gene. Appl. Environ. Microbiol. 72:4957–4963.

37. Xia, X.S., L. Liu, X. Zhang, F. Feng, M. Xue, and D.Q. Wei. Phylogenetic analysis of cyanophages, based on g20 analyses in Plateau Lakes of Yunnan, China. Unpublished.

38. Zhong, X., L. Berdjeb, and S. Jacquet. 2013. Temporal dynamics of procyanobacteria and cyanomyoviruses in two large and deep per-alpine lakes. FEMS Microbiol. Ecol. 86:312–326.

39. Zhong, X., F. Chen, S.W. Wilhelm, and R.E. Hodson. 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.