



FEATURE ARTICLE

# Water column stratification structures viral community composition in the Sargasso Sea

Dawn B. Goldsmith<sup>1</sup>, Jennifer R. Brum<sup>2,4</sup>, Max Hopkins<sup>1</sup>, Craig A. Carlson<sup>3</sup>,  
Mya Breitbart<sup>1,\*</sup>

<sup>1</sup>College of Marine Science, University of South Florida, St. Petersburg, FL 33701, USA

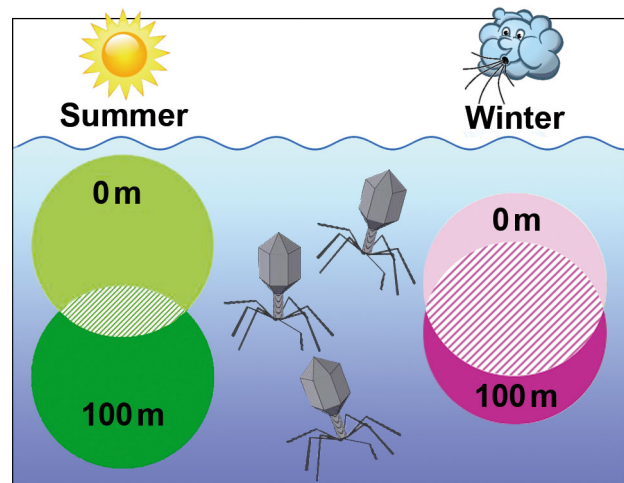
<sup>2</sup>Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA

<sup>3</sup>Department of Ecology, Evolution and Marine Biology, University of California, Santa Barbara, CA 93106-9620, USA

<sup>4</sup>Present address: Department of Microbiology, Ohio State University, Columbus, OH 43210, USA

**ABSTRACT:** A decade-long study of viral abundance at the Bermuda Atlantic Time-series Study (BATS) site recently revealed an annually recurring pattern where viral abundance was fairly uniform in the well-mixed upper water column each winter, yet a subsurface peak in viral abundance between 60 and 100 m depth developed each summer during water column stratification (Parsons et al. 2012; ISME J 6:273–284). Building upon these findings, this study tests the hypothesis that in the well-mixed period (March), the viral communities at the surface and at 100 m depth are similar in composition, while during water column stratification (September), differences in the viruses occupying these 2 depths emerge. Amplification and sequencing of 3 signature genes (*g23*, *phoH*, and the ssDNA phage major capsid protein) in addition to randomly amplified polymorphic DNA PCR gel banding patterns were used to assess the structure of viral communities. The 4 data sets revealed similar communities at the surface and 100 m in March when the upper water column was mixed, and divergent communities during September stratification. Water density was found to be a significant driver of viral community variability, with surface communities during September water column stratification significantly different from all other communities. These data demonstrate the importance of water column stratification for structuring viral community composition at the BATS site, either directly through altering the physical conditions, such as ultraviolet radiation, that the viral communities are exposed to or indirectly through structuring bacterial host communities.

\*Corresponding author: mya@usf.edu



Sargasso Sea viral communities at 0 and 100 m are similar in the well-mixed winter, but differ during summer water-column stratification.

Image: Mya Breitbart and Dawn B. Goldsmith

**KEY WORDS:** Marine · Virus · Phage · Signature gene · Diversity · Stratification · Season · Bermuda Atlantic Time-series Study

## INTRODUCTION

The Bermuda Atlantic Time-series Study (BATS) site in the Sargasso Sea lies in a transitional area between an oligotrophic, permanently stratified area to the south, and a eutrophic region to the north

© The authors 2015. Open Access under Creative Commons by Attribution Licence. Use, distribution and reproduction are unrestricted. Authors and original publication must be credited.

Publisher: Inter-Research · www.int-res.com

(Steinberg et al. 2001). This site experiences a significant seasonal pattern of temperature variability, convective mixing, and stratification (Siegel et al. 1999, Steinberg et al. 2001). In winter, convective mixing leads to mixed layer depths (MLD) that extend to between 150 and 300 m and results in nutrient entrainment into the surface layer (Michaels & Knap 1996, Steinberg et al. 2001, Lomas et al. 2013). As summer approaches, high pressure systems set into the region, resulting in lower wind stress and warming that cause the MLD to shoal to the surface (10 to 20 m) by late spring/early summer. As autumn progresses, temperatures cool, winds increase, and the mixed layer begins to extend to deeper depths again (Steinberg et al. 2001).

The predictable seasonal patterns at this site, along with the wealth of data available through the BATS program (<http://bats.bios.edu>), create an ideal opportunity for studying the temporal dynamics of marine viruses. Viruses play a key ecological role in the ocean by killing marine microbes, and their sheer abundance enables them to influence nutrient and energy cycles (Suttle 2007, Breitbart 2012). In a 10-yr study at the BATS site, Parsons et al. (2012) sampled viral abundance monthly at 12 depths from the surface to 300 m and discovered an annually recurring late summer subsurface viral maximum between 60 and 100 m depth. While viral abundance did not correlate with total bacterial abundance at the BATS site, viral abundance was positively correlated with the abundance of *Prochlorococcus*, the most abundant cyanobacterial group, and negatively correlated with the abundance of SAR11, the dominant heterotrophic bacterial lineage. These data led to the suggestion that patterns in viral abundance are strongly influenced by *Prochlorococcus* dynamics in response to water column stratification.

To further examine the seasonal viral dynamics observed by Parsons et al. (2012), this study investigated the composition of the viral community at the surface and at 100 m depth during the highly mixed period (March) and the thermally stratified period (September) at the BATS site in 3 different years. We hypothesized that in March, when the MLD often extends into the mesopelagic zone (>140 m), the surface viral community would resemble the 100 m viral community more closely than during times of water column stratification (September), when the MLD shoals into the upper euphotic zone to a depth shallower than the viral abundance maximum. Viral community composition was assessed using 4 different methods: through amplification and sequencing of 3 different signature genes (*g23*, *phoH*, and ssDNA

MCP) representing specific subsets of the viral community, and randomly amplified polymorphic DNA (RAPD) PCR for viral community profiling. The structural *g23* gene encodes the major capsid protein for T4-like myophage (Filée et al. 2005, Comeau & Krisch 2008, Chow & Fuhrman 2012). The auxiliary metabolic gene *phoH* is another core T4-like cyanophage gene (Ignacio-Espinoza & Sullivan 2012), but is also found in other myoviruses, podoviruses, and siphoviruses (Goldsmith et al. 2011). The *phoH* gene occurs in phage that infect both autotrophs and heterotrophs, and is more prevalent among marine phage than in the genomes of phage isolated from other environments (Goldsmith et al. 2011). Whereas both *g23* and *phoH* target double-stranded DNA (dsDNA) viruses, the third signature gene used in this study targets the major capsid protein (MCP) gene of single-stranded DNA (ssDNA) phage belonging to the *Gokushovirinae* subfamily (family *Microviridae*) (Roux et al. 2012, Hopkins et al. 2014). This combination of approaches for surveying the composition of portions of the viral community provided insight into the spatial and temporal variability in viral diversity in the context of extensive metadata at this well-studied marine site.

## MATERIALS AND METHODS

### Sample collection

Samples were collected from within the surface 5 m (hereafter referred to as '0 m' or 'surface') and at 100 m depth in the vicinity of the BATS site (31° 40' N, 64° 10' W) in March and September of 2008, 2010, and 2011 (see Table 1 for sampling dates, environmental metadata, and viral abundance). Samples were collected in years when the BATS program could accommodate requests for ship time and CTD usage. Although the exact depth of the subsurface peak varied among years, the depth of 100 m was chosen to ensure sampling below the mixed layer and approximate the subsurface peak in viral abundance based on prior data (Parsons et al. 2012). MLD during experimental periods was determined as the depth where sigma-*t* was equal to sea surface sigma-*t* plus an increment in sigma-*t* equivalent to a 0.3°C temperature decrease (Sprintall & Tomczak 1992). Ancillary data used to provide context for the annual patterns of MLD were retrieved from the public data set (<http://bats.bios.edu>). A broader assessment of the BATS data is presented elsewhere (Lomas et al. 2013).

Table 1. Sample collection data and environmental conditions for each sample. MLD: mixed layer depth; nd: no ssDNA MCP sequences were obtained from this sample because the PCR of this sample did not yield any bands. The source of viral concentrations is Parsons et al. (2012)

Year	Month	Day	Depth (m)	MLD (m)	Vol. water concentrated (l)	Number of sequences obtained	g23	phoH	ssDNA MCP	Viral conc. in whole water ( $\times 10^6 \text{ ml}^{-1}$ )	Temperature (°C)	Salinity	Water density ( $\text{sigma-t}$ )	Oxygen ( $\mu\text{mol kg}^{-1}$ )	Chlorophyll a ( $\text{ng kg}^{-1}$ )	Prochlorococcus ( $\times 10^4 \text{ ml}^{-1}$ )	Synechococcus ( $\times 10^3 \text{ ml}^{-1}$ )
2008	Mar	24	0	187	144	43	40	82	82	4.00	19.9	36.659	26.051	224.5	0	4.2	11.7
2008	Mar	24	100	187	125	45	38	74	74	3.67	19.4	36.657	26.189	220.8	144	11.0	40.5
2008	Sep	2-3	0	11	245	40	44	38	38	2.64	28.6	36.293	23.169	198.7	16	2.1	7.0
2008	Sep	2-3	100	11	245	46	42	95	95	2.86	19.5	36.612	26.140	207.4	63	5.1	0.7
2010	Mar	8	0	173	383	50	37	nd	nd	2.75	19.5	36.708	26.197	219.7	404	3.6	17.7
2010	Mar	8	100	173	288	34	34	27	27	2.51	19.0	36.657	26.304	197.6	191	0.9	6.4
2010	Sep	5	0	26	245	43	48	91	91	5.25	27.9	36.240	23.362	199.2	65	3.8	6.2
2010	Sep	7	100	26	90	36	48	90	90	4.25	19.7	36.650	26.106	205.2	158	2.5	0.1
2011	Mar	27	0	155	180	40	48	69	69	4.58	18.8	36.672	26.341	232.0	278	0.2	11.0
2011	Mar	27	100	155	180	43	47	85	85	4.61	18.7	36.660	26.381	224.0	212	0.3	15.6
2011	Sep	13	0	33	280	44	48	96	96	2.46	28.3	36.713	23.586	201.8	57	2.1	12.1
2011	Sep	13	100	33	280	45	46	96	96	5.74	20.9	36.796	25.908	211.8	269	9.4	4.1

## Sample processing

Approximately 200 l (Table 1; volumes ranged from 90 to 383 l, median volume 245 l) of seawater were concentrated for each sample by tangential flow filtration with 100 kDa filters (GE Healthcare) to a volume of approximately 50 ml. The viral concentrates were filtered through 0.22  $\mu\text{m}$  Sterivex filters (Millipore) to remove bacteria and stored at 4°C until further processing. Viruses were further concentrated and purified by polyethylene glycol (PEG) precipitation followed by cesium chloride density-dependent centrifugation (Thurber et al. 2009). Solid PEG 8000 was added to the viral concentrates at a final concentration of 10% (wt/vol), and the samples were stored at 4°C overnight before centrifuging for 40 min at 11 000  $\times g$  and 4°C to pellet the viruses. The pelleted viruses were resuspended in 0.02  $\mu\text{m}$ -filtered seawater and further purified through ultracentrifugation (85 852  $\times g$  with a Beckman SW40 Ti rotor for 3 h at 4°C) in a cesium chloride density gradient with layers of 1.2, 1.5, and 1.7  $\text{g ml}^{-1}$ . The viral fractions from the September 2008 samples were further concentrated with a Microcon centrifugal filter device with a pore size of 30 kDa (Millipore). Viral DNA was extracted from all samples using the formamide method, as described by Sambrook et al. (1989).

## Amplification of signature genes

Several signature genes, each corresponding to a specific phage group, were examined in order to bolster the analysis of viral diversity. For signature gene analysis, the extracted DNA was amplified using the strand displacement method of the Illustra GenomiPhi V2 DNA amplification kit (GE Healthcare) according to the manufacturer's instructions.

The *phoH* gene was amplified by PCR using viral *phoH* primers vPhoHf (5'-TGC RGG WAC AGG TAA RAC AT-3') and vPhoHr (5'-TCR CCR CAG AAA AYM ATT TT-3') (Goldsmith et al. 2011). The 50  $\mu\text{l}$  reaction mixture contained 1 U Apex *Taq* DNA polymerase (Genesee Scientific), 1  $\times$  Apex reaction buffer, 1.5 mM Apex  $\text{MgCl}_2$ , 0.5  $\mu\text{M}$  concentration of each primer, 0.2 mM deoxynucleoside triphosphates (dNTPs), 0.08% bovine serum albumin, and 1  $\mu\text{l}$  of template DNA. The reaction conditions were as follows: 5 min of initial denaturation at 95°C; 35 cycles of 1 min of denaturation (95°C), 1 min of annealing (53°C), and 1 min of extension (72°C); and 10 min of final extension at 72°C.

The *g23* gene was amplified using primers T4superF (5'-GAY HTI KSI GGI GTI CAR CCI ATG-3') and T4superR (5'-GCI YKI ARR TCY TGI GCI ARY TC-3') (Chow & Fuhrman 2012). The 50 µl PCR mixture contained 1 U Apex *Taq* DNA polymerase, 1 × Apex reaction buffer, 1.5 mM Apex MgCl<sub>2</sub>, 1.0 µM concentration of each primer, 0.2 mM dNTPs, and 1 µl of template DNA. The reaction conditions were as follows: 5 min of initial denaturation at 94°C; 35 cycles of 1 min of denaturation (94°C), 1 min of annealing (58°C), and 1 min of extension (72°C); and 10 min of final extension at 72°C.

The ssDNA virus major capsid gene was amplified using primers MCPf (5'-CCY KGK YYN CAR AAA GG-3') and MCPr (5'-AHC KYT CYT GRT ADC C-3') (Hopkins et al. 2014). The 50 µl PCR mixture contained 1 U Apex *Taq* DNA polymerase, 1 × Apex reaction buffer, 0.5 µM of each primer, 0.2 mM dNTPs and 1 µl of template DNA. The touchdown PCR conditions were as follows: 3 min of initial denaturation at 94°C; 32 cycles of 1 min of denaturation (95°C), 45 s of annealing (47°C with a 0.1°C decrease/cycle), and 1.5 min of extension (72°C); and 10 min of final extension at 72°C.

### Cloning and sequencing of signature genes

PCR products were cleaned using the UltraClean PCR Clean-Up Kit (MO BIO) following the manufacturer's instructions. After tailing with RED *Taq* DNA polymerase (Sigma-Aldrich), PCR products were cloned into vectors using the TOPO TA cloning kit for sequencing (Invitrogen) and then used to transform competent cells. The cells were then screened by M13 PCR, and the inserts in positive transformants were sequenced with the M13F primer by Beckman Genomics.

### Data analysis for signature genes

Vector and low-quality sequences were trimmed with Sequencher 4.7 (Gene Codes). The sequences were dereplicated into operational taxonomic units (OTUs) using FastGroupII at a level of 98% sequence identity with gaps (Yu et al. 2006). FastGroup representative sequences and reference sequences were aligned at the amino acid level with MUSCLE (Edgar 2004) using the default parameters as implemented by TranslatorX (Abascal et al. 2010). The alignments were cleaned with Gblocks (as implemented by TranslatorX) using the options for a less stringent selection (Castre-

sana 2000, Talavera & Castresana 2007). Back-translated nucleotide alignments were used to build maximum-likelihood phylogenetic trees with FastTree version 2.1 (Price et al. 2010) and branch supports were calculated using the Shimodaira-Hasegawa-like approximate likelihood ratio test on 1000 resamplings. Branches with support below 50 were collapsed using TreeCollapseCL 4 (Hodcroft 2013).

### RAPD PCR amplification and data analysis

As opposed to primers designed to capture specific genes, randomly amplified polymorphic DNA (RAPD) PCR provides a fingerprint of the viral community by using a random decamer primer (Winter & Weinbauer 2010). The RAPD PCR in this study utilized primer SP2 (5'-CGC AAC AGG G-3'), which was designed by Shawn Polson by identifying the most common decamers in the 80 m viral metagenome from the Sargasso Sea (Angly et al. 2006, Srinivasiah et al. 2013). Each 50 µl reaction contained 2 U of Apex *Taq* DNA polymerase (Genesee Scientific), 1 × Apex reaction buffer, 1.5 mM Apex MgCl<sub>2</sub>, 4 µM of primer, 0.2 mM dNTPs, and 1 µl of template DNA. Reaction conditions for the RAPD PCR were as follows: 10 min of initial denaturation at 94°C; 35 cycles of 3 min of annealing at 35°C, 1 min of extension at 72°C, and 30 s of denaturation at 94°C; 3 min of annealing at 35°C; and 10 min of final extension at 72°C. PCR products were cleaned using the UltraClean PCR Clean-Up Kit (MO BIO) following the manufacturer's instructions. The RAPD PCR viral community fingerprints were visualized on the Agilent 2100 Bioanalyzer (Agilent Technologies) following the manufacturer's instructions. Band sizes for RAPD PCR fragments, as determined by the Bioanalyzer, were then binned into OTUs using binning scripts (Ramette 2009). Based on the results of the automatic binner script, a window size of 10 with a shift value of 0.1 was chosen for use in the interactive binner script to determine the binning frame that resulted in highest similarity among the samples.

The nucleotide sequences reported in this paper were submitted to GenBank and assigned accession numbers KP783514 through KP784268.

### Collection of environmental data

Metadata was obtained through the BATS program (Michaels & Knap 1996) (<http://bats.bios.edu>; Table 1). The BATS monthly sampling is considered



synoptic although it extends over 3 to 5 d. Temperature, salinity, and water density were measured on the same cast that provided the water that was analyzed for 10 of the 12 samples. The other 2 samples (September 2008) are associated with environmental data measured 8 d after the water for the samples was collected. Measurements of oxygen and chlorophyll *a* concentrations, as well as *Prochlorococcus* and *Synechococcus* abundances, were used to establish biological and environmental context. Except for chlorophyll *a* in March 2010, all parameters used in this comparison were collected within 36 h of each other to produce a composite profile for each sampling period. No pigments were measured during the March 2010 cruise, so the chlorophyll *a* measurements for those samples were drawn from the next BATS cruise that measured pigments, 16 d later. Similarly, neither *Prochlorococcus* nor *Synechococcus* was measured during the March 2008 cruise, so those measurements were taken from a BATS cruise 11 d earlier.

### Statistical ordination of samples

Sample OTU abundance for each signature gene and RAPD PCR analysis were compared using non-metric multidimensional scaling (NMDS) performed using the 'metaMDS' function (default parameters) of the vegan package (Oksanen et al. 2013) in R version 2.15.2 (R Development Core Team 2013). The influence of environmental variables on sample ordination was evaluated using the functions 'envfit' to assess the significance of factor variables (including depth, month, and year) and 'ordisurf' to assess the significance of the fit of response surfaces (for all linear variables) in the vegan package (Wood 2011, Oksanen et al. 2013) using default parameters with 9999 permutations. Analysis of the ssDNA MCP signature gene showed that one sample (surface from September 2010) did not share any OTUs with any other sample. No ssDNA MCP sequences from the surface during March 2010 were available for analysis because PCR of that sample did not produce any bands. These samples were thus removed from the ordination analyses for ssDNA MCP OTU abundances.

## RESULTS AND DISCUSSION

During the March sampling, MLD was >150 m for all years of the study, thus representing a period with

a well-mixed upper water column (encompassing both the surface and 100 m sampling depths). In contrast, during each of the September sampling events, the water column was comparatively more stratified with MLD  $\leq 33$  m and the 100 m sample originating from well below the mixed layer (Table 1).

NMDS ordination of 3 signature genes encompassing varied groups of phage, as well as RAPD fingerprinting analysis using an arbitrary decamer primer known to occur frequently at this site, all support the hypothesis that the portions of the surface and 100 m viral communities examined here were much more similar when the water column was mixed during March than during stratification in September (Fig. 1). While depth and month of sampling were only significant for individual data sets (*phoH* and ssDNA MCP, respectively), water density was a significant driver of viral community variability based on analysis of all 3 signature genes as well as RAPD PCR fingerprinting (Table 2; Figs. S1 to S4 in the Supplement at [www.int-res.com/articles/suppl/a076p085\\_supp.pdf](http://www.int-res.com/articles/suppl/a076p085_supp.pdf)). Further, September surface communities were significantly different from all other communities for each data set (Fig. 1;  $p < 0.05$  for all). This suggests that water column mixing in March resulted in greater similarity between the viral communities at the surface and 100 m, while the thermal stratification of the water column in September resulted in divergence among these communities. The fact that all 4 methods generated congruent results is striking because each method potentially captures a different subset of the viral community and not all genes (or the phage which encode them, which infect many different hosts) are expected to exhibit the same dynamics in the environment. The robust seasonal variation in composition of the viral community examined here thus suggests that stratification of the water column is a dominant driving factor controlling the vertical distribution of viral communities at the BATS site.

Based on the *g23* and *phoH* genes, as well as RAPD PCR fingerprinting, viral community structure was significantly correlated with temperature, water density, and chlorophyll *a* concentration (Table 2; Figs. S1 to S3 in the Supplement), which are all indicators of water column stability. The viral communities containing the *g23* and *phoH* genes were also significantly correlated with salinity (Table 2; Figs. S1 & S2 in the Supplement). The similar patterns exhibited by communities containing *g23* and *phoH* may be due to the overlap in those communities; *g23* is a core gene in T4-like phage, while *phoH* is a core gene in T4-like cyanophage (Sullivan et al. 2010, Ignacio Espinoza

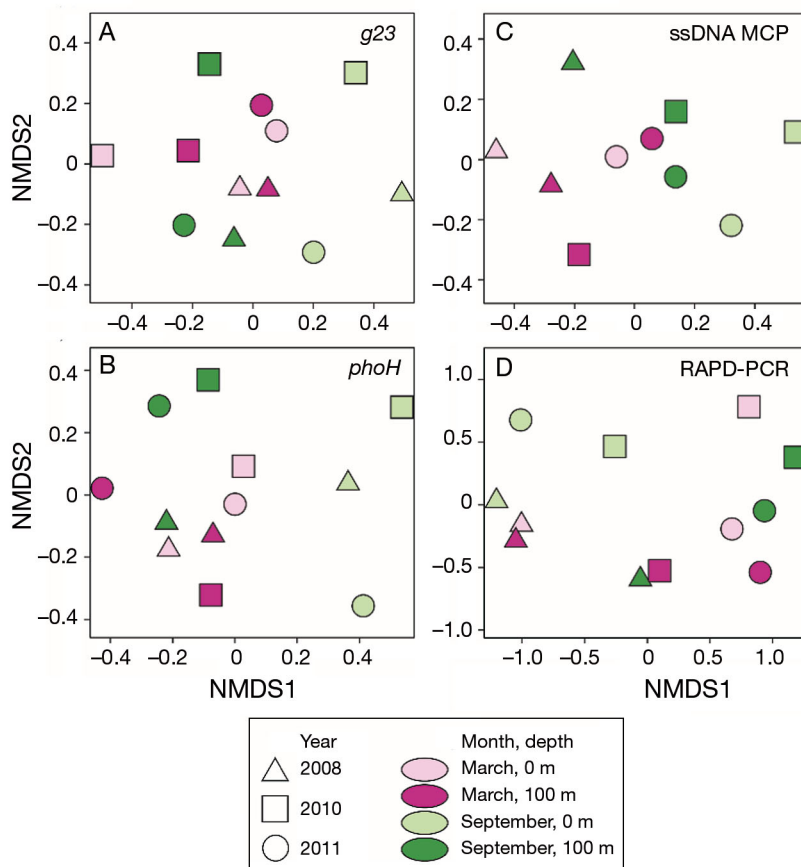


Fig. 1. Viral community composition in the Sargasso Sea: NMDS sample ordination based on OTU abundance of signature genes (A) *g23* (stress = 16.9%), (B) *phoH* (stress = 17.7%) and (C) ssDNA MCP (stress = 5.2%), and (D) RAPD-PCR gel bands (stress = 8.4%). The ssDNA MCP ordination does not include the 0 m sample from September 2008 because it did not share any OTUs with other samples, nor the 0 m sample from March 2010 because the PCR of that sample did not yield any bands

Table 2. Relationships between environmental metadata and viral community composition in the Sargasso Sea based on signature genes and RAPD-PCR. The influence of variables on NMDS sample ordination was tested using factor analysis (for factor variables) and the fit of response surfaces (for linear variables). Significant relationships are in bold ( $\alpha = 0.05$ ). Only 10 samples were used for the ssDNA MCP gene due to an outlier sample (0 m from September 2008) and a PCR that did not yield any bands (0 m from March 2010)

Variable	<i>g23</i> (n = 12)	<i>phoH</i> (n = 12)	ssDNA MCP (n = 10)	RAPD-PCR (n = 12)
Depth	p = 0.434	<b>p = 0.033</b>	p = 0.682	p = 0.130
Month	p = 0.324	p = 0.128	<b>p = 0.040</b>	p = 0.694
Year	p = 0.248	p = 0.785	p = 0.108	p = 0.074
Temperature	<b>p = 0.014</b>	<b>p &lt; 0.001</b>	p = 0.996	<b>p = 0.028</b>
Salinity	<b>p &lt; 0.001</b>	<b>p = 0.011</b>	p = 0.187	p = 0.095
Water density	<b>p = 0.007</b>	<b>p &lt; 0.001</b>	<b>p &lt; 0.001</b>	<b>p = 0.033</b>
Chlorophyll <i>a</i>	<b>p = 0.026</b>	<b>p = 0.022</b>	p = 0.489	<b>p = 0.018</b>
Oxygen	p = 0.091	p = 0.189	<b>p &lt; 0.001</b>	p = 0.576
<i>Prochlorococcus</i>	p = 0.393	p = 0.767	p = 0.872	p = 0.832
<i>Synechococcus</i>	p = 0.948	p = 0.609	p = 0.451	p = 0.565

& Sullivan 2012). In contrast, the composition of viruses with the ssDNA MCP gene was correlated only with water density and oxygen concentration (Table 2; Fig. S4 in the Supplement), indicating that water column stratification is still important in structuring the ssDNA viral community, but chlorophyll *a* concentration is not. Discrepancies may also result from the lag between sampling and metadata for a few environmental parameters and some of the samples. That fewer environmental variables are influential for the ssDNA viral community may be related to the lower statistical power in this analysis, as the PCR of one sample did not yield any bands (the surface sample from March 2010) and another sample did not share any OTUs with any other samples (the surface sample from September 2008) and needed to be removed from the analysis.

As discussed above, the NMDS-based analyses revealed that water density was a significant driver of viral community structure as determined by all 4 methods for viral community analysis. This is consistent with prior studies, which have demonstrated that physical processes drive bacterial dynamics at the BATS site (Carlson et al. 1996, Steinberg et al. 2001, Nelson et al. 2004, Palter et al. 2005, Lozier et al. 2011). During mixing periods, bacterioplankton densities are redistributed homogeneously within the mixed layer (Carlson et al. 1996, Steinberg et al. 2001). The distribution of bacterioplankton hosts and their activities in turn drive the vertical distribution of viral abundance (Parsons et al. 2012). Further, the repeatable seasonal signature in viral community composition at the BATS site based on signature genes and RAPD PCR is consistent with seasonality observed in the San Pedro Ocean Time-series evaluation of T4-like viral community composition based on the *g23* gene (Chow & Fuhrman 2012, Needham et al. 2013, Chow et al. 2014).

All viral community analyses indicated that the September surface samples were significantly different from the other communities at the BATS site. This difference from the other viral assemblages may stem from the vertical structuring of the host bacterioplankton community as the water column stratifies (Morris et al. 2005, Carlson et al. 2009, Treusch et al. 2009, Treusch et al. 2012). As the mixed layer begins to shoal after the winter deep mixing, the bacterial communities residing near the surface begin to change to distinct communities compared to those observed in the surface during vigorous mixing periods (Morris et al. 2005, Carlson et al. 2009, Treusch et al. 2009, 2012). Under this model, as the surface bacterial community changes, so would its associated viral community. Another potential explanation for differences between the September surface viral communities and all other viral communities is the greater influence of ultraviolet light at the surface during that season (Morrison & Nelson 2004). Ultraviolet radiation degrades viral particles (Wommack et al. 1996, Weinbauer & Suttle 1999, Araújo & Godinho 2009) and could thereby directly influence the observed divergence of viral communities within the stratified layer of the water column in September. Other studies have also reported variability in the degree to which different viruses respond to UV radiation (Moebus 1992, Suttle & Chen 1992, Noble & Fuhrman 1997, Wilhelm et al. 1998, Jacquet & Bratbak 2003).

While *Prochlorococcus* abundance was significantly correlated with viral abundance during a 10-yr study of the BATS site (Parsons et al. 2012), neither *Prochlorococcus* nor *Synechococcus* abundance was significantly related to viral community composition based on any of the signature genes or RAPD PCR (Table 2). This is most likely an indication that cyanobacterial community composition, and not abundance, is responsible for observed changes in cyanophage community composition (Rocap et al. 2002, Mühling et al. 2005, Zwirgmaier et al. 2008, Martiny et al. 2009). Maximum *Prochlorococcus* densities (on the order of  $10^5$  cells  $\text{ml}^{-1}$ ) occur near 60 to 80 m depth in summer and fall, and high densities persist to nearly 200 m (DuRand et al. 2001). Viruses that infect *Prochlorococcus* are likely more abundant in the depth horizon of the host abundance maximum. Therefore, if a significant portion of the viral community consists of phage that infect *Prochlorococcus*, as suggested by Parsons et al. (2012), the differential in *Prochlorococcus* densities between the surface and 100 m would account for some of the difference between the surface and 100 m viral assemblages,

since surface *Prochlorococcus* densities are often lower than they are at 100 m in late summer (DuRand et al. 2001, Malmstrom et al. 2010, Parsons et al. 2012). Prior research has shown that cyanomyophage, whose hosts include *Prochlorococcus*, are an important component of the viral community in the Sargasso Sea; they have been found to constitute up to 25% of the viral community at the surface during a summer transect (Matteson et al. 2013). Moreover, based on the *phoH* phylogenetic tree, we can reasonably infer that cyanophage are a substantial component of our sampled viral communities because many of the sequences from our sampled communities cluster in the groups with fully-sequenced cyanophage (Fig. S5 in the Supplement).

In contrast to *Prochlorococcus*, for which phage infection has been well-studied and numerous phage genomes have been sequenced, much less is known about phage infecting other members of the bacterial community at the BATS site. The most abundant heterotrophic bacteria at this site belong to the SAR11 clade (Morris et al. 2002), for which no phage were known until recently (Zhao et al. 2013). BLAST searches of oceanic viromes (including the Sargasso Sea) with these 'pelagiphage' genome sequences (named after their host, *Candidatus Pelagibacter ubique*) suggested that pelagiphage were the most abundant identifiable viruses in oceanic samples. All 4 of the cultured pelagiphage have dsDNA genomes, and only one of them (HTVC008M) contains one of the signature genes used in this study (*phoH*), so it is possible that pelagiphage were not well-covered in the analyses of this study even though they may have been recovered by RAPD PCR. In addition, it is known that SAR11 diversity is highly structured vertically at BATS (Carlson et al. 2009, Vergin et al. 2013); thus, additional tools are needed for investigating the phage of the dominant heterotrophs at this site. Finally, there is likely a high diversity of uncultured heterotrophic phage at this site, so some of the sequences generated in this study may in fact represent yet undescribed phage infecting SAR11 or other abundant heterotrophs, such as *Roseobacter*, SAR116, and OCS116 (all clades within the *Alphaproteobacteria*) and SAR86 (belonging to the *Gammaproteobacteria*).

Even less well-described than the pelagiphage are the ssDNA marine gokushoviruses, for which no cultured representatives exist. Gokushoviruses (family *Microviridae*, subfamily *Gokushovirinae*) were first documented in the marine environment by a metagenomic study in the Sargasso Sea and have recently been shown to be widespread in the oceans

(Angly et al. 2006, Tucker et al. 2011, Labonté & Suttle 2013, Hopkins et al. 2014). The hosts for the marine gokushoviruses remain almost entirely unknown, with the exception of a recent report of gokushoviruses in single-cell amplified genomes from uncultured SUP05 sulfur-oxidizing *Gammaproteobacteria* from an oxygen minimum zone in a seasonally anoxic fjord in British Columbia (Roux et al. 2014). The present study demonstrated that ssDNA phage at the BATS site are quite diverse and display temporal patterns in community composition similar to those of dsDNA viruses. As a first step towards deciphering the ecological impact of ssDNA phage, future studies should attempt to link the ssDNA phage with their potential host communities, a task that will be greatly aided by emerging single-cell methods such as phageFISH and viral tagging (Deng et al. 2012, Allers et al. 2013, Dang & Sullivan 2014, Brum & Sullivan 2015).

This study's goal was to obtain an overview of the viral communities over 2 depths, 2 seasons, and several years, by using 4 different methods. The data presented in this study should be viewed as a representation of the members of the viral community containing each signature gene (or the decamer sequence used for RAPD PCR), as opposed to an exhaustive analysis. Since the same primers were used throughout this study, the results from different samples can be directly compared. However, it is possible that the approaches used in this study do not accurately reflect the dominant members of the viral communities, which could lack the targeted signature gene sequences. Metagenomic sequencing of the viral communities is one possible avenue for overcoming these limitations (Breitbart et al. 2002, Angly et al. 2006, Brum et al. 2015, Sullivan 2015). In addition, examining the factors driving phage and host dynamics at the BATS site would benefit from better representation of environmental phage genomes in sequence databases. The signature gene phylogenetic trees (Figs. S5 to S7 in the Supplement) reveal that most of the environmental phage fall in clusters without any cultured or sequenced representatives, highlighting the gap that remains between cultured phage genomes and the dominant phage in the environment. Future work at the BATS site should also target the subsurface viral maximum depth as opposed to the static 100 m depth in order to better understand seasonal variability. Moreover, concurrent tracking of virus and host dynamics could better resolve the factors connected with the recurring seasonal patterns in viral community composition revealed here.

## CONCLUSION

Four data sets revealing the composition of portions of the viral community at the BATS site over 3 years, 2 seasons, and 2 depths supported the hypotheses that in the well-mixed period (March), viral communities at the surface and 100 m depth resemble each other, while during water column stratification (September), the viral community composition becomes vertically structured. Water density was a significant driver of viral community structure as measured by all 3 signature genes and RAPD PCR, while additional factors (temperature, salinity, chlorophyll *a*, and oxygen) were correlated with some but not all of the community analyses. Therefore, the physical factors that control water column stratification and distribution of bacterioplankton hosts are the major forces underlying viral dynamics at this well-studied oceanographic site.

*Acknowledgements.* Thanks to Camille Daniels, Darren Dunlap, Rachel Parsons, Karyna Rosario, Kimberly Pause Tucker, and the crew of the RV 'Atlantic Explorer' for logistical support with sample collection. Shawn Polson designed the RAPD PCR primers, and André Comeau designed the *g23* primers. This research was funded by grants to M.B. from the National Science Foundation (Microbial Interactions and Processes MCB-0701984 and Division of Biological Infrastructure DBI-0850206) and a Proposal Enhancement Grant from the University of South Florida. D.B.G. was supported by a Presidential Doctoral Fellowship from the University of South Florida and the Von Rosenstiel Endowed Fellowship.

## LITERATURE CITED

- Abascal F, Zardoya R, Telford M (2010) TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. *Nucleic Acids Res* 38:W7–W13
- Allers E, Moraru C, Duhaime MB, Beneze E and others (2013) Single-cell and population level viral infection dynamics revealed by phageFISH, a method to visualize intracellular and free viruses. *Environ Microbiol* 15: 2306–2318
- Angly F, Felts B, Breitbart M, Salamon P and others (2006) The marine viromes of four oceanic regions. *PLoS Biol* 4: e368
- Araújo MFF, Godinho MJL (2009) Short-term variations of virus-like particles in a tropical lake: relationship with microbial communities (bacteria, ciliates and flagellates). *Microbiol Res* 164:411–419
- Breitbart M (2012) Marine viruses: truth or dare. *Annu Rev Mar Sci* 4:425–448
- Breitbart M, Salamon P, Andresen B, Mahaffy JM and others (2002) Genomic analysis of uncultured marine viral communities. *Proc Natl Acad Sci USA* 99: 14250–14255
- Brum JR, Sullivan MB (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol* 13:147–159



- Brum JR, Ignacio-Espinoza JC, Roux S, Doucier G and others (2015) Patterns and ecological drivers of ocean viral communities. *Science* 348:1261498
- Carlson CA, Ducklow HW, Sleeter TD (1996) Stocks and dynamics of bacterioplankton in the northwestern Sargasso Sea. *Deep-Sea Res II* 43:491–515
- Carlson CA, Morris R, Parsons R, Treusch AH, Giovannoni SJ, Vergin K (2009) Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J* 3:283–295
- Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* 17:540–552
- Chow CET, Fuhrman JA (2012) Seasonality and monthly dynamics of marine myovirus communities. *Environ Microbiol* 14:2171–2183
- Chow CET, Kim DY, Sachdeva R, Caron DA, Fuhrman JA (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J* 8:816–829
- Comeau AM, Krisch HM (2008) The capsid of the T4 phage superfamily: the evolution, diversity, and structure of some of the most prevalent proteins in the biosphere. *Mol Biol Evol* 25:1321–1332
- Dang VT, Sullivan MB (2014) Emerging methods to study bacteriophage infection at the single-cell level. *Front Microbiol* 5:724
- Deng L, Gregory A, Yilmaz S, Poulos BT, Hugenholtz P, Sullivan MB (2012) Contrasting life strategies of viruses that infect photo-and heterotrophic bacteria, as revealed by viral tagging. *MBio* 3:e00373-12
- DuRand M, Olson R, Chisholm S (2001) Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep-Sea Res II* 48:1983–2003
- Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32:1792–1797
- Filée J, Tétart F, Suttle C, Krisch H (2005) Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proc Natl Acad Sci USA* 102:12471–12476
- Goldsmith DB, Crosti G, Dwivedi B, McDaniel LD and others (2011) Development of *phoH* as a novel signature gene for assessing marine phage diversity. *Appl Environ Microbiol* 77:7730–7739
- Hodcroft E (2013) TreeCollapserCL 4. <http://emmahodcroft.com/TreeCollapseCL.html>
- Hopkins M, Kailasan S, Cohen A, Roux S and others (2014) Diversity of environmental single-stranded DNA phages revealed by PCR amplification of the partial major capsid protein. *ISME J* 8:2093–2103
- Ignacio-Espinoza JC, Sullivan MB (2012) Phylogenomics of T4 cyanophages: lateral gene transfer in the ‘core’ and origins of host genes. *Environ Microbiol* 14:2113–2126
- Jacquet S, Bratbak G (2003) Effects of ultraviolet radiation on marine virus-phytoplankton interactions. *FEMS Microbiol Ecol* 44:279–289
- Labonté JM, Suttle CA (2013) Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J* 7:2169–2177
- Lomas M, Bates N, Johnson R, Knap A, Steinberg D, Carlson C (2013) Two decades and counting: 24-years of sustained open ocean biogeochemical measurements in the Sargasso Sea. *Deep-Sea Res II* 93:16–32
- Lozier MS, Dave AC, Palter JB, Gerber LM, Barber RT (2011) On the relationship between stratification and primary productivity in the North Atlantic. *Geophys Res Lett* 38, L18609, doi:10.1029/2011GL049414
- Malmstrom RR, Coe A, Kettler GC, Martiny AC, Frias-Lopez J, Zinser ER, Chisholm SW (2010) Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific oceans. *ISME J* 4:1252–1264
- Martiny AC, Tai AP, Veneziano D, Primeau F, Chisholm SW (2009) Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ Microbiol* 11:823–832
- Matteson AR, Rowe JM, Ponsoero AJ, Pimentel TM, Boyd PW, Wilhelm SW (2013) High abundances of cyanomyoviruses in marine ecosystems demonstrate ecological relevance. *FEMS Microbiol Ecol* 84:223–234
- Michaels A, Knap A (1996) Overview of the US JGOFS Bermuda Atlantic Time-series Study and the Hydrostation S program. *Deep-Sea Res II* 43:157–198
- Moebus K (1992) Laboratory investigations on the survival of marine bacteriophages in raw and treated seawater. *Helgolander Meeresun* 46:251–273
- Morris RM, Rappé MS, Connon SA, Vergin KL, Siebold WA, Carlson CA, Giovannoni SJ (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420:806–810
- Morris R, Vergin K, Cho J, Rappé M, Carlson C, Giovannoni S (2005) Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. *Limnol Oceanogr* 50:1687–1696
- Morrison JR, Nelson NB (2004) Seasonal cycle of phytoplankton UV absorption at the Bermuda Atlantic Time-series Study (BATS) site. *Limnol Oceanogr* 49:215–224
- Mühling M, Fuller N, Millard A, Somerfield P and others (2005) Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. *Environ Microbiol* 7:499–508
- Needham DM, Chow CET, Cram JA, Sachdeva R, Parada A, Fuhrman JA (2013) Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *ISME J* 7:1274–1285
- Nelson N, Siegel D, Yoder J (2004) The spring bloom in the northwestern Sargasso Sea: spatial extent and relationship with winter mixing. *Deep-Sea Res II* 51:987–1000
- Noble RT, Fuhrman J (1997) Virus decay and its causes in coastal waters. *Appl Environ Microbiol* 63:77–83
- Oksanen J, Blanchet FG, Kindt R, Legendre P and others (2013) vegan: community ecology package. <http://cran.r-project.org>
- Palter JB, Lozier MS, Barber RT (2005) The effect of advection on the nutrient reservoir in the North Atlantic subtropical gyre. *Nature* 437:687–692
- Parsons RJ, Breitbart M, Lomas MW, Carlson CA (2012) Ocean time-series reveals recurring seasonal patterns of viroplankton dynamics in the northwestern Sargasso Sea. *ISME J* 6:273–284
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490
- R Development Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna. [www.r-project.org](http://www.r-project.org)
- Ramette A (2009) Quantitative community fingerprinting methods for estimating the abundance of operational

- taxonomic units in natural microbial communities. *Appl Environ Microbiol* 75:2495–2505
- 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
  - Roux S, Enault F, Robin A, Ravet V and others (2012) Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS ONE* 7:e33641
  - Roux S, Hawley AK, Beltran MT, Scofield M and others (2014) Ecology and evolution of viruses infecting uncultivated SUP05 bacteria as revealed by single-cell- and meta-genomics. *eLife* 3:e03125
  - Sambrook J, Fritsch E, Maniatis T (1989) *Molecular cloning: a laboratory manual*, Vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
  - Siegel DA, McGillicuddy DJ, Fields EA (1999) Mesoscale eddies, satellite altimetry, and new production in the Sargasso Sea. *J Geophys Res-Oceans* 104:13359–13379
  - Sprintall J, Tomczak M (1992) Evidence of the barrier layer in the surface layer of the tropics. *J Geophys Res-Oceans* 97:7305–7316
  - Srinivasiah S, Lovett J, Polson S, Bhavsar J and others (2013) Direct assessment of viral diversity in soils by random PCR amplification of polymorphic DNA. *Appl Environ Microbiol* 79:5450–5457
  - Steinberg DK, Carlson CA, Bates NR, Johnson RJ, Michaels AF, Knap AH (2001) Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep-Sea Res II* 48:1405–1447
  - Sullivan MB (2015) Viromes, not gene markers for studying dsDNA viral communities. *J Virol* 89:2459–2461
  - Sullivan MB, Huang K, Ignacio Espinoza J, Berlin A and others (2010) Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ Microbiol* 12: 3035–3056
  - Suttle CA (2007) Marine viruses — major players in the global ecosystem. *Nat Rev Microbiol* 5:801–812
  - Suttle CA, Chen F (1992) Mechanisms and rates of decay of marine viruses in seawater. *Appl Environ Microbiol* 58: 3721–3729
  - Talavera G, Castresana J (2007) Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst Biol* 56: 564–577
  - Thurber RV, Haynes M, Breitbart M, Wegley L, Rohwer F (2009) Laboratory procedures to generate viral metagenomes. *Nat Protoc* 4:470–483
  - Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA, Giovannoni SJ (2009) Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J* 3:1148–1163
  - Treusch AH, Demir-Hilton E, Vergin KL, Worden AZ and others (2012) Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J* 6:481–492
  - Tucker KP, Parsons R, Symonds EM, Breitbart M (2011) Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. *ISME J* 5: 822–830
  - Vergin KL, Beszteri B, Monier A, Thrash JC and others (2013) High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J* 7:1322–1332
  - Weinbauer MG, Suttle CA (1999) Lysogeny and prophage induction in coastal and offshore bacterial communities. *Aquat Microb Ecol* 18:217–225
  - Wilhelm SW, Weinbauer MG, Suttle CA, Jeffrey WH (1998) The role of sunlight in the removal and repair of viruses in the sea. *Limnol Oceanogr* 43:586–592
  - Winter C, Weinbauer MG (2010) Randomly amplified polymorphic DNA reveals tight links between viruses and microbes in the bathypelagic zone of the northwestern Mediterranean Sea. *Appl Environ Microbiol* 76: 6724–6732
  - Wommack KE, Hill RT, Muller TA, Colwell RR (1996) Effects of sunlight on bacteriophage viability and structure. *Appl Environ Microbiol* 62:1336–1341
  - Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc Series B Stat Methodol* 73:3–36
  - Yu Y, Breitbart M, McNairnie P, Rohwer F (2006) Fast-GroupII: a web-based bioinformatics platform for analyses of large 16S rDNA libraries. *BMC Bioinformatics* 7:57
  - Zhao Y, Temperton B, Thrash JC, Schwalbach MS and others (2013) Abundant SAR11 viruses in the ocean. *Nature* 494:357–360
  - Zwirgmaier K, Jardillier L, Ostrowski M, Mazard S and others (2008) Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol* 10:147–161

Editorial responsibility: Jed Fuhrman,  
Los Angeles, California, USA

Submitted: April 7, 2015; Accepted: July 28, 2015  
Proofs received from author(s): September 13, 2015