ORIGINAL ARTICLE

Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists

Cheryl-Emiliane T Chow, Diane Y Kim, Rohan Sachdeva, David A Caron and Jed A Fuhrman

Department of Biological Sciences, University of Southern California, Los Angeles, CA, USA

Characterizing ecological relationships between viruses, bacteria and protists in the ocean are critical to understanding ecosystem function, yet these relationships are infrequently investigated together. We evaluated these relationships through microbial association network analysis of samples collected approximately monthly from March 2008 to January 2011 in the surface ocean (0–5 m) at the San Pedro Ocean Time series station. Bacterial, T4-like myoviral and protistan communities were described by Automated Ribosomal Intergenic Spacer Analysis and terminal restriction fragment length polymorphism of the gene encoding the major capsid protein (g23) and 18S ribosomal DNA, respectively. Concurrent shifts in community structure suggested similar timing of responses to environmental and biological parameters. We linked T4-like myoviral, bacterial and protistan operational taxonomic units by local similarity correlations, which were then visualized as association networks. Network links (correlations) potentially represent synergistic and antagonistic relationships such as viral lysis, grazing, competition or other interactions. We found that virus–bacteria relationships were more cross-linked than protist–bacteria relationships, suggestive of increased taxonomic specificity in virus–bacteria relationships. We also found that 80% of bacterial–protist and 74% of bacterial–viral correlations were positive, with the latter suggesting that at monthly and seasonal timescales, viruses may be following their hosts more often than controlling host abundance.

Introduction

Bacterial activity in the ocean is a key driver of biogeochemical cycles; this activity is mediated by bottom-up controls (for example, resource availability and competition), top-down controls (for example, predation and viral lysis) and also bacteria–bacteria interactions (for example, allelopathy or living in consortia). The microbial loop thus links bacteria, protists and viruses, creating a complex microbial community where the bacteria consume organic carbon produced by other organisms following natural death, grazing by protists or viral infection (Azam *et al.*, 1983; Sherr and Sherr, 1988; Fuhrman and Suttle, 1993; Bratbak *et al.*, 1994; Fuhrman, 1999). The dominant top-down controls, or sources of bacterial mortality, in the open ocean are thought to be viral lysis and

protistan grazing. Although widely accepted as important, the relative contribution of each control remains a subject of debate and no doubt differs based on location, time, physiological status or identity of the bacteria.

Many studies have sought to quantify grazing and viral lysis to determine the impact of top-down or bottom-up controls on structuring microbial communities (Fuhrman and Noble, 1995; Strom, 2000; Simek *et al.*, 2001; Sherr and Sherr, 2002; Evans *et al.*, 2003; Weinbauer *et al.*, 2003, 2007; Zhang *et al.*, 2007; Baudoux and Veldhuis, 2008; Longnecker *et al.*, 2010; Staniewski *et al.*, 2012). Most studies enriched or removed grazers and viruses to investigate short-term or episodic impacts from which long-term influences were inferred. For example, reduction in grazer activity affected bacterial diversity of active cells, but removal of viruses only affected rates of activity and not diversity of the active cells (Longnecker *et al.*, 2010). Another study similarly suggested that viruses may help control the abundance of rare organisms through selective mortality (Bouvier and del Giorgio, 2007), although the net influence of

viruses on bacterial communities has been reportedly mixed (Schwalbach et al., 2004; Hewson and Fuhrman, 2006). Bottom-up controls also affect bacterial diversity yet the net effects may still depend upon viral or protistan activity (Moebus, 1996; Middelboe, 2000; Gasol et al., 2002; Corno and Jürgens, 2008; Sandaa et al., 2009; Ory et al., 2010; Bouvy et al., 2011). These investigations collectively revealed close couplings between viruses, bacteria and protists, but questions remain in our understanding of how 'top' communities of viruses and protists affect bacteria at natural concentrations over long timescales.

Ecological networks of trophic interactions have historically been used to characterize complex food webs by the positive and negative interactions within (Sole and Montoya, 2001; Dunne, 2002; Montoya et al., 2006; Olesen et al., 2011). Network analysis has only recently been applied to microbes, and our ability to interpret these networks is still under development (Fuhrman and Steele, 2008; Chaffron et al., 2010; Steele et al., 2011; Eiler et al., 2012; Gilbert et al., 2012). Positive correlations may suggest co-occurrence due to (1) similar preferred conditions, (2) commensalism or (3) a mutualistic relationship between organisms cooperating within the same niche. Indirect relationships, where a third party benefits from an interaction between two others, are potentially common (Miki and Jacquet, 2008, 2010) and may also be detected as a correlated set of three or more. Indirect relationships may also appear as a single correlated pair if the indirect partner is rare or only weakly correlated. Negative correlations, or time-lagged positive ones, may suggest the presence of predation (protist–protist, protist–bacteria and bacteria–bacteria), viral lysis (virus–bacteria) and competition (any two taxa). Time-shifted correlations between viruses and bacteria could represent a succession of taxa perhaps resulting from a lysis event, whereas correlations of bacteria to environmental, viral or protistan parameters could indicate to what extent the environment, viral pressure or grazing activity drives bacterial abundance and community structure.

Here, we queried a seasonally variable, semi-oligotrophic, surface ocean bacterial community monthly over 3 years to determine the links between protistan, viral and environmental factors using culture-independent community fingerprinting methods, community similarity metrics, local similarity analysis (LSA) and construction of association networks. Past research at the San Pedro Ocean Time series (SPOT) has investigated correlations among the smallest plankton, specifically bacteria, protists and archaea over time (Fuhrman et al., 2006; Fuhrman and Steele, 2008; Steele et al., 2011). In this study, association networks (built from LS correlations between individual bacterial, protistan, T4-like myoviral and environmental parameters, including possible time lags) revealed

clusters of operational taxonomic units (OTUs) that likely reflect ecologically relevant interactions. We focused specifically on the T4-like myovirus family in lieu of the entire viral community; T4-like viruses are diverse, abundant, detectable through cultivation-independent methods and include bacteriophages of marine cyanobacteria and SAR11/ *Pelagibacter* (Filée et al., 2005; Comeau and Krisch, 2008; Clokie et al., 2010; Chow and Fuhrman, 2012; Zhao et al., 2013). Over this time series, our observations on correlated OTUs revealed: (1) many T4-like virus OTUs significantly correlated to individual bacterial OTUs (reflective of bacterial hosts susceptible to multiple viruses), (2) single viral OTUs significantly correlated with multiple bacterial OTUs (suggestive of a virus' capability to infect multiple hosts), (3) protistan OTUs significantly correlated with multiple bacterial OTUs (as evidence of non-selective grazing in the case of a phagotroph or broad bacterial use of nutrients released from phototrophs) and (4) protistan OTUs significantly correlated with a single bacterial OTU or taxonomic group (due to selective grazing or nutrient transfer). Finally, inter-correlated clusters of parameters emerged that detailed potential ecological niches and microbial guilds worth further investigation.

Materials and methods

Sample collection
Seawater (~ 20 l) was collected approximately monthly at 0 or 5 m at the University of Southern California's Microbial Observatory at SPOT (33' 33° N, 118' 24° W) and filtered for free-living protistan (0.7–20 µm), bacterial (0.22–1 µm) and viral (30 kDa–0.22 µm) community DNA from March 2008 to January 2011, as previously described (Countway et al., 2005; Fuhrman et al., 2006; Vigil et al., 2009; Countway et al., 2010; Steele et al., 2011; Chow and Fuhrman, 2012; Kim et al., 2012; Chow et al., 2013). Molecular data was unavailable for October 2008 (virus), January 2009 (all), March 2009 (bacteria), October–November 2009 (bacteria) and January 2011 (protist). Bulk seawater samples were also collected and analyzed for bacterial and viral abundance by SYBR green epifluorescence microscopy, bacterial production by thymidine and leucine incorporation, and nutrient concentrations using colorimetric methods (Chow et al., 2013).

Fingerprinting microbial communities

Bacteria and viruses. Bacterial community composition was determined by Automated Ribosomal Intergenic Spacer Analysis (ARISA) (Fisher and Triplett, 1999; Brown et al., 2005; Chow et al., 2013; Needham et al., 2013). T4-like myovirus communities were analyzed by terminal restriction

818

fragment length polymorphism (TRFLP) of g23, which encodes the major capsid protein (Chow and Fuhrman, 2012). Viral fingerprints were obtained from both terminal fragments (5′ and 3′). ARISA and g23-TRFLP products were run in duplicate on non-adjacent lanes on an ABI377 by slab gel electrophoresis with internal size standards (Bioventures, Murfreesboro, TN, USA) every 25 bp (50–900 bp) or 50 bp (900–1400 bp). Peaks were identified in DAx (van Mierlo, Inc, Eindhoven, The Netherlands). Fragments, 400–1210 bp (ARISA) and 50–500 bp (g23-TRFLP), were rounded to the nearest 0.1 bp and dynamically binned (Ruan, et al., 2006b; Chow and Fuhrman, 2012). The resulting bins were manually curated to merge bins <0.1 bp wide with the nearest neighbor; each assay was binned independently. ARISA OTUs were assigned an identity by matching ARISA lengths with known sequences and their ARISA products (Chow et al., 2013; Needham et al., 2013); terminal fragments from in silico analysis of publicly available T4-like viral genomes were used to assign identities to environmental g23-TRFLP OTUs.

Protists. Dominant taxa within protistan assemblages were characterized by 18S rDNA-based TRFLP using Euk-A and Euk570-R primers for PCR and *Hae*III for digestion (Countway et al., 2005; Vigil et al., 2009); fragments were analyzed on a Beckman CEQ 8000 (Brea, CA, USA). Fragments from *Ostreococcus* sp. and *Phaeocystis globosa* cultures were used as positive controls for calibrating and verifying fragment sizes. Protistan OTUs were identified from *in silico* digestion of 1341 18S rRNA gene sequences from October 2001 at SPOT (Kim et al., 2012).

Peak analysis for all communities. Unique fragment lengths were considered as individual OTUs. Relative abundance of each OTU was calculated by dividing a peak's area by the total area within the monthly fingerprint. Bacterial and viral OTUs <0.1% of the community and protistan OTUs <0.5% of the community were removed from further analysis, and the remaining peaks were normalized by sample to determine relative abundance per month; each community thus totaled to 100%.

Data analysis

Community similarity. Bray–Curtis similarity was determined for each microbial community independently for all monthly pairwise comparisons in PRIMER-E v6 (Clarke and Gorley, 2006). Bray–Curtis resemblance matrices were compared using RELATE (PRIMER-E), a Mantel-type test, with a Spearman correlation and 999 permutations. Correlations of Bray–Curtis similarities for adjacent month comparisons only were calculated by Pearson-product-moment (Sigmaplot11, San Jose, CA, USA).

Co-correspondence analysis and canonical correspondence analysis. Determination of covariance of microbial community data by co-correspondence analysis was completed with *coccoresp* in R (Braak and Schaffers, 2004; Simpson, 2009) on log-transformed relative abundance data for months where all three microbial community data sets were available ($n = 28$). Any OTU present in <5 months was excluded. Significance testing was completed by cross-validation with the 'leave-one-out' method and permutation tests ($n = 99$). Covariance of communities with environmental parameters was determined in R using *cca* with a stepwise model from the vegan package v2.0.2 (Oksanen et al., 2011); models were validated by analysis of variance. Estimates for chlorophyll-*a* concentrations and primary production were downloaded for the grid area surrounding SPOT from National Oceanographic and Atmospheric Administration (NOAA) Coastwatch: (a) SeaWiFS, 0.04167 degrees, West US Science Quality for Chlorophyll-*a* and (b) SeaWiFS and Pathfinder, 0.1 degrees, Global, Experimental data sets for primary productivity (Hooker and McClain, 2000). Environmental data were transformed as follows: log(value) for bacterial production by thymidine and leucine incorporation, calculated turnover time, chlorophyll-*a* (bottle) and satellite-based chlorophyll-a; log(value + 0.01) for NO_2, NO_3 and PO_4; square-root for bacterial and viral abundance and the virus:bacteria ratio; no transformation for salinity, temperature, sea surface height differential, primary production (satellite), day length and monthly change in day length. Missing environmental data were filled with the overall mean of the transformed data; all data were then normalized to a common scale (subtracted means and divided by s.d.) to account for differences in units before completing canonical correspondence analysis analyses.

LSA and network analysis. We determined LS correlations (ranked Pearson's correlations) by LS analysis (eLSA) using a linear interpolation for missing values and a delay up to 1 month (Ruan et al., 2006a; Steele et al., 2011; Xia et al., 2011, 2013). Any OTU or environmental parameter that occurred in <5 months was excluded, resulting in 227 bacterial OTUs, 376 T4-like viral OTUs (3′: 171, 5′: 205), 70 protistan OTUs and 30 environmental parameters. *P*-values were determined using statistical approximation followed by permutation testing to reduce computing time while ensuring accuracy (eLSA option: p_{mix} (Xia et al., 2013)). First, *P*-values were determined for all vs all pairwise relationships using Feller's theoretical approximation based on the approximate tail distribution of the maximum partial sum of independent identically distributed random variables (p_{theo}); second, for any pairwise

relationship with $p_{theo} < 0.05$, the more robust yet more intensive, permutation-based ($n = 2000$) P-value (p_{perm}) was determined. Only local similarity (LS) correlations with q-value < 0.10 and $p_{perm} < 0.0015$ were retained for further analysis. The q-value (or false-discovery rate) was the more stringent criteria and led to the specified P-value cutoff; by employing $q < 0.10$, no more than 10% of all remaining 'statistically significant' LS correlations may be due to error. These remaining LS correlations were visualized in Cytoscape v2.8.2 (Shannon, 2003; Cline *et al.*, 2007; Smoot *et al.*, 2011). Example networks were selected by taxonomic OTU identification (for example, cyanobacteria) or edge type (or example, correlations between specific OTUs). Random undirected networks of equal size by number of nodes and edges were calculated by the Erdős–Rényi model using the Random Network plugin in Cytoscape. Network statistics were calculated with Network Analyzer as undirected networks using the defaults (Assenov *et al.*, 2008).

Results

Monthly covariance in microbial communities and the environment

T4-like virus community structure varied less than bacterial or protistan communities, whether compared between all months (Figure 1a) or between adjacent months only (Figure 1b). Month-to-month shifts in viral, bacterial and protistan communities occurred concurrently over the 3-year period (Figure 1b). Protistan similarity patterns were significantly correlated with the bacterial community when comparing across all months, although not for communities 1 month apart (Table 1). Protistan community similarity between adjacent months (1 month lag only) was positively correlated with primary production estimates ($r = 0.579$, $P = 0.004$) and negatively correlated to sea surface temperature ($r = -0.434$, $P < 0.05$). Shifts in community composition of T4-like viruses and bacteria between adjacent months and across all months (any length lag) were also positively correlated; in addition, the 5′ (5H) and 3′ (3H) TRFLP assays for the T4-like viruses were highly correlated as expected, given that they are two related measures (Table 1). Bray–Curtis similarities for bacterial community composition between adjacent months were also negatively correlated to sea surface temperature ($r = -0.491$, $P = 0.024$) such that communities were more similar from month-to-month during colder months. Bray–Curtis similarities between adjacent months for T4-like viral communities were significantly correlated to bacterial abundance (Figure 1c) and bacterial Bray–Curtis similarities between adjacent months (Figure 1d).

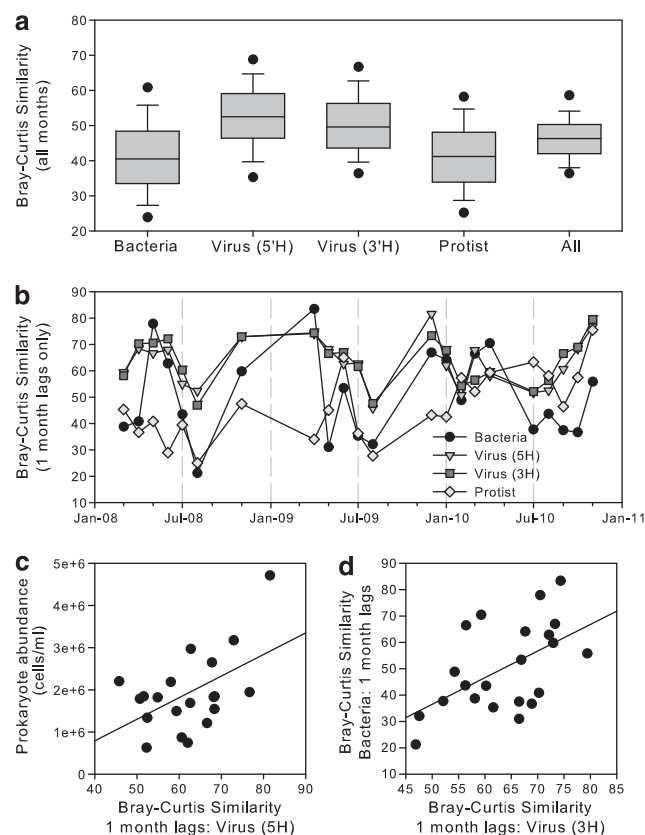Co-correspondence analysis (Braak and Schaffers, 2004) uncovered covariance of communities

Figure 1 Month-to-month shifts in Bray–Curtis similarity within microbial communities. (**a**) Average similarity within each community, observed approximately monthly, over 3 years. 'All' indicates the average similarity of all communities (that is, bacteria, viral and protistan) combined into one meta-community. Line, average similarity; box, 25th and 75th percentiles; and error bars, 10th and 90th percentiles. (**b**) Bray–Curtis similarity between adjacent months for each microbial group, plotted according to the earliest month (that is, March 2008 for comparing March 2008–April 2008). (**c**) Correlation of bacterial abundance (*y* axis) and (**d**) shifts in bacterial Bray–Curtis similarity between adjacent months (*y* axis) to viral Bray–Curtis similarity (*x* axis).

Table 1 Correlation of variability patterns between microbial communities based on Bray–Curtis similarities

	Bacteria	Virus (3′H)	Virus (5′H)	Protist
Bacteria		*r = 0.546*	*r = 0.532*	*r = 0.113*
		P = 0.01	*P = 0.01*	*P = NS.*
T4-like virus (3′-*Hinc*II)	*r = 0.21*		*r = 0.884*	*r = 0.154*
	P = 0.022		*P < 0.001*	*P = NS*
T4-like virus (5′-*Hinc*II)	*r = 0.178*	*r = 0.771*		*r = 0.088*
	P = 0.046	*P = 0.001*		*P = NS*
Protist	*r = 0.238*	*r = 0.046*	*r = 0.118*	
	P = 0.018	*P = NS*	*P = NS*	

Abbreviation: NS, not significant.
Correlations between Bray–Curtis resemblance matrices comparing all months are shown in *lower left triangle* and between communities from adjacent months are shown in *upper right triangle*. Bold text indicates statistically significant correlations.

from the relative abundance of each measured OTU (Supplementary Table S1). In our analysis, one microbial data set was considered as an

independent data set (the predictor), while the predictability of variance in a second (the response) was determined. 47% of variance of one T4-like virus TRFLP assay (3H) was predictable by the other (5H), and 20% of T4-like (3H) variability was due to changes in bacterial community composition. No T4-like community variance was significantly predictable by protistan community composition. However, protistan community variance was significantly predictable (P<0.05) by T4-like viral community variance (5H, 9.2% and 3H, 8.4%) and not by bacterial community composition. Co-correspondence analyses were unable to estimate bacterial community variance from T4-like or protistan community data, despite overall correlation in Bray–Curtis community similarities.

Microbial communities were predicted by up to five environmental parameters at $P \leqslant 0.05$ using canonical correspondence analysis when compared with the null model (Supplementary Table S1). Bacterial variation (12%) could be explained by chlorophyll-a concentration (bottle data) and salinity, and protistan community variance (11.6%) by day length and bacterial abundance. Day length, change in day length, salinity and temperature explained 28.3% of T4-like viruses (3H); viral abundance, ENSO index, day length, change in day length and temperature explained up to 33.5% of variability observed within the T4-like viral community (5H).

Correlations between individual bacterial, viral and protistan taxa in association networks
Many significant LS correlations were observed between viral, bacterial and protistan OTUs and environmental parameters (Table 2). After significance testing by permutation tests and screening by P-values and false-discovery estimates (q-values),

4365 of 223 446 (2%) possible pairwise LS values were statistically significant (Supplementary Figure S1) and formed one global network (Supplementary Figure S2). Most non-significant LS values ranged from −0.5 to 0.5 (that is, relatively weak correlations). LS correlations for both viral assays had largely similar distributions (Table 2). For simplicity, only LS correlations with the 3′ TRFLP (3H) viral OTUs were included in the following network figures (Figures 3–5, Supplementary Figures S3 and S4). We focused on 3H as month-to-month shifts in Bray–Curtis similarity, and similarities across all months between the 3H-viral and bacterial communities were significantly correlated.

Positive and negative interactions were observed between OTUs (that is, bacteria–bacteria, protist–bacteria and virus–bacteria) with co-occurring (not delayed) or time-shifted (delayed by 1 month) LS correlations. Figure 2 depicts simple networks that occurred within the whole community and the underlying relative abundance data. In Network A, a bacterial OTU and a protistan OTU were positively correlated—potentially indicative of co-occurrence, mutualism or predator-prey interaction. In Network B, the positive time-shifted LS correlation observed between one bacterial OTU and one viral OTU might reflect a lytic relationship. Negative and delayed correlations (not shown) between two viral OTUs could indicate a competitive relationship for hosts, whereas a negative correlation between OTUs may indicate lysis (virus–bacteria) or grazing (protist–bacteria). Pearson's correlation coefficients were close to LS values unless the optimal LS correlations were time-shifted (Table 3).

Bacteria–virus interactions and bacteria–protist interactions differed remarkably in their interconnectivity (Figure 3) despite each network having 18 unique components (unconnected subnetworks). The number of significant correlations between

Table 2 Distribution of significant LS correlations (edges) between all microbial OTUs and environmental parameters ($P \leqslant 0.0015$, $q<0.10$)

	Nodes	Bacteria	Virus (3′-H)	Virus (5′-H)	Protist	Bio+Chem	Phys
Bacteria	220	**791** 24 090					
T4-like virus (3′-HincII)	168	**428** 36 960	**462** 14 028				
T4-like virus (5′-HincII)	199	**353** 43 780	**979** 33 432	**529** 19 701			
Protist	61	**82** 13 420	**97** 10 248	**116** 12 139	**39** 1830		
Biological or chemical	15	**62** 3300	**68** 2520	**70** 2985	**5** 915	**12** 105	
Physical	6	**46** 1320	**93** 1008	**94** 1194	**9** 366	**10** 90	**9** 15

Abbreviations: Bio+chem, biological and chemical; Phys, physical.
'Nodes' indicates the number of OTUs or other parameters in the global network; the remaining columns indicate the number of significant pairwise correlations in bold above the total number of possible correlations between node types. 'Biological and chemical' includes: bacterial and viral abundances, nutrient concentrations, chlorophyll-a and so on; 'Physical' includes: salinity, temperature, day length, monthly change in day length and so on.
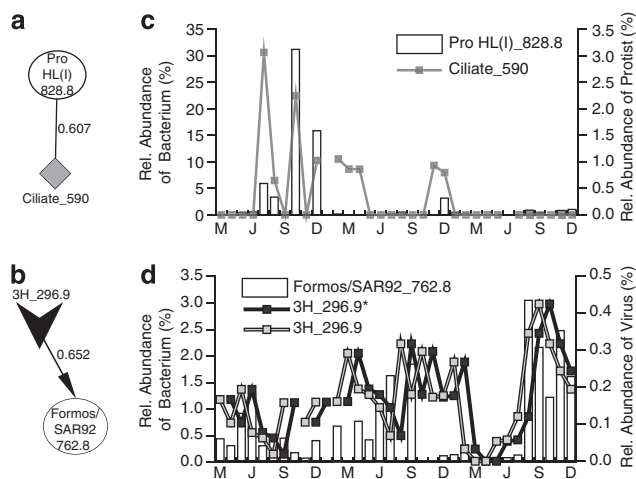
bacterial and T4-like virus OTUs (3H: 1.15% of all possible; 5H: 0.8%) outnumbered bacterial–protistan correlations (0.6%), although this may be because of the higher number of viral OTUs observed (Table 2). The protist–bacteria network (Figure 3a, Supplementary Figure S3) differed from the virus–bacteria network (Figure 3b, Supplementary Figure S4) in terms of the internal structure (that is, number of nodes and how the correlations were distributed between nodes); network statistics are summarized in Table 4. Most LS correlations were positively similar to the giant network of all OTUs; negative correlations were 19.5% ($n = 16$) of protist–bacteria and 25.9% ($n = 111$) of virus–bacteria. The global network retains small-world properties with short path lengths (number of nodes needed to link individual nodes) and a high clustering coefficient ratio such that many nodes are connected to other nodes in close-knit groups—more so than expected by chance alone. Although an OTU in the protist–bacteria network had 1.7 connections on average as opposed to 3.5 connections in the virus–bacteria network, these numbers are identical to those observed in random networks of equal size and far fewer than observed in the global network. However, network density (normalized parameter for the
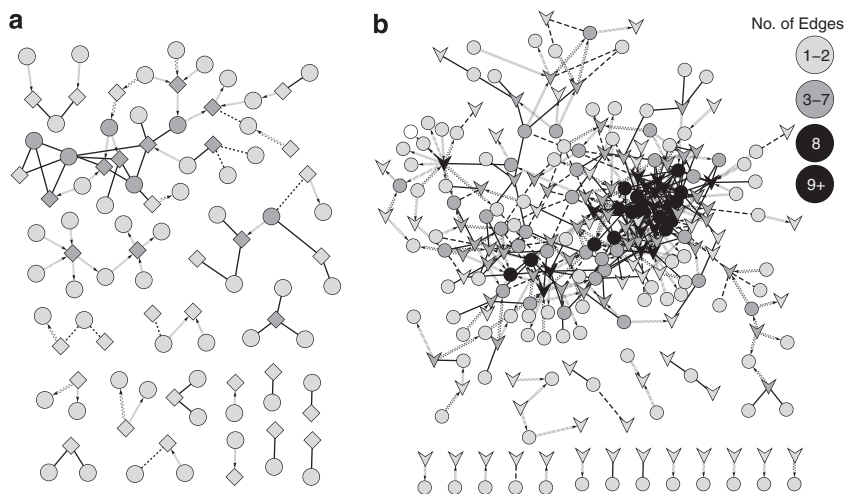

Figure 2 Two mini networks and the relative abundance of each OTU over time. Each mini network (**a** and **b**) depicts microbial OTUs as shapes (bacteria, circles; protists, diamonds; viruses, v-shapes). Lines represent statistically significant LS correlations with LS values shown: solid lines are positive correlations and arrows are delayed correlations, pointing toward lagging OTU. Relative abundance of each node is shown as a percent of each microbial community from March 2008–December 2010 for each network (**c** and **d**). Bacterial abundance is shown by the bar graph, whereas the protistan (**c**) and viral (**d**) OTU relative abundance is indicated by the line graph. * denotes the time-shifted viral OTU, as described in Table 3.

Table 3 Description of LS correlations displayed in Figure 2

Network	OTU (x)	OTU (y)	Int.	LS	Xs	Ys	Length	PCC	P_{PCC}
A	Pro_HL(I)_828.8	Ciliate_590	pu	0.607	1	1	32	0.655	$P < 0.005$
B	3H_296.9[a]	Formos/SAR92_762.8[b]	pdl	0.652	3	4	31	0.543	$P < 0.005$

Abbreviations: Int., interactions ; LS, local similarity; OTU, operational taxonomic unit; PCC and P_{PCC} are the Pearson's Correlation Coefficient with no delays and the associated *P*-value, respectively for the OTUs listed.
Interactions (Int.) indicate if correlations were positive with no time lags (pu) or positive with time lag (pdl). Xs and Ys note the month in which the LS correlation begins, and 'Length,' indicates the length of the LS correlation in months (of 34 maximum).
[a]denotes which OTU's relative abundance was shifted by 1 month in Figure 2D.
[b]indicates which OTU's relative abundance was 'fixed' in time.

Figure 3 Broad overview of interactions between (**a**) protists and bacteria only, and (**b**) T4-like viruses and bacteria only. Microbial nodes are bacteria, circles; protists, diamonds; viruses, v-shapes. Node color indicates its number of edges according to the scale shown in the upper right. Solid lines are positive correlations with no delay; dashed lines, negative correlations with no delay; sine-wave lines, negative-delayed correlations; and forward-slashed lines, positive-delayed correlations. Arrows point toward the lagging OTU. Note that correlations between similar taxa (for example, bacteria–bacteria, protist–protist and virus–virus) were omitted.
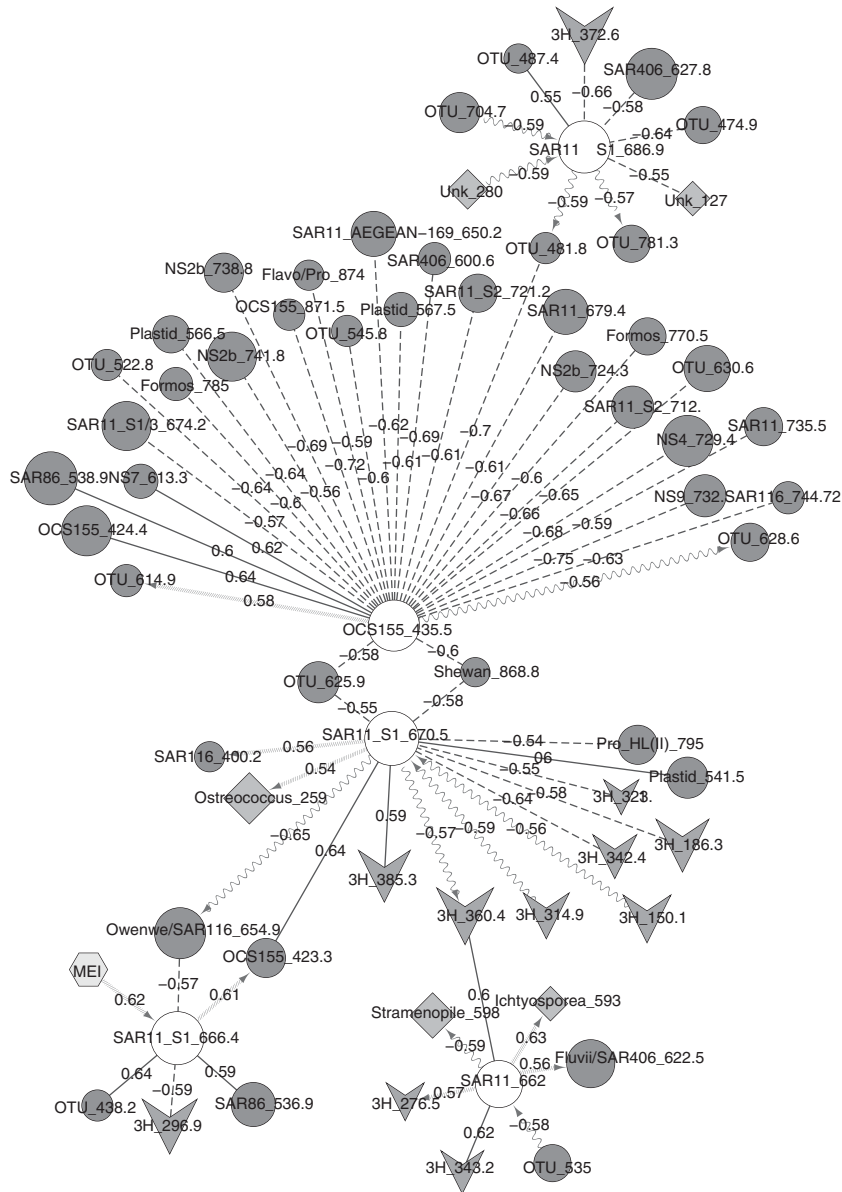
822

Figure 4 Top five bacterial OTUs differentially correlate to bacterial, viral and protistan OTUs. Top five bacterial OTUs are highlighted as white circles. All other nodes are bacteria, circles; protists, diamonds; viruses, v-shapes; abiotic, hexagon. Node labels indicate an abbreviated identity (where available) and fragment length. Note that SAR11_S1 indicates SAR11 Surface Clade 1. Solid lines are positive correlations with no delay; dashed lines, negative correlations with no delay; sine-wave lines, negative-delayed correlations; and forward-slashed lines, positive-delayed correlations. Arrows point toward the lagging OTU.

average connectivity) was slightly higher in the protist–bacteria network than virus–bacteria network. Each bacterial OTU was typically associated with a single protistan out, whereas a protistan OTU was often correlated to several bacterial OTUs to form small cliques. A viral OTU was often correlated to multiple bacterial OTUs, which were also correlated to two or more viral OTUs, and resulted in one large interconnected cluster. However, more correlated pairs of one viral to one bacterial OTU were observed than pairs of one protist to one bacterial OTU, which could skew the network density calculations when looking at average correlations per node overall. Network

heterogeneity (unevenness of the number of connections per node) is lower in protist–bacteria as compared with a random network of equal size, but higher in virus–bacteria, which would confirm a higher skew in the distribution of connections per viral OTU as opposed to protistan OTU. If secondary connections between OTUs of the same type (that is, bacteria–bacteria, protist–protist and virus–virus) were also shown, many small clusters would be connected although subgroups remained apparent.

We examined significant correlations for the five most abundant surface ocean bacterial OTUs. These four putative SAR11 and one Actinobacterium (OCS155_435.5) were significantly correlated to

Table 4 Global and local network statistics for microbial association networks

Parameters	Bacteria, protist, and virus (3H and 5H)	Bacteria, protist, and virus (3H only)	Protist to bacteria only	Virus (3H) to bacteria only
Nodes	669	465	94	248
Edges	4365	2224	82	428
Positive edges (%)	3094 (70.9%)	1652 (74.3)	66 (80.5%)	317 (74.1%)
Negative edges (%)	1274 (29.1%)	572 (25.7%)	16 (19.5%)	111 (25.9%)
No. of components	1	1	18	18
No. of components, random	1	1	25	9
Diameter (radius)	10 (5)	9 (1)	8 (1)	14 (1)
Diameter (radius), random	4 (4)	5 (4)	10 (1)	11 (1)
Connectivity				
Average number of neighbors	13.0	9.6	1.7	3.5
Network density	0.02	0.021	0.019	0.014
Likelihood for uneven distribution of edges				
Network heterogeneity	0.958	0.94	0.591	0.953
Network heterogeneity, random	0.287	0.310	0.873	0.532
Centralization	0.114	0.092	0.036	0.051
Centralization, random	0.019	0.023	0.058	0.023
Identifying small-world properties				
Average clustering coefficient (Cl)	0.241	0.227	ND	ND
Clustering coefficient, random (Cl_r)	0.02	0.022	ND	ND
Ratio of Cl/Cl_r	12.05	10.32	ND	ND
Characteristic path length (L)	3	3.523	ND	ND
Characteristic path length, random (L_r)	2.797	2.957	ND	ND
Log response ratio: Cl/Cl_r	1.08	1.01	ND	ND
Log response ratio: L/L_r	0.03	0.08	ND	ND

Abbreviation: ND, not determined.
Parameters were calculated for global networks of all microbial OTUs and environmental parameters with both (3H and 5H) and only one (3H) viral data set (network shown in Supplementary Figure S2). Local parameters are presented for the protist–bacteria and virus–bacteria sub-networks (seen in Figure 3 and Supplementary Figures S3 and S4).

several other viral, bacterial and protistan OTUs (Figure 4). This network comprised 66 nodes and 67 edges. OCS155_435.5 was negatively correlated to bacterial OTUs only. The four SAR11 OTUs were correlated to bacterial and viral OTUs, with several delayed correlations (1 month lag); SAR11_S1_666.4 was correlated to the El Niño Southern Oscillation Index (MEI). Four protistan OTUs, identified as *Ostreococcus*_259, Ichtyosporea_593, Stramenopile_598 and Unknown_280, were correlated with delay to two of the dominant SAR11 OTUs. The fifth protistan OTU, Unknown_127, was negatively correlated with no delay to SAR11_686.9.

Case study: cyanobacteria, possible grazers and viruses
Potential top-down relationships were determined from cyanobacterial, protistan and viral OTUs; this network of cyanobacterial OTUs and their correlated partners has 65 nodes with 66 correlations (Figure 5). Some cyanobacterial OTUs were connected to multiple viral OTUs and others were connected primarily to protistan or other bacterial OTUs; many virus–bacteria correlations were delayed. A non-simple path (a.k.a. a ring structure or cycle) that also included internal rings was identified for a series of correlations between cyanobacterial OTUs, other bacterial OTUs and one viral OTU. Three viral OTUs had TRFLP

patterns consistent with cultured isolates by *in silico* analysis of *g23* genes: (1) 3H_408.9: S-SM2 isolated from *Synechococcus* WH8017; (2) 3H_413.5: Syn9 or Syn19 isolated from *Synechococcus* WH8012 and WH8109, respectively; and (3) 3H_415.5: S-SSM7 isolated from *Synechococcus* WH8109 (Sullivan et al., 2003, 2010; Weigele et al., 2007). Protistan OTUs included two potential cyanobacterial grazers: (1) dinoflagellate or *Lingulodinium*-relative (though possibly a phototroph or mixotroph) and (2) a ciliate. Salinity and temperature were correlated to a potentially low-light *Prochloroccus* (OTU 907.8).

Discussion

Potential bottom-up controls on bacterial, protistan and viral communities
We undertook an integrated assessment of bacterial, protistan and T4-like viral communities over 3 years to address microbe–microbe interactions in the surface ocean and their relationship to environmental conditions. ARISA and 18S TRFLP both surveyed the entire domain of bacteria and protists, respectively, whereas g23-TRFLP focused on a specific viral family, the T4-like viruses. ARISA resolved organisms near the species-level (Brown and Fuhrman, 2005; Brown et al., 2005) and
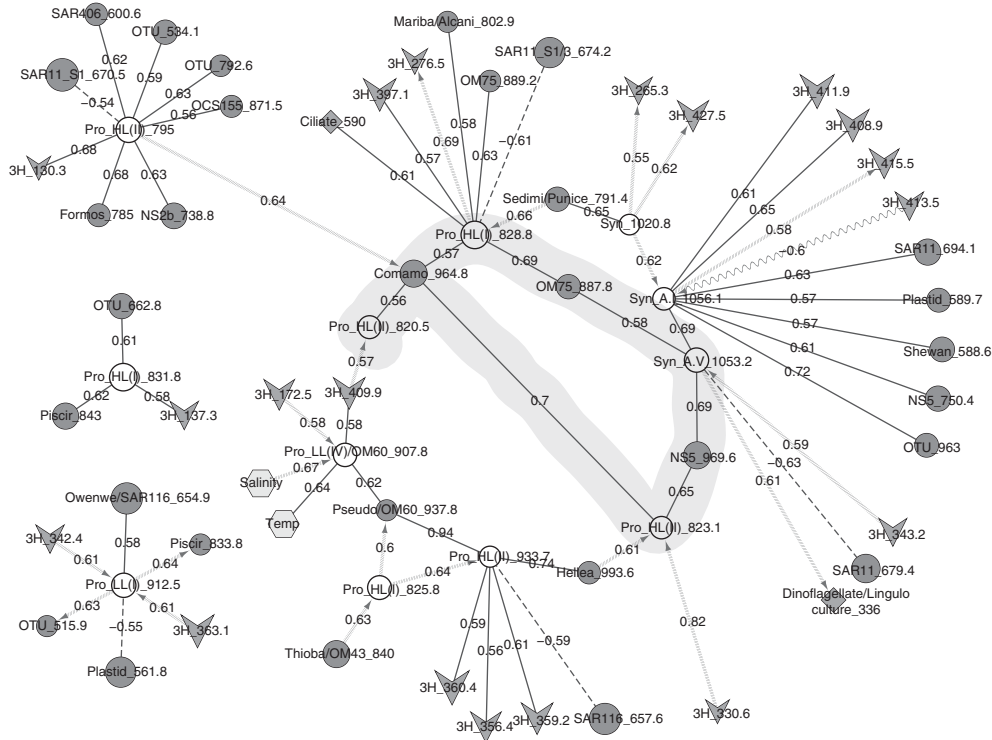
Figure 5 Cyanobacterial OTU correlations to other microbial OTUs reveal potential lytic virus–host relationships, grazing and temporal trends. *Cyanobacteria* OTUs are noted as white circles and labeled as *Prochlorococcus* (Pro) or *Synechococcus* (Syn), followed by ecotype designation (HL: high light; LL: low light; A/B: *Synechococcus* group). All other nodes are bacteria, circles; protists, diamonds; viruses, v-shapes; abiotic, hexagons. Node labels indicate an abbreviated identity (where available) and fragment length. Solid lines are positive correlations with no delay; dashed lines, negative correlations with no delay; sine-wave lines, negative-delayed correlations; and forward-slashed lines, positive-delayed correlations. Arrows point toward the lagging OTU.

detected OTUs >0.1%; g23-TRFLP detected >100 T4-like virus OTUs each month on average, each representing >0.1% of the total T4-like community (Chow and Fuhrman, 2012) and 18S rDNA TRFLP revealed dominant members (>0.5%) of the protistan community (Kim *et al.*, 2012).

Community variability was predictable by environmental factors, but only to a limited extent; each microbial community was discerned by a unique set of environmental factors. The proportion of predictable variability within the T4-like viral community (28.1%) was higher than within both protists (11.6%) and bacteria (11.9%). Bacterial variance was related to chlorophyll-*a* and salinity, whereas protistan variance was related to day length and bacterial abundance. Day length may have largely influenced the composition of the autotrophic (and mixotrophic) community, whereas prey availability (as indicated by bacterial abundance) may have been a major factor in shaping the heterotrophic (and mixotrophic) community. T4-like virus community shifts were predicted by day length, change in day length (distinguishes seasons), salinity and temperature—all of which relate to abiotic controls on abundance and physiological status of potential hosts. We were unable to significantly predict overall bacterial community structure from T4-like myoviral or protistan community structure. This likely reflects the complex

ecology of virus–bacteria–protist interactions, and that bacterial communities can be acted upon by a variety of viral types, not just the T4-likes targeted in this study. However, bacterial, protistan and T4-like viral communities were significantly correlated with one another at the community level, suggesting similar timing of responses to one another or the environment (Table 1; Figure 1).

Inference of ecological interactions from LS correlation networks
Monthly variation of community structure was apparent and the detailed inquiry that follows revealed underlying relationships between OTUs. Numerous statistically significant correlations between microbial OTUs were observed that potentially represent common, stable relationships in the surface ocean (Table 2). The structure of correlations between protistan and bacterial as compared with viral and bacterial OTUs (Figure 3, Supplementary Figures S3 and S4; Table 4) suggests increased specificity in virus–bacteria interactions. A relationship that switches partners (at a fine taxonomic resolution) often or at random is less likely to result in a significant correlation over a 3-year time series. Network statistics were likely heavily influenced by the presence of one large interconnected subnetwork between viral and bacterial OTUs with at least three apparent hubs in contrast to several

smaller hubs of protistan to bacterial OTUs. This is not unexpected. As such, grazing and viral lysis can be selective processes that result in different outcomes (Miki and Jacquet, 2008). For example, grazing can be more influenced by size rather than taxonomy of the prey (González et al., 1990; Monger and Landry, 1992; Simek and Chrzanowski, 1992; Hahn and Höfle, 1999). Taxonomically selective grazing is still thought to occur and may be represented by a protistan OTU linked to one or only a few bacterial OTUs (Figure 3a, Supplementary Figure S3). Protist–bacteria interactions could also represent associations where bacteria obtain nutrients indirectly from particular protists, for example, phytoplankton, and such potential positive (one-way) interactions could be inferred from known protistan phytoplankton (for example, Ostreococcus). Similarly, a virus' host range may be indicated by the number of bacteria 'host' OTUs correlated to a virus OTU such that more connections would suggest a broader host range (Figure 3b, Supplementary Figure S4). Note that the characterized T4-like viruses include several that are known to have a relatively broad host range, compared with other virus families (Sullivan et al., 2003). The large number of correlations between viruses and bacteria supports our observation on the correlated community-level shifts between these two groups.

Gross changes in viral abundance and community structure may have a partitioned, rather than universal impact on the bacterial community. Virus–bacteria interactions were not evenly dispersed throughout the community, as specific bacterial OTUs or clusters were more highly connected than others. Investigation into the five most abundant bacterial OTUs in the surface ocean at SPOT over a 10-year period suggests that three SAR11 OTUs may be equally influenced by interactions with viruses and other bacteria. Recently, a T4-like virus infecting SAR11 was discovered (Zhao et al., 2013); the primers used in this study would result in an uncut fragment of 332 bp based on in silico analysis. Of the potential viral OTUs observed near this length, none were correlated with any individual SAR11 OTUs in this analysis. These viral OTUs were correlated to Flavobacteria OTUs, other Alphaproteobacteria, Ichtyosporea (protist) and other viral OTUs; they were present in 16 months of our study ranging from 6% to 65.7% (3H) and in 28 months ranging from 22.4% to 60.9% (5H). The viral OTUs that did correlate with SAR11 may then represent other unknown pelagiphages. The remaining two most abundant bacteria (Actinobacteria 435 and SAR11_S1_666.4) were correlated to several bacteria, suggesting that bacteria–bacteria interactions may be more crucial to their success (Figure 4). Thus, virus–bacterial interactions defined some niches of our most dominant bacterial OTUs, whereas others were bound more by protistan or bacterial interactions.

Case study: connections between cyanobacteria, co-occurring microbes and the environment
Our network analysis revealed specific virus–cyanobacteria–protist interactions (Figure 5), which suggested that cyanobacteria were differentially responsive to top-down controls similar to their response to environmental pressures. *Prochlorococcus* and *Synechococcus,* two marine genera of cyanobacteria, are integral to marine ecosystems as key autotrophs in the microbial loop (Chisholm et al., 1988, 1992; Li, 1994; Campbell et al., 1997; Liu et al., 1997; Partensky et al., 1999; DuRand et al., 2001; Giovannoni and Vergin, 2012). It has been suggested that spatial, temporal and vertical differences in the distribution of specific ecotypes reflect physiological capabilities, adaptation to nutrient utilization and differential mortality (Moore et al., 1998; Martiny et al., 2009; Malmstrom et al., 2010; Partensky and Garczarek, 2010), which was seen in the observed correlations between cyanobacteria OTUs to salinity and temperature.

Ciliates and nanoflagellates are thought to be the predominant grazers of cyanobacteria, and newly identified groups such as lineages of marine stramenopiles are now considered as important bacterivores too (Christaki et al., 1999; Worden and Binder, 2003; Christaki et al., 2005; Massana et al., 2006; Frias-Lopez et al., 2009; Lin et al., 2012). A ciliate (OTU 590) and a dinoflagellate/*Lingulodinium* sp. OTU were correlated to a high-light *Prochlorococcus* OTU (828.8) and a *Synechococcus (group A-V)* OTU, potentially indicative of common grazing controls in the ocean. However, we cannot rule out that this dinoflagellate OTU may be a phytoplankton whose preferred conditions parallel to those of the *Synechococcus* OTU.

Cyanophage-host systems are some of the best-characterized host–virus models from the ocean. Stable, and seasonally variant, co-existing populations of viruses and their hosts have been observed in the field (Waterbury and Valois, 1993; Suttle and Chan, 1994; Marston and Sallee, 2003; Mühling et al., 2005; Sandaa and Larsen, 2006; Wang et al., 2011); isolate-based laboratory experiments have provided information on the genetic regulation of these interactions (Sullivan et al., 2003; Lindell et al., 2005; Zinser et al., 2009; Weinbauer et al., 2011). The correlations of 3H_415.5 (S-SSM7) and 3H_413.5 (Syn9 or Syn19) to Syn_A.I_1056.1 likely represent known host–virus interactions. The correlated bacterial OTU was a putative *Synechococcus;* specifically, *Synechococcus* WH 8109 (original host for Syn19 and S-SSM7) has an empirical ARISA length of 1055 bp, which falls within the OTU-labeled Syn_A.1_1056.1. Other *Synechococcus* spp. yield ARISA fragments of similar lengths. Both Syn9 and Syn19 have wide host ranges and are capable of infecting *Prochlorococcus* and *Synechococcus* (Sullivan et al., 2003). Host range data on S-SSM7 (3H_415.5) and S-SM2 (3H_408.9), to our knowledge, is unavailable, although both were

isolated from *Synechococcus sp.* isolates (Sullivan *et al.*, 2010). With that knowledge, we posit that these specific correlations between viral and putative *Synechococcus* WH8109-like OTUs 1053 and 1056 represent detection of an ongoing lytic relationship in the surface ocean.

Each cyanobacterial OTU in this network was correlated to at least one other bacterial (non-cyanobacterial) OTU, many of which were known heterotrophs. Prior studies reported that growth rates of *Synechococcus* are positively affected by lysis of co-occurring heterotrophic bacteria (Weinbauer *et al.*, 2011), and the presence of heterotrophic bacteria in culture with *Prochlorococcus* either significantly improved or inhibited growth depending on the co-cultured taxa (Sher *et al.*, 2011). The network's ring structure between these OTUs highlights potential redundancy within the cyanobacterial niches. The main ring includes a series of positively correlated cyanobacterial OTUs with and without delays, potentially illustrating a succession of cyanobacteria. All correlations point toward a positively correlated inner ring of three *Prochlorococcus,* one *Synechococcus* and three other heterotrophic bacterial OTUs (highlighted in Figure 5), suggesting that this group forms a guild or clique that follows *Synechococcus* (upper right) or other *Prochlorococcus* OTUs (from the left). Thus, these bacteria–bacteria links may identify which bacteria help form unique cyanobacterial cliques, alongside the established environmental features.

Conclusions

Monthly microbial community analysis at SPOT provided a mechanism for exploring relationships between individual bacterial, protistan and viral taxa that influence the seasonal variability in the surface ocean. Ecological relationships in the ocean are complex and the association networks presented here likely represent stable relationships between microbes observed *in situ*. eLSA has been applied to determine potential interactions that may represent boom-bust relationships over shorter daily timescales (Needham *et al.*, 2013), and others that may occur consistently and have particular ecological significance over longer monthly timescales were shown here. Connectivity of OTUs and observation of independent interconnected clusters indicated that microbial communities are full of potential niches that warrant further investigation. Protistan–bacterial associations were far fewer than virus–bacteria associations, and their connectivity may reflect relative non-selective or size-selective interactions as compared with the web of virus–bacteria interactions that may reflect a virus' host range. The dominance of positive over negative correlations suggests that on this monthly and seasonal timescale, viral OTUs may be primarily controlled

by host availability (that is, viruses following the hosts). The microbial association networks identified factors that were highly correlated to specific OTUs, such as cyanobacteria, and the persistence of those relationships over time. Our association networks support the paradigm that microbes are regulated by both bottom-up and top-down controls, and our findings add another layer of complexity to the bacterial response to changing microbial counterparts and environmental conditions.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Assenov Y, Ramírez F, Schelhorn S, Lengauer T, Albrecht M. (2008). Computing topological parameters of biological networks. *Bioinformatics* **24**: 282–284.

Azam F, Fenchel T, Field J, Gray J, Meyer-Reil L, Thingstad F. (1983). The ecological role of water-column microbes in the sea. *Mar Ecol Prog Ser* **10**: 257–263.

Baudoux A, Veldhuis M. (2008). Estimates of virus-vs grazing induced mortality of picophytoplankton in the North Sea during summer. *Aquat Microb Ecol* **52**: 69–82.

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

Bouvy M, Bettarel Y, Bouvier C, Domaizon I, Jacquet S, Le Floc'h E *et al.* (2011). Trophic interactions between viruses, bacteria and nanoflagellates under various nutrient conditions and simulated climate change. *Environ Microbiol* **13**: 1842–1857.

Braak Ter CJF, Schaffers AP. (2004). Co-correspondence analysis: a new ordination method to relate two community compositions. *Ecology* **85**: 834–846.

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

Brown MV, Fuhrman JA. (2005). Marine bacterial microdiversity as revealed by internal transcribed spacer analysis. *Aquat Microb Ecol* **41**: 15–23.

Brown MV, Schwalbach MS, Hewson I, Fuhrman JA. (2005). Coupling 16S-ITS rDNA clone libraries and automated ribosomal intergenic spacer analysis to show marine microbial diversity: development and application to a time series. *Environ Microbiol* **7**: 1466–1479.

Campbell L, Liu H, Nolla HA, Vaulot D. (1997). Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991-1994 ENSO event. *Deep-Sea Res Part 1 Oceanogr Res Pap* **44**: 167–192.

Chaffron S, Rehrauer H, Pernthaler J, Mering von C. (2010). A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res* **20**: 947–959.

Chisholm SW, Frankel SL, Goericke R, Olson RJ, Palenik B, Waterbury JB et al. (1992). Prochlorococcus marinus nov. gen. nov. sp.: an oxyphototrophic marine prokaryote containing divinyl chlorophyll a and b. *Arch Microbiol* **157**: 297–300.

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

Chow C-ET, Fuhrman JA. (2012). Seasonality and monthly dynamics of marine myovirus communities. *Environ Microbiol* **14**: 2171–2183.

Chow C-ET, Sachdeva R, Cram JA, Steele JA, Needham DM, Patel A et al. (2013). Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. *ISME J;* epub ahead of print 18 July 2013; doi:10.1038/ismej.2013.122.

Christaki U, Jacquet S, Dolan J, Vaulot D, Rassoulzadegan F. (1999). Growth and grazing on Prochlorococcus and Synechococcus by two marine ciliates. *Limnol Oceanogr* **44**: 52–61.

Christaki U, Vazquez-Dominguez E, Courties C, Lebaron P. (2005). Grazing impact of different heterotrophic nanoflagellates on eukaryotic (Ostreococcus tauri) and prokaryotic picoautotrophs (Prochlorococcus and Synechococcus). *Environ Microbiol* **7**: 1200–1210.

Clarke KR, Gorley R. (2006). *PRIMER v6: User Manual/Tutorial.* PRIMER-E: Plymouth, UK.

Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C et al. (2007). Integration of biological networks and gene expression data using Cytoscape. *Nat Protoc* **2**: 2366–2382.

Clokie MRJ, Millard AD, Mann NH. (2010). T4 genes in the marine ecosystem: studies of the T4-like cyanophages and their role in marine ecology. *Virol J* **7**: 291.

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.

Corno G, Jürgens K. (2008). Structural and functional patterns of bacterial communities in response to protist predation along an experimental productivity gradient. *Environ Microbiol* **10**: 2857–2871.

Countway PD, Gast RJ, Savai P, Caron DA. (2005). Protistan diversity estimates based on 18S rDNA from seawater incubations in the western north Atlantic1. *J Eukaryot Microbiol* **52**: 95–106.

Countway PD, Vigil PD, Schnetzer A, Moorthi SD, Caron DA. (2010). Seasonal analysis of protistan community structure and diversity at the USC Microbial Observatory (San Pedro Channel, North Pacific Ocean). *Limnol Oceanogr* **55**: 2381–2396.

Dunne JA. (2002). Food-web structure and network theory: the role of connectance and size. *Proc Natl Acad Sci USA* **99**: 12917–12922.

DuRand MD, Olson RJ, Chisholm SW. (2001). Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep-Sea Res Part 2 Top Stud Oceanogr* **48**: 1983–2003.

Eiler A, Heinrich F, Bertilsson S. (2012). Coherent dynamics and association networks among lake bacterioplankton taxa. *ISME J* **6**: 330–342.

Evans C, Archer SD, Jacquet S, Wilson WH. (2003). Direct estimates of the contribution of viral lysis and microzooplankton grazing to the decline of a Micromonas spp. population. *Aquat Microb Ecol* **30**: 207–219.

Filée J, Tétart F, Suttle CA, Krisch HM. (2005). Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proc Natl Acad Sci USA* **102**: 12471–12476.

Fisher MM, Triplett EW. (1999). Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl Environ Microbiol* **65**: 4630–4636.

Frias-Lopez J, Thompson A, Waldbauer J, Chisholm SW. (2009). Use of stable isotope-labelled cells to identify active grazers of picocyanobacteria in ocean surface waters. *Environ Microbiol* **11**: 512–525.

Fuhrman JA. (1999). Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541–548.

Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Natl Acad Sci USA* **103**: 13104–13109.

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

Fuhrman JA, Steele JA. (2008). Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat Microb Ecol* **53**: 69–81.

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

Gasol JM, Pedrós-Alió C, Vaqué D. (2002). Regulation of bacterial assemblages in oligotrophic plankton systems: results from experimental and empirical approaches. *Antonie Van Leeuwenhoek* **81**: 435–452.

Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B et al. (2012). Defining seasonal marine microbial community dynamics. *ISME J* **6**: 298–308.

Giovannoni SJ, Vergin KL. (2012). Seasonality in Ocean Microbial Communities. *Science* **335**: 671–676.

González JM, Sherr E, Sherr BF. (1990). Size-selective grazing on bacteria by natural assemblages of estuarine flagellates and ciliates. *Appl Environ Microbiol* **56**: 583–589.

Hahn MW, Höfle MG. (1999). Flagellate predation on a bacterial model community: interplay of size-selective grazing, specific bacterial cell size, and bacterial community composition. *Appl Environ Microbiol* **65**: 4863–4872.

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

Hooker SB, McClain CR. (2000). The calibration and validation of SeaWiFS data. *Prog Oceanogr* **45**: 427–465.

Kim DY, Countway PD, Yamashita W, Caron DA. (2012). A combined sequence-based and fragment-based characterization of microbial eukaryote assemblages provides taxonomic context for the Terminal Restriction Fragment Length Polymorphism (T-RFLP) method. *J Microbiol Methods* **91**: 527–536.

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

Lin Y-C, Campbell T, Chung C-C, Gong G-C, Chiang K-P, Worden AZ. (2012). Distribution patterns and phylogeny of marine stramenopiles (MAST) in the North Pacific Ocean. *Appl Environ Microbiol* **78**: 3387–3399.

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

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

Longnecker K, Wilson M, Sherr E, Sherr BF. (2010). Effect of top-down control on cell-specific activity and diversity of active marine bacterioplankton. *Aquat Microb Ecol* **58**: 153–165.

Malmstrom RR, Coe A, Kettler GC, Martiny AC, Frias-Lopez J, Zinser ER *et al.* (2010). Temporal dynamics of Prochlorococcus ecotypes in the Atlantic and Pacific oceans. *ISME J* **4**: 1252–1264.

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

Martiny AC, Tai APK, Veneziano D, Primeau F, Chisholm SW. (2009). Taxonomic resolution, ecotypes and the biogeography of Prochlorococcus. *Environ Microbiol* **11**: 823–832.

Massana R, Terrado R, Forn I, Lovejoy C, Pedrós-Alió C. (2006). Distribution and abundance of uncultured heterotrophic flagellates in the world oceans. *Environ Microbiol* **8**: 1515–1522.

Middelboe M. (2000). Bacterial Growth Rate and Marine Virus-Host Dynamics. *Microb Ecol* **40**: 114–124.

Miki T, Jacquet S. (2008). Complex interactions in the microbial world: underexplored key links between viruses, bacteria and protozoan grazers in aquatic environments. *Aquat Microb Ecol* **51**: 195–208.

Miki T, Jacquet S. (2010). Indirect interactions in the microbial world: specificities and similarities to plant–insect systems. *Popul Ecol* **52**: 475–483.

Moebus K. (1996). Marine bacteriophage reproduction under nutrient-limited growth of host bacteria. I. Investigations with six phage-host systems. *Mar Ecol Prog Ser* **14**: 1–12.

Monger BC, Landry MR. (1992). Size-selective grazing by heterotrophic nanoflagellates: an analysis using live-stained bacteria and dual-beam flow cytometry. *Arch Hydrobiol Beih* **37**: 173–185.

Montoya JM, Pimm SL, Solé RV. (2006). Ecological networks and their fragility. *Nature* **442**: 259–264.

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

Mühling M, Fuller NJ, Millard A, Somerfield PJ, Marie D, Wilson WH *et al.* (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 C-ET, 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.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB *et al.* (2011). vegan: Community Ecology Package (R package version 2.0.2)http://CRAN.R-project.org/package=vegan.

Olesen JM, Stefanescu C, Traveset A. (2011). Strong, long-term temporal dynamics of an ecological network. *PLoS ONE* **6**: e26455.

Ory P, Hartmann HJ, Jude F, Dupuy C, Del Amo Y, Catala P *et al.* (2010). Pelagic food web patterns: do they modulate virus and nanoflagellate effects on picoplankton during the phytoplankton spring bloom? *Environ Microbiol* **12**: 2755–2772.

Partensky F, Garczarek L. (2010). Prochlorococcus: advantages and limits of minimalism. *Annu Rev Marine Sci* **2**: 305–331.

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

Ruan Q, Dutta D, Schwalbach MS, Steele JA, Fuhrman JA, Sun F. (2006a). Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* **22**: 2532–2538.

Ruan Q, Steele JA, Schwalbach MS, Fuhrman JA, Sun F. (2006b). A dynamic programming algorithm for binning microbial community profiles. *Bioinformatics* **22**: 1508–1514.

Sandaa R-A, Gómez-Consarnau L, Pinhassi J, Riemann L, Malits A, Weinbauer MG *et al.* (2009). Viral control of bacterial biodiversity—evidence from a nutrient-enriched marine mesocosm experiment. *Environ Microbiol* **11**: 2585–2597.

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

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

Shannon P. (2003). Cytoscape: a software environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* **13**: 2498–2504.

Sher D, Thompson JW, Kashtan N, Croal L, Chisholm SW. (2011). Response of Prochlorococcus ecotypes to co-culture with diverse marine bacteria. *ISME J* **5**: 1125–1132.

Sherr E, Sherr BF. (1988). Role of microbes in pelagic food webs: a revised concept. *Limnol Oceanogr* **33**: 1225–1227.

Sherr E, Sherr BF. (2002). Significance of predation by protists in aquatic microbial food webs. *Antonie Van Leeuwenhoek* **81**: 293–308.

Simek K, Chrzanowski TH. (1992). Direct and indirect evidence of size-selective grazing on pelagic bacteria by freshwater nanoflagellates. *Appl Environ Microbiol* **58**: 3715–3720.

Simek K, Pernthaler J, Weinbauer M, Hornák K, Dolan JR, Nedoma J *et al.* (2001). Changes in bacterial community composition and dynamics and viral mortality

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

Simpson GL. (2009). cocorresp: Co-correspondence analysis ordination methods (R package version 0.1-9)http://cran.r-project.org/package=cocorresp.

Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**: 431–432.

Sole RV, Montoya M. (2001). Complexity and fragility in ecological networks. *Proc R Soc Lond B* **268**: 2039–2045.

Staniewski MA, Short CM, Short SM. (2012). Contrasting community versus population-based estimates of grazing and virus-induced mortality of phytoplankton. *Microb Ecol* **64**: 25–38.

Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY *et al.* (2011). Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J* **5**: 1414–1425.

Strom SL. (2000). Bacterivory: Interactions between Bacteria and their Grazers. In *Microbial Ecology of the OceansKirchman DL (edsWiley-Liss, Inc.: NY, USA, pp 351–386.*

Sullivan MB, Huang KH, Ignacio-Espinoza JC, Berlin AM, Kelly L, Weigele PR *et al.* (2010). Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. *Environ Microbiol* **12**: 3035–3056.

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

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

Vigil P, Countway P, Rose J, Lonsdale D, Gobler C, Caron D. (2009). Rapid shifts in dominant taxa among microbial eukaryotes in estuarine ecosystems. *Aquat Microb Ecol* **54**: 83–100.

Wang K, Wommack KE, Chen F. (2011). Abundance and distribution of Synechococcus spp. and Cyanophages in the Chesapeake Bay. *Appl Environ Microbiol* **77**: 7459–7468.

Waterbury JB, Valois FW. (1993). Resistance to co-occurring phages enables marine Synechococcus

communities to coexist with cyanophages abundant in seawater. Appl. Environ. *Microbiol* **59**: 3393–3399.

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

Weinbauer MG, Bonilla-Findji O, Chan AM, Dolan JR, Short SM, Simek K *et al.* (2011). Synechococcus growth in the ocean may depend on the lysis of heterotrophic bacteria. *J Plankton Res* **33**: 1465–1476.

Weinbauer MG, Christaki U, Nedoma J, Simek K. (2003). Comparing the effects of resource enrichment and grazing on viral production in a meso-eutrophic reservoir. *Aquat Microb Ecol* **31**: 137–144.

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

Worden AZ, Binder BJ. (2003). Application of dilution experiments for measuring growth and mortality rates among Prochlorococcus and Synechococcus populations in oligotrophic environments. *Aquat Microb Ecol* **30**: 159–174.

Xia LC, Ai D, Cram JA, Fuhrman JA, Sun F. (2013). Efficient statistical significance approximation for local association analysis of high-throughput time series data. *Bioinformatics* **29**: 230–237.

Xia LC, Steele JA, Cram JA, Cardon ZG, Simmons SL, Vallino JJ *et al.* (2011). Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC Syst Biol* **5**: S15.

Zhang R, Weinbauer MG, Qian P-Y. (2007). Viruses and flagellates sustain apparent richness and reduce biomass accumulation of bacterioplankton in coastal marine waters. *Environ Microbiol* **9**: 3008–3018.

Zhao Y, Temperton B, Thrash JC, Schwalbach MS, Vergin KL, Landry ZC *et al.* (2013). Abundant SAR11 viruses in the ocean. *Nature* **494**: 357–360.

Zinser ER, Lindell D, Johnson ZI, Futschik ME, Steglich C, Coleman ML *et al.* (2009). Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph, Prochlorococcus. *PLoS One* **4**: e5135.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)