Marine microbial community dynamics and their ecological interpretation

Jed A. Fuhrman, Jacob A. Cram and David M. Needham

Abstract | Recent advances in studying the dynamics of marine microbial communities have shown that the composition of these communities follows predictable patterns and involves complex network interactions, which shed light on the underlying processes regulating these globally important organisms. Such 'holistic' (or organism- and system-based) studies of these communities complement popular reductionist, often culture-based, approaches for understanding organism function one gene or protein at a time. In this Review, we summarize our current understanding of marine microbial community dynamics at various scales, from hours to decades. We also explain how the data illustrate community resilience and seasonality, and reveal interactions among microorganisms.

Phototrophic
Organisms that can transform light energy into biologically usable (chemical) forms. A photoautotroph obtains its biomass carbon from fully oxidized carbon (that is, carbon dioxide) and a photoheterotroph obtains its biomass carbon from reduced carbon (for example, organic matter).

Chemotrophic
Organisms that can obtain energy from chemical reactions. A chemoautotroph obtains its biomass carbon from carbon dioxide, and a chemoheterotroph obtains its biomass from reduced carbon.

Department of Biological Sciences and Wrigley Institute for Environmental Studies, University of Southern California, Los Angeles, California 90089–0371, USA. Correspondence to J.A.F.
e-mail: fuhrman@usc.edu
doi:10.1038/nrmicro3417
Published online
9 February 2015

Marine microbial communities (consisting of bacteria, archaea, protists, fungi and viruses) process about one-half of the global biogeochemical flux of biologically important elements, such as carbon, nitrogen, phosphorus, sulphur and iron[1–4]. These organisms include phototrophic and chemotrophic primary producers, as well as heterotrophic 'secondary' producers, which recycle dissolved organic carbon and nutrients through the microbial loop[5]. In recent years, molecular analysis has made great strides in determining the organisms that are present at a given site and how they are distributed over space and time. Most of this information is derived from phylogenetic analyses using a few informative genes, such as the 16S and 18S rRNA genes[6–8], but phylogenetic identification alone is insufficient to assess the environmental functions and ecology of the community. Although we are learning much about such functions via 'omics' (metagenomic, metatranscriptomic, metaproteomic and metabolomic analyses), such studies alone rarely provide the information that is needed to predict interactions, competition for nutrients, symbioses and other processes that determine the overall roles of the distinct organisms in the sea. Therefore, additional information, beyond that which can be derived from omics, is needed to predict community functions and interactions between organisms. Towards this goal, a great deal can be learned by evaluating the dynamics of community composition and the corresponding environmental parameters to determine the extent to which such dynamics follow predictable patterns, and potentially to show which organisms interact and how. In the context of this Review, the term dynamics is used to refer to changes in the abundance of various members (or populations) in a community, but not necessarily changes in the physiological state of these members; the term community is used here to refer to the types of microorganisms present and their relative proportions. The time-series approach, which involves evaluating the dynamics of microbial community composition and the surrounding environment over multiple time points (BOX 1), including the subsequent analysis of the large data sets that are created (BOX 2; FIG. 1), can be considered a holistic, system-wide investigation that complements the more reductionist approaches of examining the system gene by gene or protein by protein.

Recent studies of marine community dynamics over various timescales, from diel to interannual, have illuminated not only the environmental conditions that are preferred by individual microorganisms but also the likely interactions among these organisms and the emergent properties of the system as a whole. In particular, these studies have shown that marine microbial communities are dynamic (in a constant state of flux) but also resilient, which means that their behaviours are broadly predictable in terms of typical features of daily, seasonal and interannual variation in community composition. This implies that despite external forces that alter the community (such as temperature, nutrient supply and physical mixing), there are internal feedback mechanisms, including competition, viral infection

Box 1 | **Temporal dynamics versus spatial variation**

All biological systems are dynamic on one or more scales. The term 'dynamics' specifically refers to changes over time, and temporal changes in microbial communities result from growth and death, as well as the import and export of each organism present in the community. Intraspecies evolution, resulting in the emergence of variants with new combinations of traits, may also contribute to community dynamics, but this is beyond the scope of this discussion.

Dynamics are relatively straightforward in closed, artificial systems (such as bioreactors, in which import and export are constrained), but such closed systems are rare to non-existent in nature, and the sea is a particularly open and complex environment. In marine systems, dynamics are complicated by water mixing and advection via currents (which are directional) and eddies (which typically swirl). Sampling of dynamics in the sea is traditionally carried out by two alternative time-series sampling modes: Lagrangian, in which the sampling attempts to follow the prevailing current; and Eulerian, in which the sampler remains at a fixed geographic location and currents will move water past the sampling location. Lagrangian sampling attempts to track a 'parcel' of water as it moves, but in practice even Lagrangian sampling does not track a parcel of water because eddies mix the water and there are no truly stable parcels. Thus, the movement and mixing of water need to be considered when interpreting the dynamics of marine microbial communities to avoid confusing temporal dynamics with spatial variation. Most ocean time-series studies tend to use Eulerian sampling for practical and logistical reasons; Lagrangian sampling is impractical in studies that are carried out over the course of more than a few weeks. It is also important to consider horizontal and vertical scales separately, because much of the ocean is stratified vertically, with restricted vertical mixing; thus, physicochemical and biological gradients can be on scales of metres (or less) vertically.

The extent to which water mixing (which drives the import and export of organisms) alters the composition of a parcel of water is determined by the composition of the adjacent water. Thus, although horizontal mixing occurs all the time (via diffusion and eddies), a parcel of water can have a stable microbial composition as long as the microbiologically important conditions (for example, temperature and nutrient concentrations) are consistent in the surrounding water. From the limited available data, it seems that the size of a typical microbiologically coherent parcel (with a consistent microbial community composition) is in the order of 2–20 km in horizontal extent[105,106]. However, at fronts where conditions such as temperature or chlorophyll levels change abruptly, sharper gradients in microbial composition are expected. In any case, the difficulty of distinguishing changes in time from changes in space is a practical problem that biological oceanographers must live with, often not fully satisfactorily.

Notably, this discussion refers to planktonic organisms, which by definition move with the currents. Most marine bacteria, archaea and viruses are so small that they do not sink. However, some larger microorganisms (for example, larger protists such as diatoms and ciliates), and the bacteria that are attached to such particles, sink. Thus, the composition of microbial communities in the deep sea, fed primarily from sinking particles, is no doubt greatly influenced by sinking particles and the nutrients and microorganisms that they transport.

and predator–prey interactions, that help to maintain a remarkably steady 'average' community year after year, even in seasonally changing habitats. Dynamics also show that correlations differ among organisms depending on the timescale that is examined, which indicates that different processes operate at different scales. Furthermore, comparisons of dynamics among organisms help to determine which microorganisms are members of distinct ecological species, in the sense of having differences in ecological properties, and can point to what those differences are and their causes. In this Review, we explore the evidence for microbial community dynamics and the processes that lead to them, as well as the approaches to use these dynamics to learn more about potential interactions among microorganisms. The scope of this Review includes a focus on bacteria and archaea, for which whole-community molecular analyses are well developed and data sets are available. Owing to space limitations, an extensive discussion of protistan and phage dynamics is excluded, although some prominent examples are highlighted.

Heterotrophic
Organisms that obtain their biomass from reduced carbon.

Microbial loop
The microbial components of the food web — such as bacteria, archaea, protists and viruses — which together process the recycling and return of dissolved organic matter back into the 'classical' food web (the food web that is more traditionally related to the transfer of particulate organic matter, including organisms).

What processes drive dynamics at different scales?

Microorganisms change over multiple timescales and in response to different forces, including both biological and non-biological properties of the environment that drive changes in microbial community composition. As the typical average generation times of marine plankton are approximately a day in surface waters, and longer in deep waters[1], substantial changes in community composition in less than a few hours are not expected; therefore, this discussion addresses changes over timescales of hours and longer. Important timescales and likely forcing functions are outlined below, and these often overlap with each other.

Hours. The physiological response of microorganisms to changing conditions can occur rapidly, but compositional changes observed at subdaily timescales can only occur in communities in which organisms replicate relatively rapidly. Candidates for change on this timescale are copiotrophic organisms (often Gammaproteobacteria, Flavobacteriia, certain Alphaproteobacteria and others) that can grow rapidly, and although these taxa are usually rare in seawater, they can quickly become abundant under suitable conditions[9]. Community changes that occur on the scale of hours might also result from rapid selective cell death, but this has not yet been documented. Changes in community composition at this scale can result from both predictable and unpredictable causes. Among the predictable changes are the day–night cycle of primary productivity, internal clocks and interactions with vertically migrating plankton (such as during grazing). A diel periodicity of whole-community bacterial productivity[10–12] and abundance is sometimes observed, which has been suggested to be a community response to higher daytime production[10,11]. Among particular taxa, cyanobacteria have a diel pattern of abundance, which is probably due to the synchronization of cell division with the photoperiod[13]. Unpredictable short-term changes include those induced by storm or wind events (which induce upwelling and deposit allochthonous material) and those that are consequences of biological interactions, such as competition, grazing, sloppy feeding, viral infection and lysis, and cross-feeding. The response to wind events can take less than a day, which occurred, for example, during a strong event in the Red Sea in which diatoms showed an approximate 28-fold increase in abundance in 24 hours[14].

Daily to weekly. In the offshore surface ocean, estimated whole-community biomass turnover times (which are derived from microbial growth rates) range from less than a day to about a week[1]. Therefore, this timescale is suitable for the observation of microbial variation in response to environmental variation. The forces that vary substantially on this scale include weather, mesoscale oceanographic processes (for example, eddies that

Box 2 | **Methods for the analysis of time series and microbial networks**

Whole-community ordination. Whole-community ordination approaches summarize multivariate data by depiction in two (and sometimes more) dimensions[107]. In non-metric multidimensional scaling, similar samples are plotted near each other; such analysis has revealed dynamics across depths and between seasons at the Bermuda Atlantic Time-series Study (BATS)[28,29]. Canonical correspondence analysis enables samples and species to be examined in the context of seasonal factors and other factors, such as chlorophyll concentrations. It has been used to link environmental and biotic variability to bacterial community structure in a short time-series of the North Sea[108]. All ordination approaches can simplify considerable complexity, but some patterns in community dynamics can be lost, necessitating the application of other analytical methods[29,33].

Identifying seasonality and long-term patterns. Discriminant function analysis (DFA) uses machine learning to 'predict' categorical variables from community structure (using mathematics, after the event has occurred). For example, it showed a predictive relationship between the bacterial or viral community composition alone and the month of sampling, which was interpreted as demonstrating seasonality in microbial community composition[48,52]. As this application of DFA searches through the data for only a fraction of the available taxa (typically <30 bacterial or viral types) that show the desired pattern, it ignores organisms that may constitute the majority of the community if they are not seasonal. Bray–Curtis analysis pseudo-autocorrelation (FIG. 1) detects both seasonal and long-term time-series patterns by quantifying dissimilarity between a pair of samples and then plotting the time difference against community similarity. It is used to detect variability on daily, monthly and longer timescales[46,50,54,56]. The advantages of this method include statistical verifiability with Mantel and other tests, and that it is a more direct measurement of variability than DFA. However, it can fail to detect patterned variability of some taxa if the majority of the community varies on other scales, as this overwhelms the signal of rarer operational taxonomic units. Other methods include Mantel tests, which determine whether two similarity matrices are positively correlated and have been used with Bray–Curtis analysis to verify seasonal and long-term patterns. These are robust to non-normal data. Partial Mantel tests can factor out otherwise overwhelming variables, such as seasonality and long-term variability, to identify how parameters of interest influence community variability[56].

Association networks. In microbial ecology, association networks have been most commonly generated from pairwise correlations (or sometimes other mathematical relationships) among complex data sets. In addition to connections between individual microbial groups, the networks display emergent properties, such as the degree of interconnectedness, which relates to characteristics of the community associations[88,89].

General additive mixed models. General additive mixed models (GAMMs) enable the investigation of individual parameters for seasonality and other trends[56]. In time-series analysis, GAMMs tolerate autoregression within parameters (samples that are closer in time tend to be similar) and unevenly spaced data. The mathematics behind GAMMs is complex and newly developed, so they have been rarely used to date.

False discovery rate calculations. Community analysis involves relating many parameters, which greatly increases the probability of spurious correlations. Calculating *Q* values (also known as false discovery rates[109]) allows researchers to keep the numbers of false positive results low enough to make informative conclusions about data.

Fronts
Boundaries in the sea between different water masses, across which environmental conditions (for example, temperature and nutrient concentrations) change abruptly, similar to weather fronts on land where there are abrupt changes in temperature and humidity.

Copiotrophic
Organisms that generally rely on relatively high concentrations of nutrients and have rapid maximum growth rates.

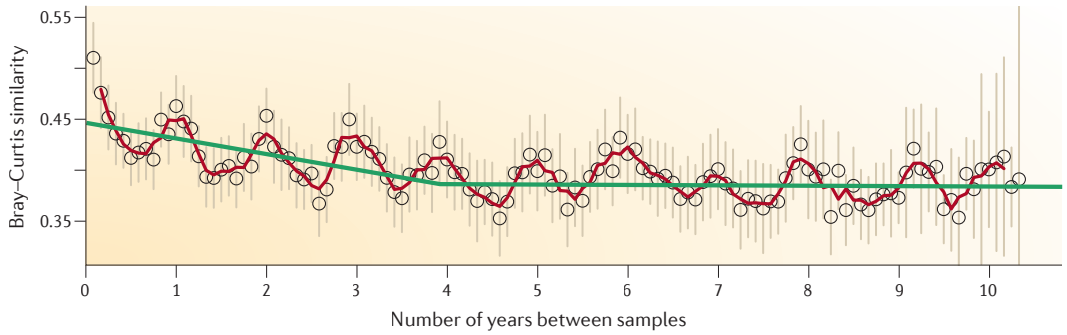
are tens to hundreds of kilometres wide), interactions with larger organisms (from protists to fish), food-web cascade effects and microbial interactions involving all types of viruses, bacteria, archaea and protists.

The daily-to-weekly timescale is appropriate for studying dynamics associated with phytoplankton blooms, which have a direct influence on the composition of bacterial and archaeal communities via cross-feeding interactions and the effects of toxins or other allelopathic substances, as well as an indirect influence through oxygen depletion and injury or death of larger organisms (if the bloom is toxic). Changes in the composition of a bacterial community in response to phytoplankton blooms is informative because the response is typically dramatic and often occurs in several stages, which can facilitate the identification of causal factors[15–17]. The species of phytoplankton present during the course of a bloom can be variable, as with the bacterial response, but in general a few copiotrophic bacterial lineages have been repeatedly shown to be involved: Flavobacteriia, *Rhodobacter* spp. and Gammaproteobacteria[17]. One study[15] that examined phytoplankton blooms on the daily-to-weekly scale in the North Sea followed the physiological and abundance successional patterns of bacteria associated with a diatom bloom. Bacteroidetes (including Flavobacteriia) were succeeded by Gammaproteobacteria (*Reinekea* spp. and SAR92), other Flavobacteriia members and Alphaproteobacteria (the Roseobacter clade-affiliated (RCA) group) over a period of weeks. The Flavobacteriia expressed enzymes that degrade polymers and particles (algal cells and their decay products), and were succeeded by other taxa that degrade polymers, which have a competitive advantage at intermediate concentrations of dissolved organic matter[15]. When total bacterial counts declined towards the end of the succession, the SAR11 cluster, which is thought to have a competitive advantage at low concentrations of dissolved organic matter, increased in relative abundance. In Monterey Bay, California, heterotrophic bacterial communities associated with phytoplankton blooms varied between seasons, depending mostly on the bacterial taxa that were already seasonally enriched[18].

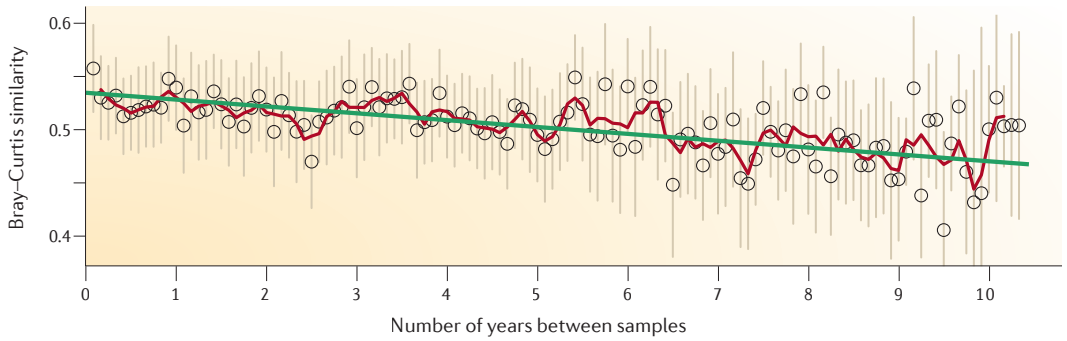
Monthly to seasonal. Forcing functions that vary at this scale include changes in solar angle (including associated changes in light intensity and ultraviolet penetration of the ocean), seasonal weather patterns (such as winds and storm frequencies), seasonal upwelling and the associated changes in nutrient availability, and stratification, the four of which are interrelated (the first is weakly related to the last three, which are closely related). Other forcing factors include changes in temperature and day length; seasonal variation in land runoff and atmospheric deposition; and interactions with larger organisms and other microbial species, which can all be influenced by the same pervasive environmental factors. Importantly, monthly sampling is the most common interval in major long-term ocean time-series studies, so this is discussed in more detail below. When referring to 'seasonal' patterns, it is implied that there is an annual repeating pattern, not just changes over a period of a few months or a year. Hence, data need to be collected over several years (ideally consecutive years) to convincingly show that a pattern is predictably seasonal rather than just changing over months. Nevertheless, patterns are often thought to be seasonal when they seem to be driven by physical and chemical factors that occur on a seasonal basis. Notably, because of seasonal winds, currents and upwelling, some seasonal changes in composition may reflect advection of different communities from adjacent waters rather than internal dynamics (BOX 1), which is not unique to this timescale.

Interannual. Forcing functions that are most variable at this scale include the El Niño Southern Oscillation, the Pacific Decadal Oscillation, stochastic year-to-year weather and/or climate variability, directional climate change

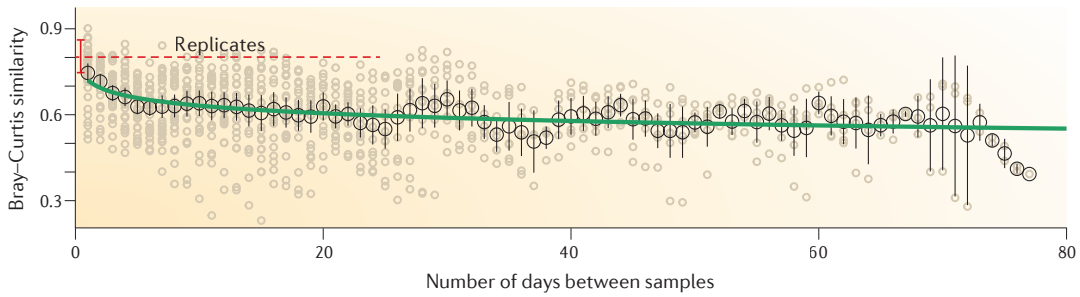
a Monthly sampling at 5 m (by ARISA)



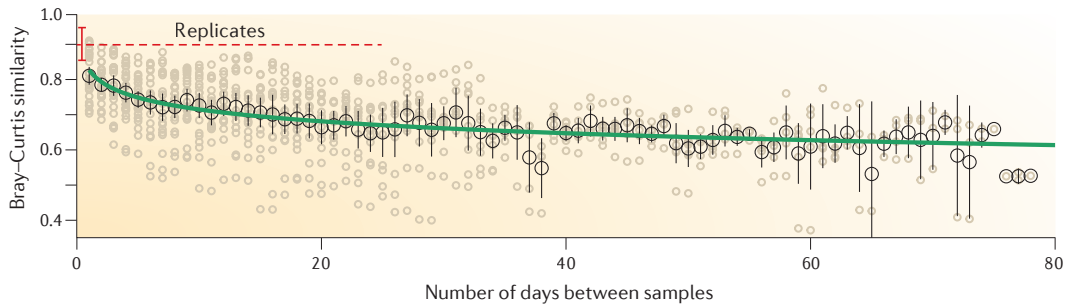
b Monthly sampling at 150 m (by ARISA)



c Daily sampling at 1 m (by ARISA)



d Daily sampling at 1 m (by 16S rRNA tag sequencing)



Grazing
In the context of plankton, grazing refers to the removal of organisms by a predatory or herbivorous organism.

Upwelling
The physical upward vertical transport of water from deeper to shallower depths (often caused by seasonal offshore winds near a coast). In the context of this Review, it refers to water and its accompanying nutrients entering the euphotic zone from below.

Allochthonous material
Material that is derived from an external source, as opposed to autochthonous material that is generated internally.

Eddies
Swirls of water motion caused by flow around objects or by instabilities inherent in the motion of density-stratified water on a rotating Earth. Ocean mesoscale eddies are typically 10–500 km in diameter.

Autoregression
Refers to the phenomenon in which samples that are collected closer to each other (in space or time) tend to be more similar to each other than those further separated.

Phytoplankton
(Also known as photosynthetic plankton). Single-cell photosynthetic organisms that form the basis of the marine food web, including cyanobacteria and many kinds of protists (such as diatoms, dinoflagellates, coccolithophores and others).

Allelopathic
An organism is termed allelopathic if it produces biochemicals that influence the growth, behaviour and reproduction of other organisms, with negative allelopathy adversely affecting the target organisms. The chemicals are generally not required for metabolism and include compounds such as antibiotics or repellants.

(trending in one direction, such as anthropogenic global warming and ocean acidification[19–21]), other anthropogenic regional environmental changes (such as hypoxia and anoxia caused by eutrophication[22]) and food-web cascade effects due to variations in larger organisms, such as fish. Other forces operating at this scale include alterations in entire ecosystems owing to overfishing, invasive species or general environmental degradation. Furthermore, there are even longer-term climatological variations at time scales much longer than decades, which emerge owing to geological and astronomical processes that are associated with changes in microbial assemblages: for example, foraminiferal sp. have been shown to vary over scales of tens to thousands of years[23,24], and diatom[25] and dinoflagellate[26] populations have trended towards species with smaller cell sizes over the past 50 million years owing to, among other variables, increased stratification (which generally decreases

◀ Figure 1 | **Resilience and predictability of community composition.** **a** | A Bray–Curtis similarity plot (Bray–Curtis pseudo-autocorrelation; see BOX 2) of samples collected on a monthly basis at a depth of 5 m for over a decade at the San Pedro Ocean Time-series (SPOT), which were analysed by ARISA (automated ribosomal intergenic spacer analysis)[46]. Each data point is the average of pairwise measurements of samples collected 'x' number of months apart: for example, the first point shows the average similarity of all pairs of samples collected 1 month apart, the second point from all samples taken 2 months apart, and so on, with the last point representing the average similarity of the first and last samples. The sinusoidal pattern, with peaks every 12 months and troughs 6 months after each peak, indicates that community composition follows a seasonal pattern. There is a downward trend over the first 4 years (straight green line with steep slope), which indicates interannual variability, whereas this trend plateaus over time, suggesting long-term stability of the average community. Error bars indicate 95% confidence intervals. **b** | A similar plot to that in part **a**, in which ARISA was used but the samples were taken at a depth of 150 m. The absence of a sinusoidal pattern indicates that seasonality in community composition is not demonstrated at this depth, and the gentle downward trend (green line) indicates a degree of long-term variability. Error bars indicate 95% confidence intervals. **c** | A similar Bray–Curtis similarity plot to those in parts **a** and **b** (generated using ARISA) but with higher resolution as a result of daily sampling over a period of 35 days at a depth of 1 m, followed by multiple samples that were taken weekly, up to a total of 85 days[54]. Small grey circles represent individual pairwise similarities, and large black circles represent the mean, with black bars showing 95% confidence intervals. That is, the first black point is the mean similarity of all samples taken 1 day apart and the grey points above and below it are individual pairwise comparisons, the second black point and grey points above and below it are from samples taken 2 days apart, and so on. **d** | A similar plot to that in part **c** but analysed by 16S rRNA tag sequencing[76]. The red horizontal dashed lines and error bars in parts **c** and **d** represent the mean and 95% confidence intervals of the Bray–Curtis similarities between replicate DNA samples (the mean is 0.8 for ARISA and 0.9 for tag sequencing). Notably, the ARISA and 16S rRNA tag sequencing results are very similar, but with a slight increase in similarity when tag sequencing is applied. Parts **a** and **b** are from REF. 56, Nature Publishing Group; and part **c** is from REF. 54, Nature Publishing Group.

Stratification
The layering of water by density, which is caused by variations in temperature and salinity. As vertical mixing eliminates stratification, the existence of stratified water indicates that there is no significant vertical mixing over the stratified depth range.

El Niño Southern Oscillation (ENSO). The effect of long-term variations in sea surface temperatures, primarily in the eastern Pacific Ocean, with anomalies lasting several months. Although the underlying causes are not fully known, it is correlated with pressure variations along the entire Pacific Basin and it affects upwelling, productivity and global weather patterns.

Pacific Decadal Oscillation
A long-term (20–30-year) oscillation of temperatures in the Pacific Ocean north of 20 degrees.

Euphotic zone
The upper layer of an aquatic water column that has sufficient light to support net photosynthesis.

nutrients in surface waters). These changes can be explored via the fossil record and by examining chemical or isotopic signatures. Paleo-oceanographic time series such as these, which retrieve useful information about past events in the ocean, are complementary to those that measure present-day dynamics.

The timescales at which marine microbial evolution (for example, through mutation, horizontal gene transfer and selection) occurs for the creation of ecologically distinct 'new' organisms are currently unclear, and it is also unclear whether such changes are important driving forces or occur as a response to driving forces, compared with compositional changes by rearrangement and selective growth among pre-existing microbial taxa.

Long-term time series

The most straightforward approach to observe the dynamics of marine microbial communities is through the use of oceanographic time-series studies that measure the composition of the microbial community. There are a few long-term studies at major time-series sites (as discussed below) and many small-scale studies, all of which typically examine the underlying physical and/or chemical oceanographic features, such as temperature, salinity, and chlorophyll and nutrient concentrations, which have a strong influence on microbial dynamics.

Bermuda. The Bermuda Atlantic Time-series Study (BATS) samples the subtropical waters (~33° N) of the northern Sargasso Sea[27]. The annual winter convective mixing of waters in this region (down to depths of 160–200 m), followed by re-stratification, leads to changes in the overall composition of the microbial community[28–33]. This is reflected in seasonal variabilities in temperature, inorganic nutrients and chlorophyll concentrations in the top 200 m of the water column[27,34], with fewer seasonal changes occurring in deeper waters. An important feature of most temperate surface waters, including those at BATS, is the spring bloom, in which re-stratification of the water column after winter mixing causes an overall increase in phytoplankton abundance and production (in other words, net growth), along with changes in the composition of the phytoplankton community[35–38]. Moreover, because phytoplankton form the base of the food web, this bloom leads to an increase in the abundance of bacterioplankton[28], which is also reflected in a changing bacterial community structure[28]. The dominant bacteria in the upper water column at BATS include seasonal subclades of SAR11, SAR86, SAR116 and the cyanobacteria *Prochlorococcus* and *Synechococcus*. Analysis of BATS communities has suggested that seasonal variation is strongest in the upper euphotic zone and that a dichotomy in surface-water communities exists, which is caused by the mixing of waters during the winter and early re-stratification, and the increase in stratification during the summer and autumn months. Recently mixed water in the euphotic zone has a higher concentration of nutrients and is dominated by phytoplankton and presumably bacterial species that are adapted to the comparatively higher nutrient levels, including the metabolites released by the phytoplankton. In the summer, the water is more stratified and a greater number of oligotrophic taxa, such as most SAR11 strains and SAR86, dominate the nutrient-poor surface waters. *Prochlorococcus* are the dominant photosynthetic organisms in the summer, and their abundance peaks in the deep chlorophyll maximum layer. The abundance of viruses also varies on a seasonal scale and positively correlates with the abundance of *Prochlorococcus* and Rhodobacteraceae, and negatively correlates with that of SAR11 and *Synechococcus*. These dynamics suggest that, compared with *Prochlorococcus*, SAR11 may be regulated differently by viruses and that viruses have a role in seasonal succession patterns[39]. Rare copiotrophic taxa have more irregular patterns than organisms that are more abundant[33]. Some notable low-abundance (rare) bacteria from the Vibrionaceae and Alteromonadaceae families occasionally bloom, substantially increasing in abundance for a short period of time and then declining[33]. Microbial community dynamics in the Sargasso Sea (and probably elsewhere) are also driven by the movement of ~200 km mesoscale eddies[40,41]. Analysis of the distribution of microbial taxa across a mode-water (anticyclonic but upwelling) mesoscale eddy near BATS showed enrichment of certain taxa (for example, roseobacters) that prefer higher levels of nutrients and phytoplankton growth near the surface, in addition to upward transport of mesopelagic species[42].

Hawaii. The Hawaii Ocean Time-series (HOT) is located in the North Pacific Gyre, 100 km north of Oahu, at the northern boundary of the tropics (~22° N)[43]. Compared

with the sea surface at BATS, the sea surface at HOT has less annual variability in temperature, mixed layer depth and inorganic nutrient content[44]. Microbial communities at HOT stratify by depth and show less seasonal variability than those at BATS[31]. Some seasonality is evident at the surface and at certain depths above 200 m (for example, seasonal variation is strong at 75 m but not at 45 or 100 m)[45], suggesting that seasonal effects may be influenced by depth-specific features. Nevertheless, many of the microbial taxa co-occurred with other taxa or were positively correlated with certain environmental parameters, such as pH and nitrogen availability[45]. In terms of the abundance of dominant populations, the surface community present at HOT during an entire year is similar to the summer surface community at BATS[31].

Southern California. The San Pedro Ocean Time-series (SPOT) is in the San Pedro Channel, which is ~20 km off the coast of Southern California. Similarly to BATS, it has a subtropical latitude (~33° N); however, the water temperature is cooler than that of BATS owing to the southbound California Current and regional seasonal upwelling. SPOT is situated above the deepest part of the San Pedro Basin at a depth of 890 m; owing to restricted horizontal water flow in the basin at depths below a few hundred metres, the oxygen concentration decreases with depth to a few micromoles per litre near the sea floor. Seasonal factors at SPOT include both seasonal mixing in the top ~40 m[46] as well as seasonal variability in current structure at the San Pedro Basin[47].

Sampling over the course of a decade has shown that the microbial communities at SPOT vary seasonally[46,48–50] and are dominated by various members of the SAR11 cluster, Actinobacteria, and the cyanobacteria *Prochlorococcus* and *Synechococcus*. The T4-like myovirus superfamily[51]— which has a broad host range and infects Cyanobacteria, Proteobacteria and other bacterial phyla — has been studied by a fingerprinting approach based on the conserved g23 gene. This has revealed that these viruses also vary seasonally at the surface and that the abundance of many correlates with that of specific bacteria[52,53]. On a daily-to-weekly scale, the bacterial and T4-like viral communities at the water surface undergo modest day-to-day changes, with several viral taxa correlating (usually with a delay of 1–5 days) with specific bacterial taxa[54]. Microbial communities in deeper waters (at depths of 150, 500 and 890 m) are dominated by dynamic Thaumarchaea[55], Marine Group A, SAR324, *Nitrospina* spp., Flavobacteriia and other unnamed 'deep' SAR11 clades[56], and only modest changes in average bacterial composition were observed at 500 m over the course of the time series[57]. Ammonia-oxidizing archaea in the lower euphotic zone correlate with members of SAR11, SAR86 and Bacteroidetes[58]. Initial examination of data collected over 8 years showed moderate variability but little seasonality in the community at 890 m, which is just above the sea floor[50]. However, the analysis of data collected over 10 years using an alternative statistical approach has shown that the entire microbial community changes seasonally at both the surface and 890 m but not at intermediate depths[56]. This suggests that sinking organic particles, which also vary according to season and originate at the ocean surface, pass through intermediate depths too rapidly to make the entire communities at those depths change on a seasonal scale. By contrast, the accumulation and degradation of such particles on the sea floor has a more pervasive influence on the composition of the community in the overlying water. Despite the lack of community seasonality at intermediate depths, individual taxa did show seasonal variation at every depth, which suggests that some microorganisms may respond to seasonal variations in particle flux[56].

Western English Channel. Station L4 in the western English Channel is located 10 km off the southwest coast of England, over a shallow continental shelf, and has a depth of 55 m[59,60]. In the northern temperate zone at ~50° N, it is situated at a boundary between oceanic and coastal environments and is influenced by estuarine outflow from Plymouth Sound and seasonal mixing in the winter to a depth of ~35 m[60,61]. Microbial populations at the surface have been studied by monthly sampling over a period of 6 years, and their composition varies strongly according to season, with large shifts in the entire community over the year and the highest diversity during the winter[61,62]. In 1 month out of the 6 years (which coincided with a diatom bloom), a *Vibrio* sp. dominated the community, even though it was rare during the rest of the time series, demonstrating the ability of some rare taxa to take advantage of ephemeral niches, similar to (but more dramatic than) what was reported at BATS[33].

Northwestern Mediterranean Sea. There are two microbial observatories in the northwestern Mediterranean Sea, located just over 100 km from each other, in the temperate zone at ~42° N. The Banyuls observatory is ~4 km off the coast of Banyuls, France, and the Blanes Bay Microbial Observatory (Spain) is to the west and ~1 km offshore. The surface waters at both stations show seasonal variability, particularly in temperature, light levels, rainfall and stratification[63], with winter mixing in the Northwest Mediterranean Sea extending down to 100 m below the surface[64]. Bacterial community seasonality, which is inferred in terms of relative species abundance, was reported over two 1-year time-series studies in Blanes Bay in 1997 and 2003; seasonal groups of bacteria included SAR11, which are also seasonal in other environments, such as BATS[65,66]. The study of archaeal communities over a 4.5-year period at Blanes Bay showed that the relative abundances of archaea and the proportions of various Group II Euryarchaeota (MGII) and Group I Thaumarchaeota (MGI) were found to follow a seasonal pattern of variation[67]. At Banyuls, the presence of both archaeal rDNA and rRNA was examined, which showed that there was seasonal variability of both abundant and rare archaea[68]. This study clearly showed that taxa within both MGI and MGII have different dynamics. For example, multiple taxa within MGI and MGII change seasonally but vary distinctly from each other, which suggests that they each have a unique

Anticyclonic
A type of motion. In oceanography, it refers to the circulation of water around a region of high pressure. In the Northern Hemisphere, water rotates clockwise when viewed from above.

Mesopelagic
The middle of the ocean water column, below the euphotic zone and above the sea floor and the 'abyssal zone', which is usually defined as the water below average depth of the world ocean (approximately 4,000 m).

Estuarine outflow
Water passing from a river (low salinity) to the sea (normal ocean salinity).

Ephemeral niches
Sites at which specific environmental conditions exist for only a relatively short period of time.

response to different combinations of environmental factors, such that they can be considered to have different ecological niches. Therefore, this is a clear example of how dynamics can contribute to our understanding of taxa and their distinct niches.

Broader interpretations

Seasonality between sites. By comparing microbial dynamics between the different sites described above, consistent patterns begin to emerge. For instance, microbial communities show more seasonal variation in illuminated surface waters than in darker deep waters, which is consistent with the seasonal variation in sunlight and the observation that there is more seasonal variation in the physical and chemical properties of surface waters. Although seasonal variation is observed at all of the major time-series sites, it is weakest at HOT, intermediate at SPOT and strongest in the English Channel. This is unsurprising, as seasonal changes in microbial community composition reflect seasonal changes in the environment, which generally increase with distance from the tropics.

Overall stability and predictability. Large data sets from time-series studies offer a unique insight into the main factors that control microbial dynamics because they enable long-term relationships to be examined, which averages out the noise from measurements of community composition and from daily or monthly fluctuations (BOX 2). A powerful approach that identifies patterns at multiple timescales involves plotting the average differences in overall community composition (as measured by the Bray–Curtis similarity) as a function of the time lag between samples[46,50,54]. High similarity in community composition over time implies temporal stability, whereas low similarity implies that the community is changing. The monthly plot for surface waters at SPOT (FIG. 1a) shows a remarkable pattern of sine-wave-like variability, with clear peaks in similarity every 12 months (the same month of the year), and minima occurring 6 months after each peak. Similarly, a strong annual periodic pattern was also reported in a 6-year study of the western English Channel[50,61]. Together, these studies reveal that communities are more similar to each other during the same season of the year (even when they are several years apart), whereas the community composition is most different when comparing opposite seasons (those 6 months apart). Furthermore, the average similarity declines a little between the shortest time lags (of up to 4 years) and then it levels off, which implies that samples that are 1 year apart are slightly more similar than those 2, 3 or 4 years apart, but once samples are 4 years apart the average similarity remains the same until they are at least 10 years apart. By contrast, the monthly plot for a depth of 150 m[56] shows no obvious seasonality and a slight decline in average similarity over a decade (FIG. 1b).

It could be argued that because the average Bray–Curtis similarity values in these plots (FIG. 1a,b for SPOT; western English Channel[50] not shown) are only about 40–50%, these communities are not particularly similar overall. However, if there were secular changes in the long term (for example, loss of species or successful invasions) or even long-term changes in the relative abundances of existing taxa, the slope of the average would continue to decline substantially with increasing lags. Notably, this does not occur in surface waters (FIG. 1a; western English Channel[50] not shown), and the average similarity only declines slightly at a depth of 150 m (FIG. 1b). This suggests that there are clear limits on the variation in community composition over the entire period studied and that some kind of ecological 'invisible hand' is operating, such that the average proportions of the various microbial taxa persist over at least a decade. The invisible hand metaphor refers to how the cumulative effect of individuals working in their own self-interest, through various feedback processes, leads to well-regulated community function without any centralized control mechanism, similar to Adam Smith's description of economic markets being regulated by combinations of individual self-interests. We propose that the invisible hand involves negative feedback mechanisms including the biological and physiochemical conditions that control microbial community composition. Some aspects of such an invisible hand are beginning to be revealed: for example, the rapid evolution and trade-offs between phage susceptibility and resistance in marine cyanobacteria[69,70], whereby phage resistance either reduces the bacterial growth rate or increases susceptibility to other phages. Meanwhile, some phages have very narrow specificities (at the strain level) and others infect several host species[71]. Another example is the 'Black Queen' hypothesis[72], which posits that particular taxa lose the ability to perform certain functions (for example, produce vitamins or detoxify reactive oxygen species), so they require services from other microorganisms in the ecosystem[73], leading to stable coexistence[74].

The sinusoidal shape of the monthly similarity curve at the sea surface (FIG. 1a), with peaks at yearly intervals and troughs at 6 months offset, indicates that seasonal factors (such as day length, temperature and wind) are clearly important in the large-scale control of the surface community but do not seem to have a strong effect at a depth of 150 m (FIG. 1b), which is unsurprising given the minimal influence of such factors at this depth. A similar sampling approach also revealed interannual variability at both SPOT and the western English Channel[46,50]. SPOT and the English Channel seem to have similar patterns of short-term, seasonal and interannual variability[50]. The strongest variability by far occurs within seasons, followed by weaker but detectable interseasonal variability and even weaker interannual variability. The strong variability within seasons suggests that factors other than temperature, day length and additional seasonal factors (such as winds or upwelling) are dominant in the control of community composition. This includes less predictable factors such as day-to-day changes in weather (that is, daily variations around the monthly climate averages), complex currents and biological interactions at several trophic levels, as discussed above.

Ecological niches
(Here we use the version of the concept popularized by G. E. Hutchinson). The multidimensional properties of the lifestyle of an organism or population, including the resources it uses, preferred temperatures, pressures, light levels and spectra, salinities, pH, redox state; its competitors, predators, prey, parasites, pathogens, symbionts; its alteration of its own habitat, and any other environmental factors that influence its growth and survival. The concept of an ecological species refers to a species that occupies a particular multidimensional niche that is distinct from others.

Bray–Curtis similarity
A pairwise comparison of two communities, whereby each community composition has been described by the proportions of comprising taxa, and the similarity is calculated as the total of all the shared proportions of all the taxa in both communities. It ranges from 0, when no taxa are shared, to 1, when all taxa are present in the same proportions in both communities.

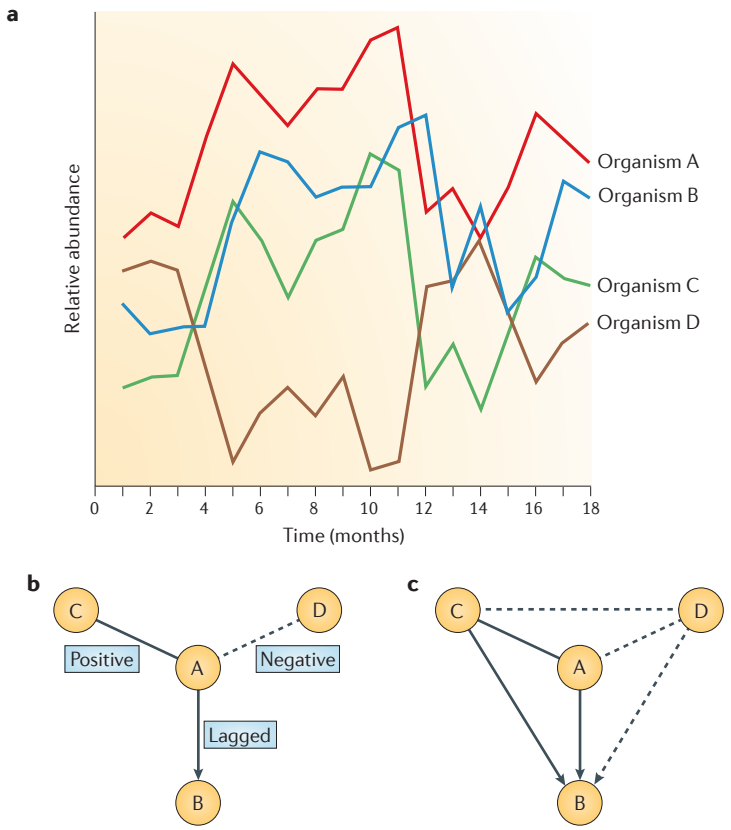


Figure 2 | **Translation of community dynamics into a microbial association network. a** | A graph of four hypothetical organisms that change in abundance over time. Note that the relative abundance of organism A is positively correlated with that of organism B but with a 1-month time lag, the abundance of A is also positively correlated with that of organism C but is negatively correlated with that of organism D. **b** | A 'hub-and-spoke' network showing how organisms B, C and D correlate with organism A (which is the hub of the network). Positive correlations are represented as solid lines, negative correlations as dashed lines and lagged associations as lines with arrows, pointing from the leading towards the lagging organism. **c** | A network showing all of the correlations between all pairs of organisms (A–D), using the solid, dashed and arrow lines defined in part **b**. In addition to what is shown in part **b**, this also shows organism B with a lagged positive correlation to organism C and a lagged negative correlation to organism D, and organism C with a negative correlation to organism D.

As opposed to the monthly similarity plot, the daily similarity plots (FIG. 1c,d) show a different scale of control, which includes bacterial automated ribosomal intergenic spacer analysis (ARISA; also known as fingerprinting)-based data (FIG. 1c) and data from the same samples evaluated by 16S rRNA tag sequencing[8,75,76] (FIG. 1d), as tag sequencing is now more widely used and is usually considered to cover more taxa than fingerprinting. Notably, the overall shapes of the curves for both approaches are the same, although the between-sample similarities (that is, values on the *y* axis) are slightly lower when using ARISA. As we are looking at relatively small temporal differences, to what extent these observed temporal variations differ from those of replicate samples needs to be considered (FIG. 1c,d). For both methods, samples that were taken 1 day apart vary from each other by about 5–10% more than the variation between replicate samples. In other words, the day-to-day variation in community composition is about 5–10% higher than the variability in the measurement. Thus, if this 5–10% variability were to occur randomly, the community would be quite different after a week and unrecognizable after a month. By contrast, the average slope of the variation between consecutive days is only about 0.2%, which implies that when the community changes, as it does moderately from day to day, the following days are much more likely to return the community towards its local average state than to keep changing it further from what it was. Again, it seems that there is an invisible hand that stabilizes the community within well-defined limits. Furthermore, this leads to the stability and predictability from month to month and season to season. Importantly, the discussion above refers to averages that are calculated by examining many samples, and the patterns that emerge (including seasonality and average day-to-day stability) may not be obvious from a small number of measurements. This is evident from the considerable scatter about the mean (FIG. 1c,d), in which any single data point can differ considerably from the mean. Therefore, the 'predictability' that emerges from these studies is, by analogy, much more similar to the predictability of monthly climate (which is a long-term average), than daily weather (which varies considerably around mean climatic values). In other words, with access to long-term data, it is possible to make long-term predictions about the likely composition of the average microbial community for a given month but not for a given day.

Automated ribosomal intergenic spacer analysis (ARISA). A microbial community fingerprinting approach, in which the highly variable intergenic spacer between the 16S and 18S rRNA is amplified by PCR, and the products, which vary from taxon to taxon, are separated and sized by a genetic fragment analyser.

Inferring associations from dynamics

Dynamic microbial systems, unlike systems that are at equilibrium or steady state, enable us to examine how different organisms change in relation to one another and to environmental conditions[49,77]. Correlations and anticorrelations among organisms and environmental properties reflect potential direct and indirect interactions. This is a powerful window into the processes that operate in complex natural microbial ecosystems, which are usually not possible to observe directly. Positive correlations may have direct causes, such as cross-feeding[78] or parasitism, or less direct causes, such as an overlapping preference for similar environmental conditions (for example, psychrotrophic organisms that may each prefer cold conditions for different reasons). Negative correlations may reflect competition, allelopathy, predator–prey interactions or a preference for mutually exclusive environmental conditions. A time series has the advantage of enabling the detection of correlations that are delayed in time; in addition to helping to separate 'cause' (which occurs early) from 'effect' (which occurs later), time-delayed correlations are probably the norm in many ecological interactions, such as classic Lotka–Volterra predator–prey or virus–host interactions. Confidence in the accuracy of inferred relationships among organisms is strengthened when the correlations are robust across locations and long timescales, although many relationships probably vary depending on timescale, season, other organisms that are present and environmental conditions[79]; such complexity is necessary to embrace and has begun to be explored[80].

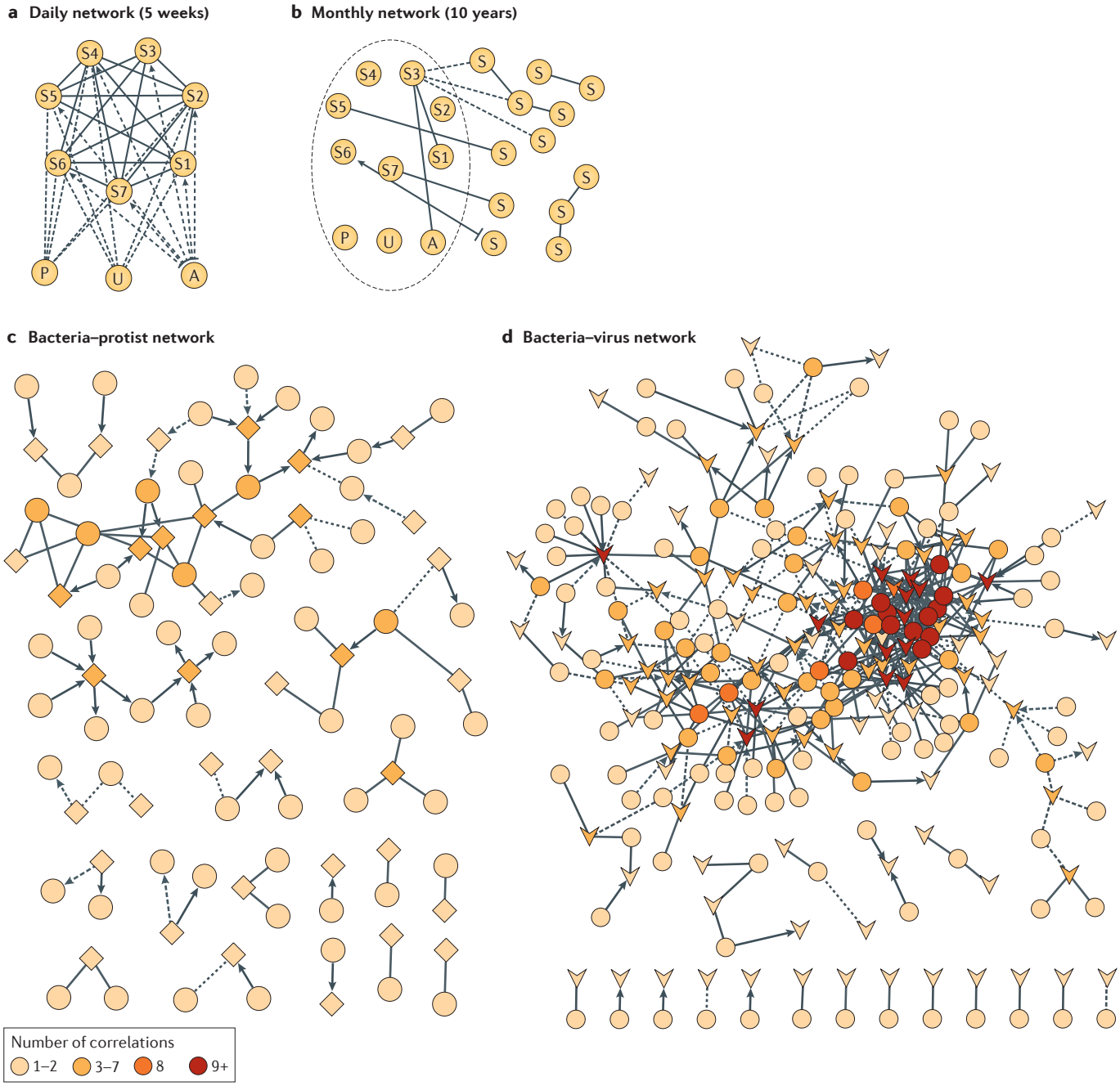


Figure 3 | **Network features change dramatically at different timescales and with the inclusion of different microorganisms.** All of the networks shown are from the San Pedro Ocean Time-series. A comparison of a daily (part **a**) and a monthly (part **b**) network shows that the same organisms (members of the SAR11 cluster) are highly intercorrelated (meaning that their abundances follow highly similar patterns) on a daily scale over the 5-week period studied, but not on the monthly scale over a decade[46,54,90]. The daily network shows the most interconnected bacterial module (highly intercorrelated taxa)[54], and the monthly network includes the same taxa (within the dashed oval) in addition to other SAR11 taxa[46]. 'S' indicates SAR11, with S1, S2, S3, and so on, indicating SAR11 subgroups that are the same in both networks; 'A' indicates Actinobacteria; 'P' indicates *Photobacterium* and 'U' indicates unidentified bacterium. Few taxa from the daily module are correlated on the monthly scale, showing that relationships vary greatly between scales. The bacteria–protist network (part **c**) has remarkably different connectivity compared with the bacteria–virus network (part **d**) from the same set of samples[53]. We propose that this reflects a higher specificity in the interactions between bacteria and viruses compared with that between bacteria and protists. Circles represent bacterial automated ribosomal intergenic spacer analysis (ARISA)-based taxa, V-shapes represent T4-like myoviruses and diamonds represent protistan taxa. Solid lines represent positive correlations, dashed lines represent negative correlations and arrows indicate a 1-month delayed correlation. The darkness of the nodes in parts **c** and **d** indicates the number of edges (or correlations) connecting that node (see the key), which is particularly high for some bacteria and viruses in part **d**. Part **a** is from REF. 54, Nature Publishing Group; and parts **c** and **d** are from REF. 53, Nature Publishing Group.

Fingerprinting
A range of molecular genetic approaches that quickly profile the diversity of microbial communities in a sample. Instead of directly identifying individuals by their actual gene sequences, fingerprinting is based on the length of genes or sequence properties (such as the location of a restriction site), and it shows how many variants of a gene are present and in what proportions.

Box 3 | **What level of phylogenetic resolution to use?**

When studying the dynamics of a community, microbial ecologists typically strive to differentiate between organisms that have different functions or meet the definition of an 'ecological species': that is, an organism that occupies a distinct ecological niche. However, there are alternative concepts for defining microbial species, and the nature of microbial speciation and species-level taxonomy has been the subject of much discussion and debate[110–114]. Many microbiologists consider two organisms to belong to the same species if their 16S rRNA sequences show 97% similarity; however, it has been repeatedly noted that distinct ecological species are often more similar than this[9,115,116]. Owing to its universality and relative ease of use and interpretation, high-throughput 16S sequencing is often an informative first step to gain an understanding the composition of a community. However, because 16S rRNA is so highly conserved, many species with distinct ecological characteristics are difficult or impossible to distinguish by typical tag sequencing. For example, distinguishing between the different ecotypes of *Prochlorococcus* (such as those adapted to high-light versus low-light conditions) is not possible with standard 16S rRNA high-throughput methods that use percentage similarity as a cut-off. Furthermore, the genus *Prochlorococcus* and the ecologically distinct genus *Synechococcus* are 98% similar in terms of their 16S rRNA sequences; hence, out of necessity, studies of cyanobacteria have used alternative approaches, such as the sequencing of protein-coding genes or the intergenic spacer (ITS). In addition, there are new, non-cluster-based 16S rRNA analyses that are able to resolve close, but often ecologically distinct, relatives by careful examination of characteristic single-base variants in multiple samples[117,118].

As a prominent illustration of the importance of microdiversity among marine microbial populations, a recent study at the Bermuda Atlantic Time-series Study (BATS) revealed hundreds of co-occurring *Prochlorococcus* strains, some of which were found to be identical by ITS sequencing but differed in terms of the genes encoding proteins involved in nutrient uptake and surface structures (possibly owing to the selective pressures to evade phage infection or grazing). These subpopulations showed seasonal variability, suggesting that their ecological niches differ, even though they belong to the same ITS types[93]. It is likely that *Prochlorococcus* are not exceptional in terms of such microdiversity; a more broadly diverse clade, SAR11, has also been shown to have considerable microdiversity at the genomic level[119,120]. As sequencing costs continue to decline, the use of shotgun metagenomics and single-cell genomics promises to continue to reveal other genomic variants and ecological specializations within lineages through the analysis of suitable highly resolving genes and comparisons of multiple samples or via reverse ecology population genomics[112].

Although oceanic time-series studies are currently the most common applications of correlational network analysis, the application of this approach to spatial distributions has become common in other environments (such as soil[81–83]) and can reveal meaningful correlations between taxa and environmental conditions.

Microbial association networks are a useful and convenient tool to visualize, display and analyse statistical associations[84] between many parameters simultaneously[85–87], including the abundance of members in a microbial community[88,89]. Graphical visualization (FIG. 2) is important because of the difficulty in understanding multiple correlations when observed in tabular form. Association networks usually reveal clusters of co-occurring organisms, which suggests shared niches, symbiosis and other relationships that may be indirect[61,88,90]. Studies at SPOT have shown that association networks display different patterns at different timescales (FIG. 3). For example, in a set of daily measurements at SPOT, the abundances of several SAR11 taxa correlated strongly with each other over several weeks[54] (FIG. 3a); however, these same SAR11 taxa generally did not correlate well with each other on monthly sampling scales over several years[49] (FIG. 3b). This suggests that the typical day-to-day environmental changes result in all of these SAR11 types responding in a similar manner; however, subtle differences — for example, temperature preferences or interactions with other organisms that change considerably from month to month — allow different SAR11 types to dominate at different times of the year. In addition, bacteria–protist networks differ from bacteria–virus networks[53], such that bacteria–virus networks are tightly interconnected, whereas bacteria–protist networks are more loosely connected (FIG. 3c,d). This might be explained by the increased specificity of phages for their bacterial hosts compared with the lower specificity in interactions between bacteria and protists, irrespective of whether the protists are a source of nutrition (for example, phytoplankton) or a predator (for example, bacterivorous flagellata).

Defining an 'ecological species', and why we should care. The observed predictability of community composition within clearly delineated ranges supports the conclusion that most microorganisms occupy well-defined niches, and overall there is limited functional redundancy because most microorganisms are not readily replaced with another species as operationally defined by the methods used (BOX 3). This is further supported by the microbial association networks, in which even close relatives can show different connections to other community members (FIG. 3b). These observations suggest that microorganisms are similar to species of better-known animals and plants, in which close relatives with mostly similar functions constitute different ecological species, in some ways analogous to Darwin's finches[91]. Needless to say, the way that we define microbial taxa or species in a practical sense, and questions about how closely we want to differentiate close relatives (BOX 3), has a huge influence on how we parse and interpret our data. Modest changes can greatly alter conclusions; for example, all marine cyanobacteria would be considered globally distributed if we consider the commonly used 97% 16S rRNA similarity cut-off, but they have distinct spatial and temporal distribution patterns (owing to specific oceanographic and biotic properties), which are revealed by the finer resolution of analyses based on intergenic spacer sequences or other highly resolving sequences[92,93] (BOX 4). It should be noted that what is often called 'redundancy' in modern microbial ecology often refers to shared omic annotation: that is, two genes are considered to be redundant if they have the same or a similar annotation. However, the relevant biological variation that defines species (that is, those that occupy distinct niches) might be controlled by the affinities or kinetics of proteins rather than their basic functions. For example, although two competing species might share most of their genes (in terms of annotations), the species that possesses the transporter with a higher affinity for a potentially limiting nutrient (such as phosphate) would outcompete the other species in low-phosphate conditions, whereas the competitor species would win at higher phosphate concentrations if it has a higher uptake rate. Therefore, for ecological interpretations and relevant characterization of microbial

Tag sequencing
A high-throughput method for determining the microbial community composition in a sample via PCR and sequencing of a portion of the rRNA genes (sometimes other genes).

Lotka–Volterra
The most well-known mathematical model of predator–prey interactions, in which predator and prey usually oscillate with a time offset between them.

Ecotypes
Genetically distinct organisms that are adapted to specific environmental conditions. The term is often used in microbiology to avoid the difficulties of formally defining species.

Box 4 | **Dynamics of marine cyanobacteria: abundance and diversity**

Prochlorococcus and *Synechococcus* are suitable models for investigating the processes that influence the abundance and distribution of bacterial populations in the ocean, as they are globally important (together they contribute ~25% of global photosynthesis) and their environmental distributions are known via flow cytometry[121,122] and molecular methods[123–125]. Different ecotypes within each genus, often distinguished by intergenic spacer (ITS) sequences[92,124], dominate different global oceanic regions and depths[92] predictably enough that global dynamics can be reasonably modelled[126]. As many representatives are cultivable (including their phages), the physiological nature of the ecotypes and both bottom-up controls (resources) and top-down controls (phages and grazing) can be assessed. Even though *Synechococcus* and *Prochlorococcus* are closely related, their ecological niches are vastly different. In general, *Prochlorococcus* dominate in low-nutrient (and often warmer) conditions, whereas *Synechococcus* dominate in coastal and nutrient-rich conditions[127]. Their temporal dynamics also differ, showing varied preferences for seasons and environmental conditions. For example, in the well-studied subtropical Gulf of Aqaba in the Red Sea, a successional pattern occurs where seasonal upwelling triggers rapid growth of eukaryotic phytoplankton[14,128], followed within days to weeks by an increase in *Synechococcus*. *Prochlorococcus* dominate the community following summer stratification[129,130], and during the autumn-to-winter transition *Synechococcus* replace *Prochlorococcus* as water column turbulence increases. At the Bermuda Atlantic Time-series Study (BATS), the successional pattern is similar; however, a peak in *Synechococcus* does not occur in the autumn. At the Hawaii Ocean Time-series (HOT), *Prochlorococcus* dominate all year round[131]. In contrast to oligotrophic regions, in temperate coastal waters (such as Chesapeake, Woods Hole and San Diego), *Synechococcus* are often the dominant cyanobacterium, even in the summer, but display seasonal variability in total abundance, and their abundance is often lowest during the winter[125,132,133]. In Monterey Bay, California, *Synechococcus* show a general negative correlation with the abundance of eukaryotic phytoplankton, which peak during spring upwelling[99]. During summer stratification, *Synechococcus* comprise several temporally variable ecotypes[123]. Daily data from an autumn study showed that *Synechococcus* abundance fluctuates rapidly and is negatively correlated with chlorophyll levels[99].

Beyond cyanobacterial seasonal dynamics at the genus level that are probably related to 'bottom-up' chemical and physical conditions, there are 'top-down' influences from co-occurring and co-evolving cyanophages[132,134–136], which often outnumber their host. The extensive cyanobacterial genomic microdiversity, particularly in genomic islands, is associated with phage resistance; high host microdiversity presents a wide array of 'moving targets' for phages and thus may result in phage resistance at the community level, which may be one way that cyanobacteria avoid having their populations reduced greatly by phage infection[69,70]. Indeed, several microdiverse *Synechococcus* variants[123,137] show seasonal dynamics that often correspond to phage genotype dynamics[136,138], and observations of environmental *Prochlorococcus* phages suggest that they are positively correlated with *Prochlorococcus* abundance[139]. Phages obviously affect these communities, but we are still left with an important question: to what extent and on what scales do environmental cyanophage populations manipulate and change the host strains present, or alternatively, under what conditions (if any) do phages essentially follow the abundance of successful host strains within these communities, without substantially altering the community composition that has been primarily determined by other bottom-up and top-down controls? Recent methodological advances have enabled the identification and enumeration of infective cyanophages in the sea, and following the cyanobacterial and cyanophage dynamics should help to shed light on controls of these populations and the ecological implications[95,140]. It will be some time before we know how well these cyanobacterial systems reflect marine bacteria in general.

traits (which are required to define a species), we often need to go beyond basic annotations that are commonly used in omics today, as finer details often matter.

Conclusions and future prospects

This Review highlights several studies of microbial community dynamics, which together with the consideration of variations in the physical, chemical and biological properties of the environment provide valuable insights into the functioning of marine microbial systems. For example, each distinct microorganism or species (BOX 3) occupies a unique or largely unique niche that defines its relationship with other organisms; marine microbial systems show resilience and long-term stability in terms of their average community composition despite considerable short-term variation; and associations between the various players in the ecosystem over different timescales seem to be driven by different underlying causes. The patterns and interactions that emerge are reminiscent of those that occur among animals and plants, despite variation occurring on shorter timescales for microorganisms. Thus, it is now possible to address major questions from classical ecology about the factors and feedback loops that promote stability or variability in marine microbial networks.

The field is advancing rapidly, especially in terms of the volume of environmental sequences being analysed, and it is challenging for data analysis to keep pace with the generation of data. As is typical for marine microbiology, technological advances will lead the way forward to addressing new versions of timeless questions. For example, how do top-down, bottom-up and sideways (competitive) controls influence microbial communities and what are the functional implications of the diversity and dynamics of microbial communities? Researchers are moving beyond analysis of community composition alone to also integrate data on functional genes, transcripts, proteins and metabolites. Emerging technologies, such as cell sorting, coupled with genomics can provide insights into trophic interactions, population-level diversity and dynamics[94,95]. Continued improvements of ecogenomic sensors[96], which enable unattended sample collection with matching contextual data at high temporal resolution[96], will continue to provide physiological[97,98] and ecological[99] insights. Biomarkers — for

Reverse ecology population genomics
An approach that uses genomic information from metagenomes or isolates to assess gene flow and the nature of populations and ecological units (for example, 'species').

Ecogenomic sensors
Devices that are typically used by autonomous *in situ* sampling systems to detect genome characteristics to facilitate the identification of organisms or other factors (such as nutrients, temperature and others).

example, specific proteins (such as IdiA) or sphingolipids — facilitate the assessment of limiting factors and microbial processes, such as iron stress or viral infection, and will be valuable in exploring functional implications and the forces influencing microbial communities[100–102]. Finally, continued software development for data management and processing is needed to utilize the ever-increasing amount of data (especially omic data) that are generated. In particular, development and application of algorithms to detect multicomponent interactions will help to explore the complexity inherent in microbial communities[88]. Improvements in the assembly, identification and binning of environmental omic data that help to identify the origins of taxa (including the parsing of close relatives), and improvements in functional annotation for well-designed experiments will continue to be major drivers of future progress in this field[103,104].

1. Ducklow, H. W. in *Microbial Ecology of the Oceans* (ed. Kirchman, D. L.) 85–120 (Wiley-Liss Inc., 2000).
2. Suttle, C. A. Marine viruses — major players in the global ecosystem. *Nature Rev. Microbiol.* **5**, 801–812 (2007).
3. Falkowski, P. G., Fenchel, T. & Delong, E. F. The microbial engines that drive earth's biogeochemical cycles. *Science* **320**, 1034–1039 (2008).
4. Zehr, J. P. & Kudela, R. M. Nitrogen cycle of the open ocean: from genes to ecosystems. *Ann. Rev. Mar. Sci.* **3**, 197–225 (2011).
5. Azam, F., Fenchel, T., Field, J. G., Gray, J. S. & Thingstad, T. F. The ecological role of water-column microbes in the sea. *Mar. Ecol. Prog. Ser.* **10**, 257–263 (1983).
6. Pace, N. R., Stahl, D. A., Lane, D. J. & Olsen, G. J. in *Advances in Microbial Ecology* (ed. Marshall, K. C.) 1–55 (Springer, 1986).
7. Sogin, M. L. *et al.* Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proc. Natl Acad. Sci. USA* **103**, 12115–12120 (2006).
8. Caporaso, J. G. *et al.* Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J.* **6**, 1621–1624 (2012).
9. Fuhrman, J. A. & Hagström, Å. in *Microbial Ecology of the Oceans* (ed. Kirchman, D. L.) 45–90 (John Wiley & Sons Inc., 2008).
10. Kuipers, B., van Noort, G. J., Vosjan, J. & Herndl, G. J. Diel periodicity of bacterioplankton in the euphotic zone of the subtropical Atlantic Ocean. *Mar. Ecol. Prog. Ser.* **201**, 13–25 (2000).
11. Fuhrman, J. A., Eppley, R. W., Hagstrom, A. & Azam, F. Diel variations in bacterioplankton, phytoplankton, and related parameters in the Southern California Bight. *Mar. Ecol. Prog. Ser.* **27**, 9–20 (1985).
12. Gasol, J. M. *et al.* Diel variations in bacterial heterotrophic activity and growth in the northwestern Mediterranean Sea. *Mar. Ecol. Prog. Ser.* **164**, 107–124 (1998).
13. Jacquet, S. *et al.* Diel patterns of growth and division in marine picoplankton in culture. *J. Phycol.* **37**, 357–369 (2001).
14. Iluz, D. *et al.* Short-term variability in primary productivity during a wind-driven diatom bloom in the Gulf of Eilat (Aqaba). *Aquat. Microb. Ecol.* **56**, 205–215 (2009).
15. Teeling, H. *et al.* Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**, 608–611 (2012).
This comprehensive study describes the use of classical biological oceanographic research and omics analyses to determine the bacterioplankton response to a phytoplankton bloom in the North Sea.
16. Klindworth, A. *et al.* Diversity and activity of marine bacterioplankton during a diatom bloom in the North Sea assessed by total RNA and pyrotag sequencing. *Mar. Genomics* **18**, 185–192 (2014).
17. Buchan, A., LeCleir, G. R., Gulvik, C. A. & González, J. M. Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Rev. Microbiol.* **12**, 686–698 (2014).
18. Rich, V. I., Pham, V. D., Eppley, J., Shi, Y. & DeLong, E. F. Time-series analyses of Monterey Bay coastal microbial picoplankton using a "genome proxy" microarray. *Environ. Microbiol.* **13**, 116–134 (2011).
19. Venrick, E. L., McGowan, J. A., Cayan, D. R. & Hayward, T. L. Climate and chlorophyll *a*: long-term trends in the central north Pacific Ocean. *Science* **238**, 70–72 (1985).
20. Hutchins, D. A., Mulholland, M. R. & Feixue, F. Nutrient cycles and marine microbes in a CO_2-enriched ocean. *Oceanography* **22**, 128–145 (2009).
21. Karl, D. M. & Church, M. J. Microbial oceanography and the Hawaii Ocean Time-series programme. *Nature Rev. Microbiol.* **12**, 1–15 (2014).
22. Diaz, R. J. & Rosenberg, R. Spreading dead zones and consequences for marine ecosystems. *Science* **321**, 926–929 (2008).
23. Hendy, I. L. & Kennett, J. P. Dansgaard–Oeschger cycles and the California Current System: planktonic foraminiferal response to rapid climate change in Santa Barbara Basin, Ocean Drilling Program hole 893A. *Paleoceanography* **15**, 30–42 (2000).
24. Kennett, J. P. & Venz, K. Late Quaternary climatically related planktonic foraminiferal assemblage changes: Hole 893A, Santa Barbara Basin, California. *Proc. ODP Sci. Rep.* **146**, 281–293 (1995).
25. Finkel, Z. V., Katz, M. E., Wright, J. D., Schofield, O. M. E. & Falkowski, P. G. Climatically driven macroevolutionary patterns in the size of marine diatoms over the Cenozoic. *Proc. Natl Acad. Sci. USA* **102**, 8927–8932 (2005).
26. Finkel, Z. V. *et al.* A universal driver of macroevolutionary change in the size of marine phytoplankton over the Cenozoic. *Proc. Natl Acad. Sci. USA* **104**, 20416–20420 (2007).
27. Michaels, A. F. *et al.* Seasonal patterns of ocean biogeochemistry at the U.S. JGOFS Bermuda Atlantic time-series study site. *Deep Sea Res. Part I Oceanogr. Res. Pap.* **41**, 1013–1038 (1994).
28. Morris, R. M. *et al.* Temporal and spatial response of bacterioplankton lineages to annual convective overturn at the Bermuda Atlantic Time-series Study site. *Limnol. Oceanogr.* **50**, 1687–1696 (2005).
29. Treusch, A. H. *et al.* Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J.* **3**, 1148–1163 (2009).
30. Morris, R. M., Frazar, C. D. & Carlson, C. A. Basin-scale patterns in the abundance of SAR11 subclades, marine Actinobacteria (OM1), members of the Roseobacter clade and OCS116 in the South Atlantic. *Environ. Microbiol.* **14**, 1133–1144 (2012).
31. Giovannoni, S. J. & Vergin, K. L. Seasonality in ocean microbial communities. *Science* **335**, 671–676 (2012).
This thoughtful review examines the dynamics of the predominant phylogentic groups of microorganisms at the major time-series sites of BATS, HOTS and SPOT.
32. Vergin, K. L. *et al.* High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J.* **7**, 1322–1332 (2013).
33. Vergin, K. L., Done, B., Carlson, C. & Giovannoni, S. J. Spatiotemporal distributions of rare bacterioplankton populations indicate adaptive strategies in the oligotrophic ocean. *Aquat. Microb. Ecol.* **71**, 1–13 (2013).
A detailed study showing how relatively rare bacterial taxa at BATS have distinctive dynamics, which probably reflects adaptations.
34. Steinberg, D. K. *et al.* Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. *Deep Sea Res. Part II* **48**, 1405–1447 (2001).
35. Sverdrup, H. U. On conditions for the vernal blooming of phytoplankton. *J. Cons. Int. Explor. Mer.* **18**, 287–295 (1953).
36. Ryther, J. H. & Menzel, D. W. Primary production in the southwest Sargasso Sea, January–February 1960. *Bull. Mar. Sci. Gulf Caribb.* **11**, 381–388 (1961).
37. Menzel, D. W. & Ryther, J. H. Annual variations in primary production of the Sargasso Sea off Bermuda. *Deep Sea Res.* **7**, 282–288 (1961).
38. DuRand, M. D., Olson, R. J. & Chisholm, S. W. Phytoplankton population dynamics at the Bermuda Atlantic Time-series station in the Sargasso Sea. *Deep Sea Res. Part II* **48**, 1983–2003 (2001).
39. Parsons, R. J., Breitbart, M., Lomas, M. W. & Carlson, C. A. Ocean time-series reveals recurring seasonal patterns of virioplankton dynamics in the northwestern Sargasso Sea. *ISME J.* **6**, 273–284 (2012).
40. Ewart, C. S., Meyers, M. K., Wallner, E. R., McGillicuddy, D. J. & Carlson, C. A. Microbial dynamics in cyclonic and anticyclonic mode-water eddies in the northwestern Sargasso Sea. *Deep Sea Res. Part II* **55**, 1334–1347 (2008).
41. Sweeney, E. N., McGillicuddy, D. J. & Buesseler, K. O. Biogeochemical impacts due to mesoscale eddy activity in the Sargasso Sea as measured at the Bermuda Atlantic Time-series Study (BATS). *Deep Sea Res. Part II* **50**, 3017–3039 (2003).
42. Nelson, C. E., Carlson, C. A., Ewart, C. S. & Halewood, E. R. Community differentiation and population enrichment of Sargasso Sea bacterioplankton in the euphotic zone of a mesoscale mode-water eddy. *Environ. Microbiol.* **16**, 871–887 (2014).
43. Karl, D. M. & Lukas, R. The Hawaii Ocean Time-series (HOT) program: background, rationale and field implementation. *Deep Sea Res. Part II* **43**, 129–156 (1996).
44. Church, M. J., Lomas, M. W. & Muller-Karger, F. Sea change: charting the course for biogeochemical ocean time-series research in a new millennium. *Deep Sea Res. Part II* **93**, 2–15 (2013).
45. Eiler, A., Hayakawa, D. H. & Rappé, M. S. Non-random assembly of bacterioplankton communities in the subtropical north pacific ocean. *Front. Microbiol.* **43**, 129–156 (2011).
46. Chow, C.-E. T. *et al.* Temporal variability and coherence of euphotic zone bacterial communities over a decade in the Southern California Bight. *ISME J.* **7**, 2259–2273 (2013).
47. Hickey, B. M. Circulation over the Santa Monica-San Pedro Basin and Shelf. *Prog. Oceanogr.* **30**, 37–115 (1992).
48. Fuhrman, J. A. *et al.* Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc. Natl Acad. Sci. USA* **103**, 13104–13109 (2006).
49. Fuhrman, J. A. & Steele, J. A. Community structure of marine bacterioplankton: patterns, networks, and relationships to function. *Aquat. Microb. Ecol.* **53**, 69–81 (2008).
50. Hatosy, S. M. *et al.* Beta diversity of marine bacteria depends on temporal scale. *Ecology* **94**, 1898–1904 (2013).
This meta-analysis uses data from SPOT and the western English Channel to examine the determinants, rate and scales of temporal variation of microbial communities at these sites.
51. Filée, J., Tetart, F., Suttle, C. A. & Krisch, H. M. Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *Proc. Natl Acad. Sci. USA* **102**, 12471–12476 (2005).
52. Chow, C.-E. T. & Fuhrman, J. A. Seasonality and monthly dynamics of marine myovirus communities. *Environ. Microbiol.* **14**, 2171–2183 (2012).
53. Chow, C.-E. T., Kim, D. Y., Sachdeva, R., Caron, D. A. & Fuhrman, J. A. Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J.* **8**, 816–829 (2014).
54. Needham, D. M. *et al.* Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *ISME J.* **7**, 1274–1285 (2013).
Through daily sampling and molecular genetic analysis of surface waters near SPOT, this study reveals that most bacteria and viral taxa vary around a mean community composition over weeks to months, with many correlations among taxa, whereas modest changes occur on a daily scale.

55. Beman, J. M., Sachdeva, R. & Fuhrman, J. A. Population ecology of nitrifying archaea and bacteria in the Southern California Bight. *Environ. Microbiol.* **12**, 1282–1292 (2010).
56. Cram, J. A. *et al.* Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J.* http://dx.doi.org/10.1038/ismej.2014.153 (2014).
This study examines five depths at SPOT and shows that, similar to the surface waters, the deepest waters exhibit seasonal changes.
57. Hewson, I., Steele, J. A., Capone, D. G. & Fuhrman, J. A. Remarkable heterogeneity in meso- and bathypelagic bacterioplankton assemblage composition. *Limnol. Oceanogr.* **51**, 1274–1283 (2006).
58. Beman, J. M., Steele, J. A. & Fuhrman, J. A. Co-occurrence patterns for abundant marine archaeal and bacterial lineages in the deep chlorophyll maximum of coastal California. *ISME J.* **5**, 1077–1085 (2011).
59. Pingree, R. D. & Griffiths, D. K. Tidal fronts on the shelf seas around the British Isles. *J. Geophys. Res.* **83**, 4615–4622 (1978).
60. Southward, A. J. *et al.* Long-term oceanographic and ecological research in the Western English Channel. *Adv. Mar. Biol.* **47**, 1–105 (2005).
61. Gilbert, J. A. *et al.* Defining seasonal marine microbial community dynamics. *ISME J.* **6**, 298–308 (2012).
This 10-year study of the western English Channel shows the strongest seasonality of any of the long-term time series (> 5 years), and shows that seasonal changes, not trophic interactions, can explain most of the variation that occurs.
62. Gilbert, J. A. *et al.* The seasonal structure of microbial communities in the Western English Channel. *Environ. Microbiol.* **11**, 3132–3139 (2009).
63. Duarte, C. M., Agustí, S., Kennedy, H. & Vaqué, D. The Mediterranean climate as a template for Mediterranean marine ecosystems: the example of the northeast Spanish littoral. *Prog. Oceanogr.* **44**, 245–270 (1999).
64. D'Ortenzio, F. *et al.* Seasonal variability of the mixed layer depth in the Mediterranean Sea as derived from *in situ* profiles. *Geophys. Res. Lett.* **32**, 1–4 (2005).
65. Schauer, M., Balagué, V., Pedrós-Alió, C. & Massana, R. Seasonal changes in the taxonomic composition of bacterioplankton in a coastal oligotrophic system. *Aquat. Microb. Ecol.* **31**, 163–174 (2003).
66. Alonso-Sáez, L. *et al.* Seasonality in bacterial diversity in north-west Mediterranean coastal waters: assessment through clone libraries, fingerprinting and FISH. *FEMS Microbiol. Ecol.* **60**, 98–112 (2007).
67. Galand, P. E., Gutiérrez-Provecho, C., Massana, R., Gasol, J. M. & Casamayor, E. O. Inter-annual recurrence of archaeal assemblages in the coastal NW Mediterranean Sea (Blanes Bay Microbial Observatory). *Limnol. Oceanogr.* **55**, 2117–2125 (2010).
68. Hugoni, M. *et al.* Structure of the rare archaeal biosphere and seasonal dynamics of active ecotypes in surface coastal waters. *Proc. Natl Acad. Sci. USA* **110**, 6004–6009 (2013).
One of the few studies that examines the dynamics of marine archaeal communities, using both genomics and transcriptomics.
69. Avrani, S., Wurtzel, O., Sharon, I., Sorek, R. & Lindell, D. Genomic island variability facilitates *Prochlorococcus*–virus coexistence. *Nature* **474**, 604–608 (2011).
70. Marston, M. F. *et al.* Rapid diversification of coevolving marine *Synechococcus* and a virus. *Proc. Natl Acad. Sci. USA* **109**, 4544–4549 (2012).
71. Sullivan, M. B., Waterbury, J. B. & Chisholm, S. W. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **424**, 1047–1052 (2003).
72. Morris, J. J., Lenski, R. E. & Zinser, E. R. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *mBio* **3**, e00036-12 (2012).
73. Morris, J. J., Johnson, Z. I., Szul, M. J., Keller, M. & Zinser, E. R. Dependence of the cyanobacterium *Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the ocean's surface. *PLoS ONE* **6**, e16805 (2011).
74. Morris, J. J., Papoulis, S. E. & Lenski, R. E. Coexistence of evolving bacteria stabilized by a shared black queen function. *Evolution* **68**, 2960–2971 (2014).
75. Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K. & Schloss, P. D. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* **79**, 5112–5120 (2013).
76. Bartram, A. K., Lynch, M. D. J., Stearns, J. C., Moreno-Hagelsieb, G. & Neufeld, J. D. Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads. *Appl. Environ. Microbiol.* **77**, 3846–3852 (2011).
77. Fuhrman, J. A. Microbial community structure and its functional implications. *Nature* **459**, 193–199 (2009).
78. Morris, B. E. L., Henneberger, R., Huber, H. & Moissl-Eichinger, C. Microbial syntrophy: interaction for the common good. *FEMS Microbiol. Rev.* **37**, 384–406 (2013).
79. Petanidou, T., Kallimanis, A. S., Tzanopoulos, J., Sgardelis, S. P. & Pantis, J. D. Long-term observation of a pollination network: fluctuation in species and interactions, relative invariance of network structure and implications for estimates of specialization. *Ecol. Lett.* **11**, 564–575 (2008).
80. Kara, E. L., Hanson, P. C., Hu, Y. H., Winslow, L. & McMahon, K. D. A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, Wisconsin, USA. *ISME J.* **7**, 680–684 (2013).
81. Barberán, A., Bates, S. T., Casamayor, E. O. & Fierer, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* **6**, 343–351 (2012).
82. De Menezes, A. B. *et al.* Network analysis reveals that bacteria and fungi form modules that correlate independently with soil parameters. *Environ. Microbiol.* http://dx.doi.org/10.1111/1462-2920.12559 (2014).
83. Lupatini, M. *et al.* Network topology reveals high connectance levels and few key microbial genera within soils. *Front. Environ. Sci.* **2**, 10 (2014).
84. Ruan, Q. *et al.* Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. *Bioinformatics* **22**, 2532–2538 (2006).
85. Xia, L. C. *et al.* Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. *BMC Syst. Biol.* **5** (Suppl. 2), S15 (2011).
86. Xia, L. C., Ai, D., Cram, J. A., Fuhrman, J. A. & Sun, F. Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. *Bioinformatics* **29**, 230–237 (2013).
87. Durno, W. E., Hanson, N. W., Konwar, K. M. & Hallam, S. J. Expanding the boundaries of local similarity analysis. *BMC Genomics* **14** (Suppl. 1), S3 (2013).
88. Faust, K. & Raes, J. Microbial interactions: from networks to models. *Nature Rev. Microbiol.* **10**, 538–550 (2012).
89. Cram, J. A., Sun, F. & Fuhrman, J. A. Marine bacterial, archaeal, and protistan association networks *Encyclopedia of Metagenomics* http://dx.doi.org/10.1007/978-1-4614-6418-1_721-3 (2014).
90. Steele, J. A. *et al.* Marine bacterial, archaeal and protistan association networks reveal ecological linkages. *ISME J.* **5**, 1414–1425 (2011).
The first example of a three-domain microbial association network study, which shows multiple associations (presumably including trophic interactions) and network properties both within and between the domain members.
91. Grant, P. R. *Ecology and Evolution of Darwin's Finches* (Princeton Univ. Press, 1986).
92. Johnson, Z. I. *et al.* Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**, 1737–1740 (2006).
93. Kashtan, N. *et al.* Single-cell genomics reveals hundreds of coexisting subpopulations in wild *Prochlorococcus*. *Science* **344**, 416–420 (2014).
94. Yoon, H. S. *et al.* Single-cell genomics reveals organismal interactions in uncultivated marine protists. *Science* **332**, 714–717 (2011).
95. Deng, L. *et al.* Viral tagging reveals discrete populations in *Synechococcus* viral genome sequence space. *Nature* **513**, 242–245 (2014).
96. Scholin, C. A. What are "ecogenomic sensors"? A review and thoughts for the future. *Ocean Sci.* **6**, 51–60 (2010).
97. Ottesen, E. A. *et al.* Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science* **345**, 207–212 (2014).
98. Ottesen, E. A. *et al.* Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc. Natl Acad. Sci. USA* **110**, E488–E497 (2013).
99. Robidart, J. C. *et al.* Seasonal *Synechococcus* and *Thaumarchaeal* population dynamics examined with high resolution with remote *in situ* instrumentation. *ISME J.* **6**, 513–523 (2012).
This field study used an autonomous *in situ* sample collection and processing instrument to show the detailed dynamics of important cyanobacterial and archaeal lineages in Monterey Bay.
100. Vardi, A. *et al.* Viral glycosphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* **326**, 861–865 (2009).
101. Webb, E. A., Moffett, J. W. & Waterbury, J. B. Iron stress in open-ocean cyanobacteria (*Synechococcus*, *Trichodesmium*, and *Crocosphaera*) identification of the IdiA protein. *Appl. Environ. Microbiol.* **67**, 5444–5452 (2001).
102. Martínez Martínez, J., Poulton, N. J., Stepanauskas, R., Sieracki, M. E. & Wilson, W. H. Targeted sorting of single virus-infected cells of the coccolithophore *Emiliania huxleyi*. *PLoS ONE* **6**, e22520 (2011).
103. Albertsen, M. *et al.* Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nature Biotech.* **31**, 533–538 (2013).
104. Howe, A. C. *et al.* Tackling soil diversity with the assembly of large, complex metagenomes. *Proc. Natl Acad. Sci. USA* **111**, 4904–4909 (2014).
105. Hewson, I., Steele, J. A., Capone, D. G. & Fuhrman, J. A. Temporal and spatial scales of variation in bacterioplankton assemblages of oligotrophic surface waters. *Mar. Ecol. Prog. Ser.* **311**, 67–77 (2006).
106. Lie, A. A. Y., Kim, D. Y., Schnetzer, A. & Caron, D. A. Small-scale temporal and spatial variations in protistan community composition at the San Pedro Ocean Time-series station off the coast of southern California. *Aquat. Microb. Ecol.* **70**, 93–110 (2013).
107. Ramette, A. Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol.* **62**, 142–160 (2007).
108. Sapp, M., Wichels, A., Wiltshire, K. H. & Gerdts, G. Bacterial community dynamics during the winter-spring transition in the North Sea. *FEMS Microbiol. Ecol.* **59**, 622–637 (2007).
109. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445 (2003).
110. Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G. & Hanage, W. P. The bacterial species challenge: making sense of genetic and ecological diversity. *Science* **323**, 741–746 (2009).
111. Cohan, F. M. & Perry, E. B. A systematics for discovering the fundamental units of bacterial diversity. *Curr. Biol.* **17**, R373–R386 (2007).
112. Shapiro, B. J. & Polz, M. F. Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol.* **22**, 235–247 (2014).
113. Achtman, M. & Wagner, M. Microbial diversity and the genetic nature of microbial species. *Nature Rev. Microbiol.* **6**, 431–440 (2008).
114. Doolittle, W. F. Population genomics: how bacterial species form and why they don't exist. *Curr. Biol.* **22**, R451–R453 (2012).
115. Staley, J. T. The bacterial species dilemma and the genomic–phylogenetic species concept. *Philos. Trans. R. Soc. B* **361**, 1899–1909 (2006).
116. Hunt, D. E. *et al.* Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**, 1081–1085 (2008).
117. Eren, A. M. *et al.* Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol. Evol.* **4**, 1111–1119 (2013).
118. Tikhonov, M., Leach, R. W. & Wingreen, N. S. Interpreting 16S metagenomic data without clustering to achieve sub-OTU resolution. *ISME J.* **9**, 68–80 (2015).
119. Schwalbach, M. S., Tripp, H. J., Steindler, L., Smith, D. P. & Giovannoni, S. J. The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ. Microbiol.* **12**, 490–500 (2010).
120. Wilhelm, L. J., Tripp, H. J., Givan, S. A., Smith, D. P. & Giovannoni, S. J. Natural variation in SAR11 marine bacterioplankton genomes inferred from metagenomic data. *Biol. Direct* **2**, 27 (2007).

121. Olson, R. J., Vaulot, D. & Chisolm, S. W. Marine phytoplankton distributions measured using shipboard flow cytometry. *Deep Sea Res.* **32**, 1273–1280 (1985).
122. Chisolm, S. W. *et al.* A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**, 340–343 (1988).
123. Paerl, R. W., Turk, K. A., Beinart, R. A., Chavez, F. P. & Zehr, J. P. Seasonal change in the abundance of *Synechococcus* and multiple distinct phylotypes in Monterey Bay determined by rbcL and narB quantitative PCR. *Environ. Microbiol.* **14**, 580–593 (2012).
124. Ahlgren, N. A. & Rocap, G. Diversity and distribution of marine *Synechococcus*: multiple gene phylogenies for consensus classification and development of qPCR assays for sensitive measurement of clades in the ocean. *Front. Microbiol.* **3**, 213 (2012).
125. Tai, V. & Palenik, B. Temporal variation of *Synechococcus* clades at a coastal Pacific Ocean monitoring site. *ISME J.* **3**, 903–915 (2009).
126. Follows, M. J. & Dutkiewicz, S. Modeling diverse communities of marine microbes. *Ann. Rev. Mar. Sci.* **3**, 427–451 (2011).
127. Partensky, F., Blanchot, J. & Vaulot, D. in *Marine Cyanobacteria* Vol. 19 (eds Charpy, L. & Larkum, A. W. D.) 457–475 (Institut Océanographique, 1999).
128. Labiosa, R. G., Arrigo, K. R., Genin, A., Monismith, S. G. & van Dijken, G. The interplay between upwelling and deep convective mixing in determining the seasonal phytoplankton dynamics in the Gulf of Aqaba: evidence from SeaWiFS and MODIS. *Limnol. Oceanogr.* **48**, 2355–2368 (2003).
129. Lindell, D. & Post, A. F. Ultraphytoplankton succession is triggered by deep winter mixing in the Gulf of Aqaba (Eilat), Red Sea. *Limnol. Oceanogr.* **40**, 1130–1141 (1995).
130. Al-Najjar, T., Badran, M. I., Richter, C., Meyerhoefer, M. & Sommer, U. Seasonal dynamics of phytoplankton in the Gulf of Aqaba, Red Sea. *Hydrobiologia* **579**, 69–83 (2006).
131. Malmstrom, R. R. *et al.* Temporal dynamics of *Prochlorococcus* ecotypes in the Atlantic and Pacific oceans. *ISME J.* **4**, 1252–1264 (2010).
A detailed study of the dynamics of the most abundant clade of marine phytoplankton, which is found in both major ocean basins.
132. Waterbury, J. B. & Valois, F. W. Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl. Environ. Microbiol.* **59**, 3393–3399 (1993).
A classic study of multiple *Synechococcus* isolates and their phages, which shows that the dominant phage at any given time does not generally infect the *Synechococcus* strain that is dominant at that time.
133. Wang, K., Wommack, K. E. & Chen, F. Abundance and distribution of *Synechococcus* spp. and cyanophages in the Chesapeake Bay. *Appl. Environ. Microbiol.* **77**, 7459–7468 (2011).
134. Chen, F. *et al.* Diverse and dynamic populations of cyanobacterial podoviruses in the Chesapeake Bay unveiled through DNA polymerase gene sequences. *Environ. Microbiol.* **11**, 2884–2892 (2009).
135. Marston, M. F. & Sallee, J. L. 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 (2003).
136. Clasen, J. L. *et al.* Diversity and temporal dynamics of Southern California coastal marine cyanophage isolates. *Aquat. Microb. Ecol.* **69**, 17–31 (2013).
137. Post, A. F. *et al.* Long term seasonal dynamics of *Synechococcus* population structure in the Gulf of Aqaba, Northern Red Sea. *Front. Microbiol.* **2**, 131 (2011).
138. Mühling, M. *et al.* Genetic diversity of marine *Synechococcus* and co-occurring cyanophage communities: evidence for viral control of phytoplankton. *Environ. Microbiol.* **7**, 499–508 (2005).
139. Dekel-Bird, N. P., Sabehi, G., Mosevitzky, B. & Lindell, D. Host-dependent differences in abundance, composition and host range of cyanophages from the Red Sea. *Environ. Microbiol.* http://dx.doi.org/10.1111/1462-2920.12569 (2014).
140. Deng, L. *et al.* Contrasting life strategies of viruses that infect photo- and heterotrophic bacteria, as revealed by viral tagging. *mBio* **3**, e00373-12 (2012).

Acknowledgements
The authors thank present and past members of the Fuhrman Lab, especially C. Chow, and three anonymous reviewers. This work was supported by the NSF grants 1031743 and 1136818, and the Gordon and Betty Moore Foundation through Grant GBMF3779 to J.A.F.

Competing interests statement
The authors declare no competing interests.