Year-round dynamics of amplicon sequence variant communities differ among eukaryotes, *Mimiviridae*, and prokaryotes in a coastal ecosystem

Florian Prodinger,[a] Hisashi Endo,[a] Yoshihito Takano,[b] Yanze Li,[a] Kento Tominaga,[c] Tatsuhiro Isozaki,[c] Romain Blanc-Mathieu,[a,d] Yasuhiro Gotoh,[e] Hayashi Tetsuya,[e] Etsunori Taniguchi,[f] Keizo Nagasaki,[b] Takashi Yoshida,[c] and Hiroyuki Ogata[a]#

[a] Institute for Chemical Research, Kyoto University, Gokasho, Uji, Japan

[b] Faculty of Science and Technology, Kochi University, Nankoku, Kochi, Japan

[c] Graduate School of Agriculture, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto, Japan

[d] Laboratoire de Physiologie Cellulaire & Végétale, CEA, Univ. Grenoble Alpes, CNRS, INRA, IRIG, Grenoble, France

[e] Department of Bacteriology, Faculty of Medical Sciences, Kyushu University, Maidashi, Higashi-ku, Fukuoka, Japan;

[f] Kochi Prefectural Fisheries Research Institute, Susakishi, Kochi, Japan;

Running Head: Year-round microbe dynamics in a coastal ecosystem

#Address correspondence to Hiroyuki Ogata, ogata@kuicr.kyoto-u.ac.jp.

Abstract word count: 199

Main text word count: 6102

Abstract

Coastal seawater is the habitat of diverse microbial communities. These communities are affected by seasonal environmental changes and fluctuating nutrient availability, as well as competitive and cooperative interspecific interactions. The complex interplay of these factors affects elemental cycles and therefore, the environment, which in turn affects microbial communities. In this work, we investigated the seasonal dynamics of communities of eukaryotes, a major group of double-stranded DNA viruses infecting eukaryotes (i.e., *Mimiviridae*), as well as prokaryotes in the Uranouchi Inlet, Kochi, Japan. This inlet harbors several microalgae that recurrently form blooms throughout the year. We performed metabarcoding using ribosomal RNA genes and a *Mimiviridae* DNA polymerase gene as marker genes and counted cells of major algal species in 43 seawater samples collected during 20 months. Communities of eukaryotes, *Mimiviridae*, and prokaryotes characterized at the amplicon sequence variant (ASV) level showed similar seasonal cycles but differences in several aspects such as the recovery rate after a year. The ASV persistence level differed among eukaryote, *Mimiviridae*, and prokaryote communities. *Mimiviridae* ASVs were less persistent than were eukaryotic ASVs, and prokaryotic ASVs were the most persistent. These observations are discussed in the context of survival strategies of the major microbes in these three communities.

Importance

Microbial community structure changes over time through interactions among microbes and environmental factors. We investigated the community dynamics of microbial eukaryotes, giant DNA viruses infecting eukaryotes (i.e., *Mimiviridae*), and prokaryotes in an enclosed coastal inlet by year-round observations. The community dynamics of the three microbial groups

showed similar seasonal cycling patterns but differed in their speed of change and recovery rate. The differences paralleled the distinct recurrence/persistence of operational taxonomic units (i.e., amplicon sequence variants, ASVs) among the microbe groups. Therefore, the length of the "memory" of ASV community structure differs among the three microbial groups in the study area.

Introduction

Microbes – cellular microorganisms and viruses – in marine ecosystems are responsible for biogeochemical cycling (1–3), which is driven by diverse members of microbial communities. The composition of microbial communities changes over time and space in response to environmental factors and through species interactions. Therefore, to understand a marine ecosystem, it is fundamentally important to study the structures of microbial communities, their dynamics, and the mechanism that drives the changes in their communities in the ecosystem.

Previous studies have explored the temporal dynamics of eukaryotic (4–6), prokaryotic (1, 7) (8–11), both eukaryotic and prokaryotic (12, 13), and viral communities (14). Other studies have investigated co-variations among these cellular communities (15–17) or cellular and viral communities (18–24). Overall, these studies have revealed similarities and differences in community dynamics among different microbe groups. Yearly cycles and the recurrence of certain operational taxonomy units at 12-month intervals have been detected for eukaryotes in the north western Mediterranean Sea (6), prokaryotes in the English Channel (1, 7), and small viruses at a near-coastal site of southern California (14) through long-term observations. A short-term study that compared prokaryotic and eukaryotic communities in Massachusetts Bay revealed a higher rate of decline in community similarity for eukaryotes than for prokaryotes (16). Furthermore, changes in abiotic factors were found to have a lower impact on the eukaryotic community than on prokaryotic communities in the Mediterranean Sea (6).

In this study, we investigated the community structures of microbial eukaryotes, viruses in the family *Mimiviridae* (phylum *Nucleocytoviricota*), and prokaryotes in the Uranouchi Inlet in Kochi prefecture, Japan. The Uranouchi Inlet is a semi-enclosed, eutrophic inlet, which is located at the south eastern side of Shikoku Island, Japan. The climate of Shikoku Island is

humid subtropical and shows clear seasonality, with contrasting high and low water temperatures in summer and winter, respectively. The contrasting seasons provide abiotic conditions that change throughout the year. The inlet is known for recurrent harmful blooms of eukaryotic microalgae including *Karenia mikimotoi* (Dinophyceae), *Chattonella* spp. (Raphidophyceae), and *Heterosigma akashiwo* (Raphidophyceae), which occasionally damage the local fishery and aquaculture. No previous studies have investigated the dynamics of the microbial eukaryotic community in this inlet.

Viruses are thought to be important top-down regulators of communities of microorganisms, including eukaryotes. For example, the collapse of eukaryotic phytoplankton blooms is associated with viruses (25–28). The sudden increase of virions during the period of bloom decline suggests that viruses are top-down regulators of phytoplankton populations as well (29, 30). The family *Mimiviridae* is a major group of large DNA viruses that infect diverse eukaryotes in marine environments (31). A recent study revealed an association between eukaryote communities and *Mimiviridae* in a global metagenomic dataset (32). Another study detected cooccurring pairs of eukaryotes and large DNA viruses including members of the *Mimiviridae* through time-series sampling for more than 1 year in Northern Skagerrak, South Norway (18). These spatial and temporal associations highlight the tight ecological interactions between diverse *Mimiviridae* and eukaryotes, that is, the obvious dependance of *Mimiviridae* replication on the activity of their hosts and the possible top-down regulatory roles of these viruses on the dynamics of the host community as a consequence of lytic infection. However, long-term studies on the associations between these viruses and eukaryotes over time are still limited. We recently established a highly sensitive metabarcoding method for *Mimiviridae*, which successfully detected diverse *Mimiviridae* in the Uranouchi Inlet (33). In this study, the

community of *Mimiviridae* was analyzed with this method and compared with communities of microbial eukaryotes to better understand their interactions.

We also investigated the dynamics of the prokaryotic community. *Mimiviridae* and eukaryotes directly interact with each other through infection, while eukaryotes and prokaryotes interact with each other either directly by grazing and parasitism, or indirectly by metabolization of cell debris and nutrient-rich cellular exudates. An example of the indirect interaction between eukaryotic and prokaryotic communities can be seen during algal blooms. The decline of algal blooms is associated with the release of cellular debris and other organic substrates (29), which become important resources for heterotrophic prokaryotes (34). Therefore, we considered that the prokaryotic community serves as a contrasting reference for the comparison of communities between directly interacting eukaryotes and *Mimiviridae*.

We performed metabarcoding of eukaryotes, prokaryotes, and *Mimiviridae* using 18S and 16S rRNA genes and the *Mimiviridae polB* gene as marker genes, based on 43 seawater samples collected during 20 months from January 2017 to September 2018. Microbial communities were characterized at the amplicon sequence variant (ASV) level, which is the highest possible resolution for operational taxonomic grouping of amplicon sequences. We characterized the similarities and differences in community dynamics among eukaryotes, *Mimiviridae*, and prokaryotes.

Results

Seasonal variations in abiotic and biotic environmental factors. Abiotic factors (i.e., temperature and salinity) and the concentrations of nutrients (i.e., nitrogen species and phosphate), dissolved oxygen, and chlorophyll *a* were recorded over 6 years (Fig. S1).

Temperature and salinity showed the most pronounced seasonality among all recorded biotic and abiotic factors. The differences in seasonality were less pronounced for other factors, although the concentrations of nutrients and chlorophyll *a* were higher in summer and autumn than in winter and spring (Fig. S2).

Seasonal variation of bloom-forming phytoplankton. The abundance of six harmful algal bloom (HAB) -forming phytoplankton (i.e., *Heterocapsa circularisquama* (Dinophyceae), *K. mikimotoi* (Dinophyceae), *Cochlodinium polykrycoides* (Dinophyceae), *Chattonella* spp. (Raphidophyceae), *H. akashiwo* (Raphidophyceae), and *Pseudochattonella verruculosa* (Ochrophyta, Dictyochophyceae)) were recorded by microscopic cell counting from 2013 to 2019. Blooms of three species (i.e., *H. akashiwo*, *K. mikimotoi*, *Chattonella* spp.) occurred in the inlet and showed an approximately seasonal pattern: *H. akashiwo* bloomed in late spring (April or May), *K. mikimotoi* bloomed in early summer (~June), and *Chattonella* spp. bloomed in the middle of summer (~July). This pattern was observed in 5 of the 6 years we monitored (Fig. S2M). In 2014, 2015, and 2018, a second *H. akashiwo* bloom occurred (Fig. S1G). The highest daily cell counts of these species were in 2018 or 2019 (Fig. S1G). Blooms of *H. circularisquama*, *C. polykrycoides*, and *P. verruculosa* did not occur every year and their maximum abundances were relatively low (29–4,900 cells·mL^{-1}) (Fig. S1G, and Fig. S2M).

Overview of sequence data. We generated from 36 to 43 metabarcoding datasets for four different communities (i.e., eukaryotic NanoPlus (>3 µm) and PicoNano (0.8–3 µm), prokaryotes (0.2–0.8 µm), and *Mimiviridae* (0.2–0.8 µm)) from a total of 43 sample sets. After discarding datasets with <8,000 reads, 36 to 41 datasets were retained for each community (i.e., eukaryotic NanoPlus community: 41; eukaryotic PicoNano community: 36; prokaryotes: 38; *Mimiviridae*: 41). These metabarcoding datasets were composed of 5.2 million eukaryotic reads, 3.4 million

prokaryotic reads, and 5.7 million *Mimiviridae* reads (Table 1). The subsampled datasets showed similar curve shapes and slopes at 8,000 reads, indicating sufficient sequencing depth for all four communities (Fig. S3).

The eukaryotic reads were grouped into 6,413 ASVs. Only 1,501 ASVs (23.4%) were represented in both NanoPlus (3,356 ASVs) and PicoNano (4,559 ASVs) datasets, but those ASVs accounted for 97.3% and 90.1% of total reads in the NanoPlus and PicoNano datasets, respectively. The *Mimiviridae* reads were grouped into 6,261 ASVs, whilst the prokaryotic reads were grouped into 4,478 ASVs. The subsampled datasets also showed relatively low richness for eukaryotes in the NanoPlus size fraction (1,486 ASVs) compared with eukaryotes in the PicoNano size fraction (2,184 ASVs), prokaryotes (2,383 ASVs) and *Mimiviridae* (3,693 ASVs). Shannon's diversity fluctuated throughout the sampling period (Fig. S4). The eukaryotic community in the NanoPlus size fraction was less diverse than other communities during most of the sampling period (Table 1, Fig. S4).

Composition of eukaryote, prokaryote, and *Mimiviridae* communities. Eukaryotic ASVs were classified into 16 groups (from phylum to genus levels) as shown in Figure 1. These 16 groups accounted for 98.7% of the eukaryotic reads from the NanoPlus dataset (Fig. 1A). The group with the highest relative read abundance was Metazoa (28.6%, 625 ASVs), followed by Diatomea (25.1%, 234 ASVs) and Dinoflagellata (22.5% including Dinophyceae, 469 ASVs). These three groups accounted for 77.0% of total reads. Each of the other 13 groups accounted for 0.06% to 6.6% of total reads. The relative abundance of ASVs of HAB-forming genera and families, namely *Heterosigma*, *Chattonella*, and Kareniaceae, was 0.26%, 2.1%, and 5.8%, respectively (Fig. 1A and Fig. 1B). Each of these groups was dominated by a single ASV; the most abundant ASV in each group accounted for more than 96.1% of the total reads in that group.

The 16 selected eukaryotic groups accounted for 88.8% of the eukaryotic reads in the PicoNano dataset. The most abundant groups were “other Alveolata” (Alveolata that were not classified as Dinoflagellata; 17.5%, 1201 ASVs), Metazoa (17.2%, 376 ASVs), and “other Stramenopiles” (Stramenopiles that were not classified as Diatomea or Raphidophyceae, 12.4%, 716 ASVs); these three groups accounted for 47.1% of all reads (Fig. 1B). Each of the remaining 13 groups contributed from 0.14% to 7.7% of total reads. The HAB-forming groups Kareniaceae, *Chattonella*, and *Heterosigma* accounted for 7.7%, 1.5%, and 0.75% of total reads, respectively. The most abundant ASVs in these three groups contributed more than 88% of their respective groups. Diatomea, Dinophyceae, and Haptophyta contributed 7.3%, 6.4%, and 5.9% of total reads, respectively. All other groups contributed less than 5%. Chlorophytes (including Trebouxiophyceae, Prasinophytae, Mamiellophyceae, and Tetraselmis) accounted for 7.2% of all reads.

Prokaryotic reads were dominated by three phyla (i.e., Proteobacteria, Bacteroidetes, and Cyanobacteria), which represented 89.3% of all prokaryotic reads (Fig. 1D). The most abundant phylum was Proteobacteria (2,164 ASVs and 53.6% of total reads), 72.1% of which belonged to Alphaproteobacteria. Bacteroidetes (992 ASVs and 20.4% of total reads) was dominated by Bacteroidia (967 ASVs, 91.4% of Bacteroidetes reads). Cyanobacteria was dominated by Oxyphotobacteria (99.9% of Cyanobacteria reads). Three other phyla (Verrucomicrobiae, Actinobacteria, and Planctomycetes) contributed 4.8%, 3.8%, and 0.5% of total prokaryotic reads, respectively. Overall, these six phyla contributed 98.4% of prokaryotic reads.

We classified the 5,701 *Mimiviridae* ASVs (of the 6,261 total *Mimiviridae* ASVs) into 13 clades based on their phylogeny (Fig. S5). The majority of amplicon reads (61.9%) were assigned to clades 6 (518 ASVs, 24.0%), 3 (665 ASVs, 20.7%), and 5 (535 ASVs, 17.2%). The

other ten clades contributed from 0.3% to 8.4% of total *Mimiviridae* reads. Six out of the 13 *Mimiviridae* clades included previously studied *Mimiviridae* (clades 1, 3, 8, 9, 10, 13) (Fig. S5). Of these six clades, clade 3, which includes Organic lake phycodnavirus 1 and 2 (35), had the largest number of reads (20.7%). Clade 1 (7.8% of *Mimiviridae* reads) included haptophytes (Prymnesiophyceae) infecting *Mimiviridae* (36–38). Clade 10 (2.4% of *Mimiviridae* reads) contained the predatory Choanoflagellate infecting virus ChoanoV1 (39). Each of the other clades (clades 8, 9, 13) with known *Mimiviridae* contributed less than 2% of *Mimiviridae* reads.

Seasonal changes in microbial communities. The non-metric multidimensional scaling (NMDS) ordination of the datasets revealed similar seasonal patterns for the four communities (Fig. 2). Datasets from the same month clustered together in the plot, except for June datasets, thereby forming a yearly cycling of community succession. Consistent with this observation, the community dynamics were significantly correlated among the four microbial communities (correlation coefficients >0.75, $p < 10^{-4}$; Fig. S6). The average Sørensen-Dice dissimilarity between datasets from the same season ranged from 0.58 to 0.78 (Fig. 3), while that between datasets from opposite seasons (i.e., winter *vs*. summer or spring *vs*. fall) ranged from 0.77 to 0.94), indicative of greater dissimilarity of communities between opposite seasons (Fig. 3). The spring and fall communities differed from each other as much as the summer and winter communities (Fig. 3). The dissimilarities within the same season were significantly lower than those between spring and fall ($p < 10^{-16}$) and between summer and winter ($p < 10^{-16}$). However, dissimilarities between summer and winter communities did not significantly differ from those between spring and fall communities ($p > 0.1$), except that spring *vs.* fall dissimilarities were higher than summer *vs*. winter dissimilarities for the prokaryotic community ($p < 10^{-6}$).

Distinct patterns in seasonal dynamics among eukaryote, prokaryote, and *Mimiviridae* communities. The annual cycles of the communities were characterized by plotting the community dissimilarity between samples against the length of time separating the sampling events (Fig. 4A–D). In addition, we calculated the dissimilarities between communities sampled at intervals of approximately 1-month (i.e., 30 ± 10 days; hereafter D30), approximately 6 months (i.e., 182 ± 30 days; D182), and approximately 1 year (i.e., 365 ± 30 days; D365).

The level of these dissimilarity measures (i.e., D30, D182, and D365) was consistent with the yearly cycle pattern observed for the four communities (Fig. 4E–G). D182 was the largest among these measures for all communities (NanoPlus, PicoNano, prokaryotes, *Mimiviridae*), and the levels of D30 and D365 were comparable. However, these measures substantially differed when compared between communities ($p < 10^{-6}$), except for the comparisons between *Mimiviridae* and PicoNano communities for D30 and D365. These measures were the highest for *Mimiviridae* and PicoNano communities, followed by the NanoPlus community. The prokaryotic community showed the lowest dissimilarity values for all measures (i.e., D30, D182, and D365).

Recurrence and persistence of ASVs. To further characterize the dynamics of the eukaryote, prokaryote, and *Mimiviridae* communities, we investigated temporal patterns in the occurrence of individual ASVs. First, we analyzed the temporal pattern of the four most abundant ASVs in each community (Fig. 5). This analysis revealed that the dynamics of the selected prokaryotic ASVs clearly differed from those of other ASVs. The prokaryotic ASVs were characterized by a relatively low maximum relative abundance (<31%), whilst the dynamics of the most abundant ASVs from other communities (i.e., NanoPlus, PicoNano, and *Mimiviridae*) were characterized by sharp peaks in relative abundance (up to 70%) (Fig. 5). The most abundant prokaryotic ASV was assigned to SAR11 (Clade 1a) and showed 10%–30%

relative abundance more than 7 months from late autumn to early spring (present in 32 samples in total).

This result prompted us to examine recurrence (i.e., number of samples in which an ASV was detected) and persistence (i.e., number of successive samples in which an ASV was detected) of ASVs. In this analysis, we considered all the ASVs with >1% relative abundance in at least one sample. Among the four microbial groups, prokaryotes showed the highest level of ASV recurrence, and *Mimiviridae* showed the lowest (Fig. 6). Similarly, prokaryotes showed longer ASV persistence compared with the other microbe groups (Fig. S7). For instance, three prokaryotic ASVs (of 73 prokaryotic ASVs) were continuously present in more than 19 datasets, while there were no such long-lasting ASVs in the *Mimiviridae* or eukaryote groups.

Discussion

Eukaryotic communities in the study area were dominated by Metazoa, Alveolata, and Stramenopiles. These lineages are commonly found in coastal eukaryotic communities (4, 5, 16), although detailed compositions of microbial eukaryote communities in coastal areas vary among different locations (4, 5, 16). A notable feature of the study area is the occurrence of HABs (Fig. 1C), which were detected in our sequence data (Fig. 1A, B). The prokaryotic communities in the Uranouchi Inlet (Fig. 1D) were similar to those reported for other coastal areas, where the communities are mainly composed of Proteobacteria, Bacteroidetes, and Cyanobacteria (10, 15, 16), but different from those in brackish water (40) or freshwater (12), which are dominated by the same phyla, but different classes. In the present study, the *Mimiviridae* communities were characterized using a previously developed highly sensitive metabarcoding method (33), which was used for the first time on a time-series sample set. This revealed diverse *Mimiviridae* (6,261 ASVs, 464.6 ASVs per sample on average) that were grouped into 13 clades (Fig. 1E).

All four microbial communities showed seasonal cycles (Fig. 2). Such seasonal changes in microbial communities have been reported previously for eukaryotes (4, 6), bacteria (11, 13), large DNA viruses (18, 20), and smaller viruses (14, 19, 24). The greatest dissimilarity in communities in our datasets was between communities sampled at 6-month intervals (Fig. 4F), while communities were more similar when compared at approximately 12-month intervals (Fig. 4G), indicating that community structure recovered after a year. The high dissimilarity between summer and winter datasets was expected, since physicochemical parameters such as temperature are very different between these two seasons. However, for all microbe groups, the microbial communities in spring *vs.* fall were as dissimilar as the communities in summer *vs.* winter (Fig. 3), even though the differences in water conditions (e.g., temperature, salinity, and nutrient concentrations) were less pronounced between fall and spring than between summer and winter (Fig. S1, S2). We also observed that the community compositions were relatively similar between samples collected in successive months (Fig. 2 and Fig. 4E). These observations suggest that the dynamics of microbial communities cannot be explained entirely by abiotic factors, but are governed by a Markovian process in which the current biotic state substantially influences the next biotic state. This pattern is obvious in macro-ecosystems; for example, a forest looks different in spring and autumn even on days with similar weather. Microbial communities with a much higher turnover rate appear to be no exception to this rule. A previous study on oceanic microbial communities suggested that biotic factors (such as the subset of communities) are better predictors of microbial community composition than are abiotic factors (such as geography, nutrient concentration, and temperature) (41). This reinforces the classical notion that the sum of biotic and abiotic factors forms the environment that affects individual microorganisms. We suggest that the microbial communities in the Uranouchi Inlet at a given

time strongly influence the formation of the next generation of communities. A similar observation and idea has been proposed for post-bloom communities (17).

All the studied microbial communities became similar to the original communities after a year, but the recovery level of community structure was far below 100% (i.e., 19% to 44% recovery, Fig. 4). In other words, more than half of the ASVs in the original communities were undetected by deep sequencing 1 year later. Previous long-term studies on prokaryotic (1) and viral communities (14, 24) detected a similar trend that continued for several years. Therefore, it seems legitimate to postulate that the marine microbial community structure at the ASV level never perfectly recovers over time (Fig. 7A). This implies a significant role of the rare biosphere in the formation of a new community structure through their capacity to bring forth new abundant microbes over time (6, 14). The rare biosphere may be invisible, even with state-of-the-art metabarcoding approaches as used in this study. The technical limitations in detecting the rare biosphere combined with its capacity to substantially replace abundant populations may explain why microbes cannot completely recover their community structure, although migration and evolution could also have significant effects. Interestingly, Furman et al. found that the dissimilarity of bacterial community structure increased during the first four years of analysis (1). After this period, the dissimilarity level still fluctuated (according to seasonal dynamics) but its basal level remained constant (1). This suggests the existence of a “memory” (or a historical trace) of the original community structure that influences the structures of successive communities during a short period of time (one to a few years) (Fig. 7A).

Although a pattern of annual cycling and partial recovery after a year was commonly observed for the four microbial communities in this study, the degree and rate of changes in the community structure differed among the eukaryotic, prokaryotic, and *Mimiviridae* communities

(Fig. 4). A schematic diagram in Fig. 7B illustrates the conceptual and qualitative differences in the observed community dynamics among eukaryotes, prokaryotes, and *Mimiviridae* (Fig. 7B). The D182 and D365 values were highest for *Mimiviridae* and lowest for prokaryotes. Thus, in terms of "memory" of community structure, *Mimiviridae* have the shortest memory, eukaryotes have medium memory, and prokaryotes have the longest memory. The *Mimiviridae* community was characterized by rapid change and a low level of recovery after a year. In contrast, a previous study on the seasonal changes of small viruses (i.e., mostly bacterial viruses) detected strong stability of the small virus community, with 95% of viral contigs present in all samples (14). The prokaryotic community changed slowly and showed the highest level of recovery after a year. A previous study conducted on a shorter time scale also found prokaryotic communities to be more stable and to change more slowly than eukaryotic communities (16).

Our analyses of the recurrence and persistence of ASVs among eukaryote, prokaryote, and *Mimiviridae* communities provide further insight into the differences in community dynamics among microbial groups. The recurrence and persistence levels of ASVs (Fig. 6, Fig. S7) and the relative abundance of major ASVs (Fig. 5) differed among the microbial groups. Overall, prokaryotic ASVs showed higher recurrence and longer persistence than ASVs in the other microbial groups. In contrast, *Mimiviridae* ASVs showed the lowest recurrence and persistence levels. The abundant prokaryotic ASVs were dominant throughout much of the year, whilst their maximum relative abundance was lower (<30%) than that of dominant ASVs in other microbial groups (Fig. 5). In contrast, abundant *Mimiviridae* ASVs were found only in a few samples but they had higher relative abundance (21.1%–71.3%). The pattern of eukaryotic ASVs occurrence was between these extremes. The four most abundant ASVs in the PicoNano

and the NanoPlus groups had high relative abundance (23.5%–90.3%) but were present in more samples than were the most abundant *Mimiviridae* ASVs (Fig. 5).

The differences in the community/ASV dynamics among eukaryotes, *Mimiviridae*, and prokaryotes may result from the limitations of metabarcoding. It is worth noting that ASVs may have different resolutions across microbial groups. For instance, the result that bacterial ASVs showed longer persistence than eukaryote and *Mimiviridae* ASVs may because their resolution was lower. However, if we assume that this is the case, the comparatively high abundance peaks for the major eukaryotic and *Mimiviridae* ASVs become difficult to explain, because their resolution was assumed to be higher than that of bacterial ASVs. Thus, the differences in dynamics may reflect some essential features of populations belonging to the respective microbial communities. In the following paragraph, we attempt to explain these differences in light of the possible strategies of individual populations within those communities.

The population dynamics of microbes largely depend on resource availability and their survival strategies, which have historically been classified into "*r*" and "*K*" strategies. The original interpretation of *r/K* selection theory states that organisms maximize their fitness through either high reproduction rates (high *r*), or the ability to constantly exist in relatively high numbers (high carrying capacity, *K*) (42). However, in microbial ecology, the *r*/*K* nomenclature has often been used qualitatively (42). That is, "*r*-strategists" are specialized to rapidly consume resources and reproduce quickly (42) so that they thrive through an influx of nutrients (15). In contrast, *K*-strategists adapt to a stable environment with limited resources (42). They tend to show slow growth rates and their population dynamics are more stable than those of *r*-strategists (43).

In our study, several of the major eukaryotic ASVs were present over a long period of time (i.e., the *Skeletonema* sp. ASV, the *Oithona davisae* ASV and the Dinoflagellata ASV; Fig. 5), reminiscent of *K*-strategists. The Dinoflagellata ASV in the NanoPlus fraction was abundant throughout several months in spring during oligotrophic conditions. The *Oithona davisae* ASV (a copepod preying on plankton) was present throughout the year with no pronounced peak in its relative abundance. In contrast, other abundant ASVs (e.g., the Symbiodinium ASV) and those assigned to HAB species (e.g., the *Chattonella* spp. ASVs, the *K. mikimotoi* ASVs, and the *H. akashiwo* ASVs; Fig. 1A and B) emerged only occasionally but with a high relative abundance (>20%). The dynamics of the HAB ASVs were consistent with their cell counts (Fig. 1C). They formed blooms only during the months with high nutrient availability (Fig. S1, Fig. S2). This ecological characteristic is typical of *r*-strategists (42, 43).

The majority of *Mimiviridae* ASVs showed low persistence and occasional high peaks in relative abundance. The most abundant *Mimiviridae* ASVs occurred only in a few samples but had high relative abundance of over 70% (Fig. 5). These dynamics are reminiscent of *r*-strategists such as HAB species that show a rapid boom and bust appearance, and are consistent with the view that the presence of viruses depends on the presence of their hosts, not vice-versa. The lower recurrence and persistence of *Mimiviridae* ASVs than eukaryote ASVs may be due to the existence of multiple *Mimiviridae* that can infect the same host species. A recent study suggested that *Mimiviridae* that infect *K*-strategist eukaryotes have long infection and latent periods to maximize the chance of vertical transmission, since the chances of horizontal transmissions are low due to the low host density (44). Viruses of *K*-strategist eukaryotes are more likely to exist inside the host cell (i.e., a long latent period), while viruses of *r*-strategist eukaryotes are more likely to be found outside their host (i.e., they show higher virulence and

tend to produce virions more quickly) (44). The size fraction we used for the detection of viral sequences (0.2–0.8 µm) may be enriched in free virions, which might have led to the preferential detection of *r*-strategist *Mimiviridae*. A few *Mimiviridae* ASVs showed a relatively long persistence (occurrence in more than ten continuous samples) (Fig. S6) or a high recurrence level (Fig. 6). Investigation of active *Mimiviridae* (e.g., by RNA-Seq from the PicoNano and NanoPlus fractions) may reveal *Mimiviridae* ASVs showing *K*-strategist dynamics.

The most abundant prokaryotic ASVs had comparatively low relative abundances (Fig. 5), but were found in many samples (Fig. 6). The most abundant prokaryotic ASV belonged to SAR11, which is considered to be a *K*-strategist due to its ubiquitous presence in the ocean and low maximum growth rate (45, 46). As expected, this ASV was abundant during months with low nutrient availability with no pronounced sharp peaks in abundance. Several bacterial genera are *r*-strategists (e.g., *Ulvibacter*, *Formosa*, and *Polaribacter*) (8, 17, 47), but they were not highly represented in our dataset. In this study, prokaryotic communities were more stable than were eukaryotic and *Mimiviridae* communities. Overall, the distinct community dynamics among eukaryotes, *Mimiviridae*, and prokaryotes revealed in our study may be due to the differences in the prevalence of *r vs. K* strategies among these microbial groups.

In summary, our results revealed a seasonal cycle for eukaryote, *Mimiviridae*, and prokaryote communities in the Uranouchi Inlet. The seasonality cannot be explained solely by abiotic factors and is suggested to be shaped through a Markovian process of successive communities (i.e., “biotic effect”) under the influence of abiotic factors. We also detected differences in the period of “community memory” among eukaryotes, *Mimiviridae*, and prokaryotes, which were correlated with the recurrence and persistence of ASVs in the respective

communities. These differences might be due to distinct ecological traits among the major eukaryotes, *Mimiviridae*, and prokaryotes in the study area.

Materials and Methods

Sampling and DNA extraction. Seawater was collected at four stations in Uranouchi Inlet, Kochi Prefecture, Japan (i.e., station “j”: 33°25'43.2"N 133°22'49.5"E, station “m”: 33°25'60.0"N 133°24'38.3"E, station “f”: 33°26'33.6"N 133°24'41.8"E, station “u”: 33°25'49.7"N 133°24'01.4"E). Time-series sampling was performed sporadically (average interval, 15 days) with a higher sampling frequency during summer. In total, 43 sample sets (i.e., 43 seawater samples, with three size fractions each) were obtained over 20 months from 5th January 2017 (170105-j) to 25th September 2018 (180925-j). The samples were denoted with the sampling date followed by the station’s abbreviation in “yymmdd-station” format.

Seawater (10 L) was collected from 5-m depth, transported to the laboratory, and then filtered through 3.0-µm (diameter 142 mm) (polycarbonate, Merck, Darmstadt, Germany) and 0.8-µm (diameter 142 mm) membranes (polycarbonate, Merck). The filtrate (1 L) was further filtered through a 0.22-µm filtration unit (Sterivex, polycarbonate, Merck). The filters were kept at −80°C until DNA extraction with the Proteinase-K method for the 0.22-µm filtration units (48) and the xanthogenate-sodium dodecyl sulfate method for larger size fractions (0.8-µm and 3-µm filters) (49).

PCR amplification. The 18S ribosomal RNA gene primers (forward “V8f”: ATAACAGGTCTGTGATGCCCT, reverse “1510r”: CCTTCYGCAGGTTCACCTAC) (50) targeting the V8/V9 region were used on both the 0.8–3 µm and >3 µm size fractions. Hereafter, the 0.8–3 µm and >3 µm fractions are referred to as the PicoNano and NanoPlus fractions, respectively. The PCR analysis was performed by mixing 5 µL of 1 µmol·L^{-1} of both primers and 2.5 µL of 0.25 ng·µL^{-1} template DNA with 12.5 µL 2× KAPA HiFi HotStart ReadyMix (Roche, Basel, Switzerland). The thermal program was 98°C for 3 min, followed by 25 cycles of

98°C for 20 s, 65°C for 15 s, 72°C for 15 s, and 72°C for 10 min. The 16S rRNA gene amplification was performed by mixing 5 µL of 1 µmol·L^{-1} of both primers (forward "Pro-16S NGS": CCTACGGGNBGCASCAG, reverse "Pro-16S NGS": GACTACNVGGGTATCTAATCC) (51) with 2 µL of 1 ng·μL^{-1} sample DNA (0.2–0.8 µm size fraction) and 12 µL 2× KAPA HiFi HotStart ReadyMix. The thermal program was 95°C for 3 min, then 25 cycles of 95°C for 30 s, 55°C for 30 s, 72°C for 30 s, and final extension at 72°C for 5 min. For the *Mimiviridae* DNA polymerase gene amplification, we used a recently published method that has been described in detail elsewhere (i.e., "MP10v2" in combination with "protocol 3") (33).

Clean up and sequencing. The 18S rRNA gene and 16S rRNA gene amplicons were purified with Agencourt AMPure XP beads (Beckman Coulter, Inc., Brea, CA, USA) following Illumina's library preparation protocol. The *Mimiviridae polB* gene amplicons were purified by gel extraction (33). After amplicon purification, dual indices were attached by PCR. The second purification was performed according to Illumina's library preparation protocol for the cellular amplicons and with gel extraction for *polB* amplicons. The final library concentration was 10 pM with a 30% PhiX spike. The sequencing data were generated with paired-end sequencing (2 × 300 nucleotides) on the MiSeq platform (Illumina, Inc., San Diego, CA, USA).

Analysis of prokaryotic and eukaryotic amplicon raw reads. The data were analyzed separately using QIIME 2 (52) with default settings unless noted otherwise. The pipelines for both prokaryotes and eukaryotes were nearly identical. The raw reads were imported to QIIME 2 (version 2020.2.0) and primers were removed with Cutadapt (53). Forward and reverse reads were merged using the vsearch plugin of QIIME 2 with a minimum overlap of 35 bases and up to five accepted mismatches. Low-quality merged reads were removed (minimum PHRED score of

10) and the vsearch plugin was used to perform dereplication and chimera checking against the SILVA 132 database (54). Reads were truncated at 250 bp. We used dada2 for ASV generation in QIIME 2 (55). The ASVs were taxonomically assigned at 90% nucleotide identity using "classify-consensus-vsearch" in QIIME 2 against the SILVA 132 database (taxonomic majority, 97%) for either 16S amplicon reads or 18S amplicon reads.

Analysis of *Mimiviridae polB* amplicon raw reads. *Mimiviridae polB* ASVs were generated by a newly developed pipeline named MAPS2, which is mostly based on R (3.6.3) (www.R-project.org) and the "dada2" package. The functions of dada2 were used with default settings unless specified otherwise. Raw reads were loaded using the "filterAndTrim" command, truncated at a length of 240 bp (both forward and reverse reads), and two expected read errors were accepted. Primer sequence regions were trimmed (i.e., forward reads: 40 bp and reverse reads: 38 bp). The chance of sequencing error was calculated with dada2's "learnErrors" function and sequences were dereplicated with the "derepFastq" function. The "dada" function was used on the reads, and the reads were merged with the "mergePairs" function. Chimeric sequences were removed using the "nochim" function. The nucleotide sequences and the ASV table were then exported. The nucleotide sequences were searched against a database of 11,413 *polB* sequences, 1,004 of which were viral *Mimiviridae* sequences, using BLASTx (2.9.0)(56) (like in the original MAPS (57)). Sequences with the best hit to viral sequences were processed further (E-value $<10^{-5}$). Both the original nucleotide sequence and the amino acid sequences (translated by BLASTx) were saved. MAFFT (7.453) (58) was used to align the reads with those in a reference file containing 1109 PolB sequences from different organisms and viruses. The aligned PolB sequences were added to a reference tree (32) using pplacer (1.1.alpha19) (59). We removed ASVs that were not within the *Mimiviridae* branch of the reference tree with R (Fig.

S5). The ASVs within the *Mimiviridae* branch were split into 13 clades for community structure analysis (Fig. S5). Finally, a Python (3.7.5) (www.python.org) script was used to cut the *Mimiviridae* ASV sequences in a common region. The ASVs that became identical as a result of this process were clustered using cd-hit (clustering at 100% nucleotide identity).

Community composition and diversity analysis. Communities were compared with R with the "vegan" (2.5-6) (60) and "ggplot2" (3.2.1) (https://ggplot2.tidyverse.org) packages. Singleton ASVs and ASVs that were not taxonomically assigned and datasets with fewer than 8,000 taxonomically assigned reads were removed before analyses of community composition (Fig. 1). Ecological analysis was performed after subsampling with "rarefy" at 8,000 reads. An ASV was determined to be present in the dataset if it yielded more than 0.1% of reads after subsampling. Community dissimilarity was measured using the Sørensen-Dice dissimilarity test with vegan's "vegdist" function (method = "bray", binary = T), as follows:

$$dissimilarity_{SD} = \frac{2|community_A \cap community_B|}{|community_A| + |community_B|}$$

Mantel test was performed using R's "mantel" command (5 million permutations). The NMDS analysis was conducted with R's "cmdscale" command, and stress was calculated using the "sammon" command (MASS package). The similarity of communities among seasons was calculated by binning datasets of several months: winter (December to February), spring (March to May), summer (June to August), and fall (September to November). The *p* values for seasonal differences were calculated with R's "t.test" command, if more than 100 values were compared we used the "pairwise.t.test" command with Benjamini–Hochberg correction.

Cell counts, biotic, and abiotic data. Environmental metadata (temperature, salinity, phytoplankton cell counts, and concentrations of nutrients, organic and inorganic forms of nitrogen and phosphorus, dissolved oxygen, and chlorophyll *a*) for the Uranouchi Inlet from

2013 to 2019 were obtained from the observation records of the Kochi Prefectural Fisheries Research Institute with appropriate ethics approval. Since these records aim to monitor and forecast HABs in the sampling area, cell count data were only collected for HAB species. We used the daily maximum cell count data in this study. Different species are considered to be blooming at different cellular concentrations, because bloom identification is linked to heightened fish mortality according to the Prefectural Fisheries Research Institute (https://www.pref.kochi.lg.jp/soshiki/040409). Fish mortality rates increase when *Chattonella* spp. reaches cellular concentrations as low as 10 to 100 cells mL^{-1}; when *K. mikimotoi* (an unarmored, solitary dinoflagellate, 24–40 µm long, 20–32 µm wide (61)) reaches concentrations ranging from several 100 to 1,000 cells·mL^{-1}; and when *H. akashiwo* (kidney shaped, motile dinoflagellate, 8–25 µm long (61)) reaches concentrations starting from 10.000 cells·mL^{-1}.

Data availability. The raw sequencing data of this study have been deposited in the DNA database of Japan (accession number DRA010976). The environmental data and cell count data are available in the Microsoft Excel file (dataset S1). The amplicon data analysis pipelines for eukaryotes, prokaryotes, and *Mimiviridae* have been deposited online and are available for download (www.github.com/FlorianProdinger/pipeline_18S, www.github.com/FlorianProdinger/pipeline_16S, www.github.com/FlorianProdinger/MAPS2).

Acknowledgments

F.P. performed experiments, bioinformatics analyses, and wrote the manuscript. H.E., T.Y. and H.O. designed the work and helped to interpret the results. Y.G. and H.T. contributed to sequencing. R.B.M and Y.L. contributed to bioinformatics analysis. K.T. and I.T. contributed to experimental analyses. Y.T., K.N., and E.T. contributed to sampling. All authors contributed to writing and agreed to the final manuscript. Funding for this research was provided by The Canon

Foundation (No. 203143100025), JSPS/KAKENHI (Nos. 26430184, 17H03850, 18H02279, 16H06279 (PAGS)), Scientific Research on Innovative Areas from the Ministry of Education, Culture, Science, Sports and Technology (MEXT) of Japan (Nos. 16H06429, 16K21723, 16H06437), The Kyoto University Foundation, and the Collaborative Research Program of the Institute for Chemical Research, Kyoto University (Nos. 2019-33, 2018-31, 2017-25, 2016-28). Computational work was completed at the SuperComputer System, Institute for Chemical Research, Kyoto University. We thank the Kochi Prefectural Fisheries Research Institute for providing environmental metadata. We thank Jennifer Smith, PhD, from Edanz Group (https://en-author-services.edanz.com/), for editing a draft of this manuscript.

The authors declare no conflict of interest.

References

1. Fuhrman JA, Cram JA, Needham DM. 2015. Marine microbial community dynamics and their ecological interpretation. 3. Nat Rev Microbiol 13:133–146.
2. Not F, Siano R, Kooistra WHCF, Simon N, Vaulot D, Probert I. 2012. Chapter One - Diversity and Ecology of Eukaryotic Marine Phytoplankton, p. 1–53. *In* Piganeau, G (ed.), Advances in Botanical Research. Academic Press.
3. Brum JR, Sullivan MB. 2015. Rising to the challenge: accelerated pace of discovery transforms marine virology. 3. Nat Rev Microbiol 13:147–159.
4. Gran-Stadniczeñko S, Egge E, Hostyeva V, Logares R, Eikrem W, Edvardsen B. 2019. Protist Diversity and Seasonal Dynamics in Skagerrak Plankton Communities as Revealed by Metabarcoding and Microscopy. J Eukaryot Microbiol 66:494–513.
5. Chen W, Pan Y, Yu L, Yang J, Zhang W. 2017. Patterns and Processes in Marine Microeukaryotic Community Biogeography from Xiamen Coastal Waters and Intertidal Sediments, Southeast China. Front Microbiol 8.
6. Giner CR, Balagué V, Krabberød AK, Ferrera I, Reñé A, Garcés E, Gasol JM, Logares R, Massana R. 2019. Quantifying long-term recurrence in planktonic microbial eukaryotes. Mol Ecol 28:923–935.
7. Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B, Huse S, McHardy AC, Knight R, Joint I, Somerfield P, Fuhrman JA, Field D. 2012. Defining seasonal marine microbial community dynamics. ISME J 6:298–308.
8. Teeling H, Fuchs BM, Bennke CM, Krüger K, Chafee M, Kappelmann L, Reintjes G, Waldmann J, Quast C, Glöckner FO, Lucas J, Wichels A, Gerdts G, Wiltshire KH, Amann

RI. 2016. Recurring patterns in bacterioplankton dynamics during coastal spring algae blooms. eLife 5.

9. Milici M, Deng Z-L, Tomasch J, Decelle J, Wos-Oxley ML, Wang H, Jáuregui R, Plumeier I, Giebel H-A, Badewien TH, Wurst M, Pieper DH, Simon M, Wagner-Döbler I. 2016. Co-occurrence Analysis of Microbial Taxa in the Atlantic Ocean Reveals High Connectivity in the Free-Living Bacterioplankton. Front Microbiol 7.

10. Sakami T, Watanabe T, Kakehi S, Taniuchi Y, Kuwata A. 2016. Spatial variation of bacterial community composition at the expiry of spring phytoplankton bloom in Sendai Bay, Japan. Gene 576:610–617.

11. Ward CS, Yung C-M, Davis KM, Blinebry SK, Williams TC, Johnson ZI, Hunt DE. 2017. Annual community patterns are driven by seasonal switching between closely related marine bacteria. ISME J 11:1412–1422.

12. Bock C, Salcher M, Jensen M, Pandey RV, Boenigk J. 2018. Synchrony of Eukaryotic and Prokaryotic Planktonic Communities in Three Seasonally Sampled Austrian Lakes. Front Microbiol 9.

13. Chafee M, Fernàndez-Guerra A, Buttigieg PL, Gerdts G, Eren AM, Teeling H, Amann RI. 2018. Recurrent patterns of microdiversity in a temperate coastal marine environment. 1. ISME J 12:237–252.

14. Ignacio-Espinoza JC, Ahlgren NA, Fuhrman JA. 2020. Long-term stability and Red Queen-like strain dynamics in marine viruses. 2. Nat Microbiol 5:265–271.

15. Santi I, Tsiola A, Dimitriou PD, Fodelianakis S, Kasapidis P, Papageorgiou N, Daffonchio D, Pitta P, Karakassis I. 2019. Prokaryotic and eukaryotic microbial community responses

to N and P nutrient addition in oligotrophic Mediterranean coastal waters: Novel insights from DNA metabarcoding and network analysis. Mar Environ Res 150:104752.

16. Martin-Platero AM, Cleary B, Kauffman K, Preheim SP, McGillicuddy DJ, Alm EJ, Polz MF. 2018. High resolution time series reveals cohesive but short-lived communities in coastal plankton. Nat Commun 9:266.

17. Needham DM, Fuhrman JA. 2016. Pronounced daily succession of phytoplankton, archaea and bacteria following a spring bloom. 4. Nat Microbiol 1:1–7.

18. Gran-Stadniczeñko S, Krabberød AK, Sandaa R-A, Yau S, Egge E, Edvardsen B. 2019. Seasonal Dynamics of Algae-Infecting Viruses and Their Inferred Interactions with Protists. Viruses 11.

19. Pagarete A, Chow C-ET, Johannessen T, Fuhrman JA, Thingstad TF, Sandaa RA. 2013. Strong Seasonality and Interannual Recurrence in Marine Myovirus Communities. Appl Environ Microbiol 79:6253–6259.

20. Sandaa R-A, E. Storesund J, Olesin E, Lund Paulsen M, Larsen A, Bratbak G, Ray JL. 2018. Seasonality Drives Microbial Community Structure, Shaping both Eukaryotic and Prokaryotic Host–Viral Relationships in an Arctic Marine Ecosystem. Viruses 10.

21. Johannessen TV, Larsen A, Bratbak G, Pagarete A, Edvardsen B, Egge ED, Sandaa R-A. 2017. Seasonal Dynamics of Haptophytes and dsDNA Algal Viruses Suggest Complex Virus-Host Relationship. Viruses 9.

22. Arkhipova K, Skvortsov T, Quinn JP, McGrath JW, Allen CC, Dutilh BE, McElarney Y, Kulakov LA. 2018. Temporal dynamics of uncultured viruses: a new dimension in viral diversity. ISME J 12:199.

23. Needham DM, Chow C-ET, Cram JA, Sachdeva R, Parada A, Fuhrman JA. 2013. Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. 7. ISME J 7:1274–1285.

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

25. Tomaru Y, Tarutani K, Yamaguchi M, Nagasaki K. 2004. Quantitative and qualitative impacts of viral infection on a Heterosigma akashiwo (Raphidophyceae) bloom in Hiroshima Bay, Japan. Aquat Microb Ecol 34:227–238.

26. Moniruzzaman M, LeCleir GR, Brown CM, Gobler CJ, Bidle KD, Wilson WH, Wilhelm SW. 2014. Genome of brown tide virus (AaV), the little giant of the Megaviridae, elucidates NCLDV genome expansion and host–virus coevolution. Virology 466–467:60–70.

27. Lehahn Y, Koren I, Schatz D, Frada M, Sheyn U, Boss E, Efrati S, Rudich Y, Trainic M, Sharoni S, Laber C, DiTullio GR, Coolen MJL, Martins AM, Van Mooy BAS, Bidle KD, Vardi A. 2014. Decoupling physical from biological processes to assess the impact of viruses on a mesoscale algal bloom. Curr Biol CB 24:2041–2046.

28. Tarutani K, Nagasaki K, Yamaguchi M. 2000. Viral Impacts on Total Abundance and Clonal Composition of the Harmful Bloom-Forming Phytoplankton Heterosigma akashiwo. Appl Environ Microbiol 66:4916–4920.

29. Kuhlisch C, Schleyer G, Shahaf N, Vincent F, Schatz D, Vardi A. 2020. Viral infection of algal blooms leaves a halogenated footprint on the dissolved organic matter in the ocean. bioRxiv 2020.09.08.287805.

30. Nagasaki K, Tomaru Y, Tarutani K, Katanozaka N, Yamanaka S, Tanabe H, Yamaguchi M. 2003. Growth Characteristics and Intraspecies Host Specificity of a Large Virus Infecting the Dinoflagellate Heterocapsa circularisquama. Appl Environ Microbiol 69:2580–2586.

31. Hingamp P, Grimsley N, Acinas SG, Clerissi C, Subirana L, Poulain J, Ferrera I, Sarmento H, Villar E, Lima-Mendez G, Faust K, Sunagawa S, Claverie J-M, Moreau H, Desdevises Y, Bork P, Raes J, de Vargas C, Karsenti E, Kandels-Lewis S, Jaillon O, Not F, Pesant S, Wincker P, Ogata H. 2013. Exploring nucleo-cytoplasmic large DNA viruses in Tara Oceans microbial metagenomes. ISME J 7:1678–1695.

32. Endo H, Blanc-Mathieu R, Li Y, Salazar G, Henry N, Labadie K, de Vargas C, Sullivan MB, Bowler C, Wincker P, Karp-Boss L, Sunagawa S, Ogata H. 2020. Biogeography of marine giant viruses reveals their interplay with eukaryotes and ecological functions. Nat Ecol Evol 1–11.

33. Prodinger F, Endo H, Gotoh Y, Li Y, Morimoto D, Omae K, Tominaga K, Blanc-Mathieu R, Takano Y, Hayashi T, Nagasaki K, Yoshida T, Ogata H. 2020. An Optimized Metabarcoding Method for Mimiviridae. 4. Microorganisms 8:506.

34. Gómez-Pereira PR, Schüler M, Fuchs BM, Bennke C, Teeling H, Waldmann J, Richter M, Barbe V, Bataille E, Glöckner FO, Amann R. 2012. Genomic content of uncultured Bacteroidetes from contrasting oceanic provinces in the North Atlantic Ocean. Environ Microbiol 14:52–66.

35. Zhang W, Zhou J, Liu T, Yu Y, Pan Y, Yan S, Wang Y. 2015. Four novel algal virus genomes discovered from Yellowstone Lake metagenomes. Sci Rep 5:15131.

36. Gallot-Lavallée L, Pagarete A, Legendre M, Santini S, Sandaa R-A, Himmelbauer H, Ogata H, Bratbak G, Claverie J-M. 2015. The 474-Kilobase-Pair Complete Genome Sequence of

CeV-01B, a Virus Infecting Haptolina (Chrysochromulina) ericina (Prymnesiophyceae). Genome Announc 3.

37. Johannessen TV, Bratbak G, Larsen A, Ogata H, Egge ES, Edvardsen B, Eikrem W, Sandaa R-A. 2015. Characterisation of three novel giant viruses reveals huge diversity among viruses infecting Prymnesiales (Haptophyta). Virology 476:180–188.

38. Santini S, Jeudy S, Bartoli J, Poirot O, Lescot M, Abergel C, Barbe V, Wommack KE, Noordeloos AAM, Brussaard CPD, Claverie J-M. 2013. Genome of Phaeocystis globosa virus PgV-16T highlights the common ancestry of the largest known DNA viruses infecting eukaryotes. Proc Natl Acad Sci U S A 110:10800–10805.

39. Needham DM, Yoshizawa S, Hosaka T, Poirier C, Choi CJ, Hehenberger E, Irwin NAT, Wilken S, Yung C-M, Bachy C, Kurihara R, Nakajima Y, Kojima K, Kimura-Someya T, Leonard G, Malmstrom RR, Mende DR, Olson DK, Sudo Y, Sudek S, Richards TA, DeLong EF, Keeling PJ, Santoro AE, Shirouzu M, Iwasaki W, Worden AZ. 2019. A distinct lineage of giant viruses brings a rhodopsin photosystem to unicellular marine predators. Proc Natl Acad Sci 201907517.

40. Fazi S, Baldassarre L, Cassin D, Quero GM, Pizzetti I, Cibic T, Luna GM, Zonta R, Del Negro P. 2020. Prokaryotic community composition and distribution in coastal sediments following a Po river flood event (northern Adriatic Sea, Italy). Estuar Coast Shelf Sci 233:106547.

41. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, Chaffron S, Ignacio-Espinosa JC, Roux S, Vincent F, Bittner L, Darzi Y, Wang J, Audic S, Berline L, Bontempi G, Cabello AM, Coppola L, Cornejo-Castillo FM, d'Ovidio F, Meester LD, Ferrera I, Garet-Delmas M-J, Guidi L, Lara E, Pesant S, Royo-Llonch M, Salazar G, Sánchez P,

Sebastian M, Souffreau C, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Coordinators TO, Gorsky G, Not F, Ogata H, Speich S, Stemmann L, Weissenbach J, Wincker P, Acinas SG, Sunagawa S, Bork P, Sullivan MB, Karsenti E, Bowler C, Vargas C de, Raes J. 2015. Determinants of community structure in the global plankton interactome. Science 348.

42. Andrews JH, Harris RF. 1986. r- and K-Selection and Microbial Ecology, p. 99–147. *In* Marshall, KC (ed.), Advances in Microbial Ecology. Springer US, Boston, MA.

43. Endo H, Ogata H, Suzuki K. 2018. Contrasting biogeography and diversity patterns between diatoms and haptophytes in the central Pacific Ocean. Sci Rep 8:1–13.

44. Blanc-Mathieu R, Dahle H, Hofgaard A, Brandt D, Kalinowski J, Ogata H, Sandaa R-A. 2020. The genome of a persistent giant algal virus encodes an unprecedented number of genes involved in energy metabolism. bioRxiv 2020.07.30.228163.

45. Suttle CA. 2007. Marine viruses — major players in the global ecosystem. 10. Nat Rev Microbiol 5:801–812.

46. Giovannoni SJ. 2017. SAR11 Bacteria: The Most Abundant Plankton in the Oceans. Annu Rev Mar Sci 9:231–255.

47. Teeling H, Fuchs BM, Becher D, Klockow C, Gardebrecht A, Bennke CM, Kassabgy M, Huang S, Mann AJ, Waldmann J, Weber M, Klindworth A, Otto A, Lange J, Bernhardt J, Reinsch C, Hecker M, Peplies J, Bockelmann FD, Callies U, Gerdts G, Wichels A, Wiltshire KH, Glöckner FO, Schweder T, Amann R. 2012. Substrate-Controlled Succession of Marine Bacterioplankton Populations Induced by a Phytoplankton Bloom. Science 336:608–611.

48. Frias-Lopez J, Shi Y, Tyson GW, Coleman ML, Schuster SC, Chisholm SW, Delong EF. 2008. Microbial community gene expression in ocean surface waters. Proc Natl Acad Sci U S A 105:3805–3810.

49. Yoshida T, Yuki Y, Lei S, Chinen H, Yoshida M, Kondo R, Hiroishi S. 2003. Quantitative Detection of Toxic Strains of the Cyanobacterial Genus Microcystis by Competitive PCR. Microbes Environ 18:16–23.

50. Bradley IM, Pinto AJ, Guest JS. 2016. Design and Evaluation of Illumina MiSeq-Compatible, 18S rRNA Gene-Specific Primers for Improved Characterization of Mixed Phototrophic Communities. Appl Environ Microbiol 82:5878–5891.

51. Takahashi S, Tomita J, Nishioka K, Hisada T, Nishijima M. 2014. Development of a Prokaryotic Universal Primer for Simultaneous Analysis of Bacteria and Archaea Using Next-Generation Sequencing. PLOS ONE 9:e105592.

52. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet C, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope E, Silva RD, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley G, Janssen S, Jarmusch AK, Jiang L, Kaehler B, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MG, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton J, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Michael S Robeson II, Rosenthal P, Segata N, Shaffer M, Shiffer

A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, Hooft JJ van der, Vargas F, Vázquez-Baeza Y, Vogtmann E, Hippel M von, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CH, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2018. QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science. e27295v2. PeerJ Inc.

53. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17:10–12.

54. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590–D596.

55. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581–583.

56. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

57. Li Y, Hingamp P, Watai H, Endo H, Yoshida T, Ogata H, Li Y, Hingamp P, Watai H, Endo H, Yoshida T, Ogata H. 2018. Degenerate PCR Primers to Reveal the Diversity of Giant Viruses in Coastal Waters. Viruses 10:496.

58. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780.

59. Matsen FA, Kodner RB, Armbrust EV. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics 11:538.

60. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Sólymos P, Stevens MHH, Wagner H. 2012. vegan: Community Ecology Package.

61. Omura T, Iwataki M, Borja VM, Takayama H, Fukuyo Y. Marine Phytoplankton of theWestern Pacific.

Tables and figures

Table 1: Summary of generated datasets.

	Raw reads	Reads	Richness all reads	Richness per samples subsampled	Average richness subsampled
NanoPlus	3,613,399	3,001,267	3,356	52 – 253	120 ± 51
PicoNano	2,692,119	2,180,138	4,559	115 – 410	278 ± 68
Prokaryotes	4,044,427	3,427,158	4,478	141 – 415	275 ± 57
Mimiviridae	45,106,746	5,712,580	6,261	108 – 394	246 ± 72

Notes: Table contains read and ASV numbers for all datasets with >8,000 reads. Singletons and unassigned ASVs were removed from datasets.

Average richness is shown with one standard deviation.

Figure 1: Relative abundance of different eukaryotic lineages, prokaryotic phyla, and viral clades. Bar plots show relative read percentages of generated sequencing data for eukaryotic lineages of (A) NanoPlus and (B) PicoNano size fraction, and (C) daily cell count of HAB species. (D) Relative abundance of prokaryotic phyla. (E) Relative abundance of *Mimiviridae*. Colors in the background represent seasons (blue: winter, green: spring, red: summer, and yellow: fall).

Figure 2: NMDS ordination of different communities. Strong seasonality for every community was detected at station "j". Samples clustered together according to sampling months. Different colors show different months, shapes represent years. Stress of the NMDS was between 0.09 and 0.13 (prokaryotic NMDS: 0.09, PicoNano NMDS: 0.11, NanoPlus NMDS: 0.13, *Mimiviridae* NMDS: 0.11).

Figure 3: Dissimilarity metrics according to seasons. Violin plot shows dissimilarities of samples collected in opposite seasons (i.e., summer *vs.* winter or spring *vs.* fall). Communities were equally dissimilar between spring and fall as between summer and winter. Average Sørensen-

Dice dissimilarity was lower between datasets from the same season (PicoNano: 0.72 ± 0.08, NanoPlus: 0.78 ± 0.06, Prokaryotes: 0.58 ± 0.09, and *Mimiviridae*: 0.78 ± 0.11; average dissimilarity ± 1 standard deviation) than between datasets from opposite seasons (PicoNano: 0.83 ± 0.04, NanoPlus: 0.89 ± 0.04, Prokaryotes: 0.77 ± 0.09, and *Mimiviridae*: 0.94 ± 0.03).

Figure 4: Sørensen-Dice dissimilarity metrics for the four different communities. (A–D) Dot plots showing pairwise dissimilarities of all samples plotted against sampling interval (days). (E–F) Boxplots showing dissimilarity and average dissimilarity of the four communities at three different time points.

Figure 5: Four most abundant ASVs of the subsampled datasets of each microbe-group and their taxonomic assignment. Eukaryotes are plotted in green and blue, prokaryotes in red, and *Mimiviridae* in purple. Eukaryotes are plotted in blue if their ASV was one of the four most abundant ASVs in both size fractions.

Figure 6: Recurrence of abundant ASVs. Line plots show number of samples in which abundant ASVs were present (i.e., recurrence). (A) Percentage of abundant ASVs and how often they were found. (B) Recurrence as indicated by absolute number of abundant ASVs and number of samples in which they were present. Most *Mimiviridae* ASVs were found in fewer than ten samples, while roughly half of all eukaryotic ASVs were found in fewer than ten samples. Most prokaryotic ASVs were found in more than half of all samples. An ASV was considered abundant if it generated at least 1% relative reads in one sample.

Figure 7: Schematic, visual representation of community memory model. (A) Seasonally dependent dissimilarity of one community over a longer sampling period with seasonal fluctuations (dark cyan) and fading community memory (grey). (B) Line plots showing changes in community dissimilarity over the short period represented in our datasets. *Mimiviridae* communities (purple) were more dissimilar than were eukaryote (green) and prokaryote (red) communities.

Fig. S1: Cell counts, biotic, and abiotic data collected over 6 years. Blue dotted lines indicate sampling period during which amplicon sequencing was performed.

Fig. S2: Monthly and seasonal summary of cell counts, biotic, and abiotic data. (A–L) Boxplots showing monthly and seasonal biotic and abiotic data. (M) Bar plot showing monthly sum of daily cell counts throughout 6 years.

Fig. S3: Rarefaction curves of all samples. (A) Rarefaction curves for subsampled datasets. (B) Rarefaction curves for data without subsampling.

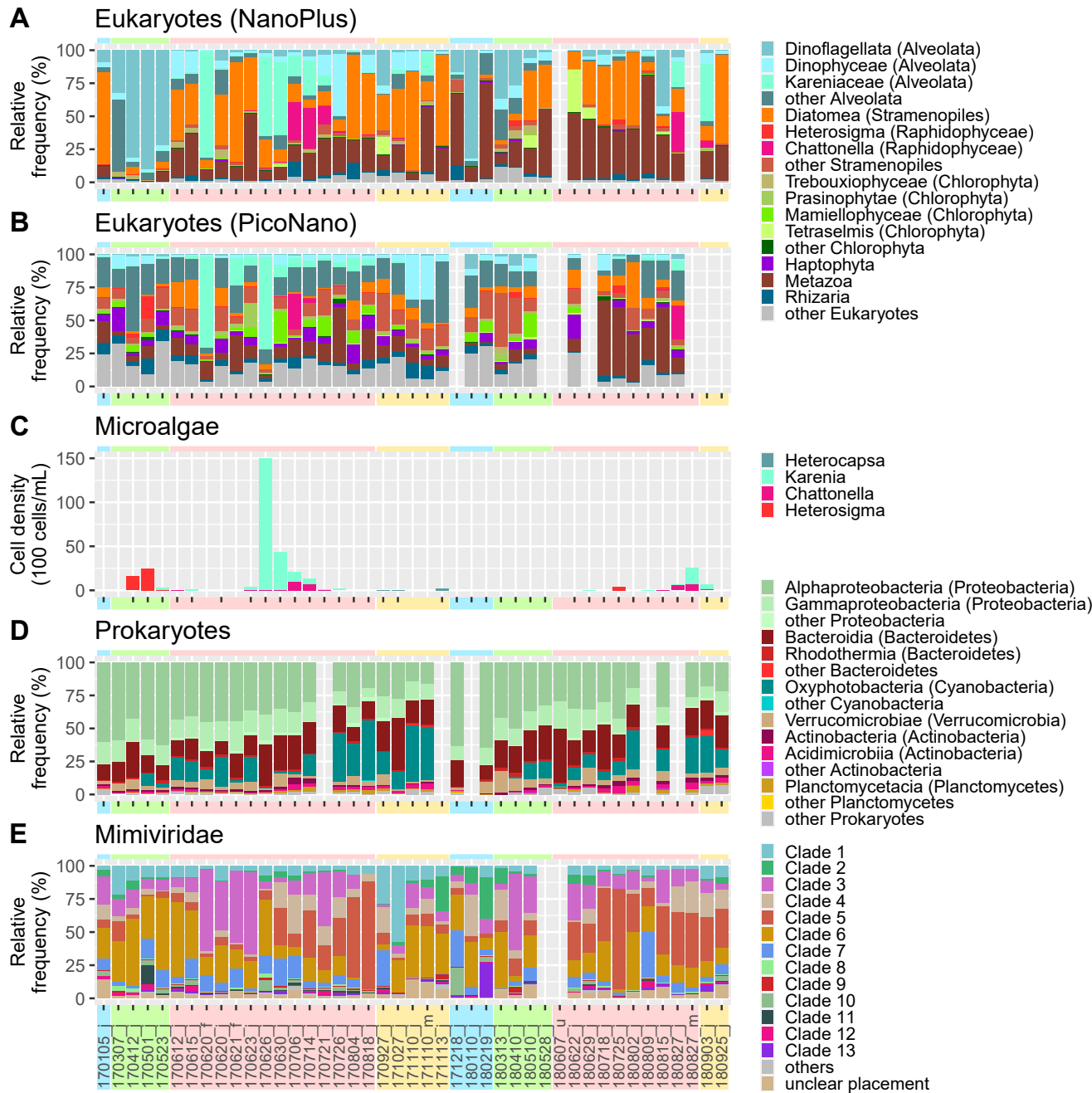
Fig. S4: Visualization of different diversity indices across all datasets subsampled at 8,000 reads. (A) Shannon's diversity index; (B) Richness of different communities; (C) Pielou's evenness index.

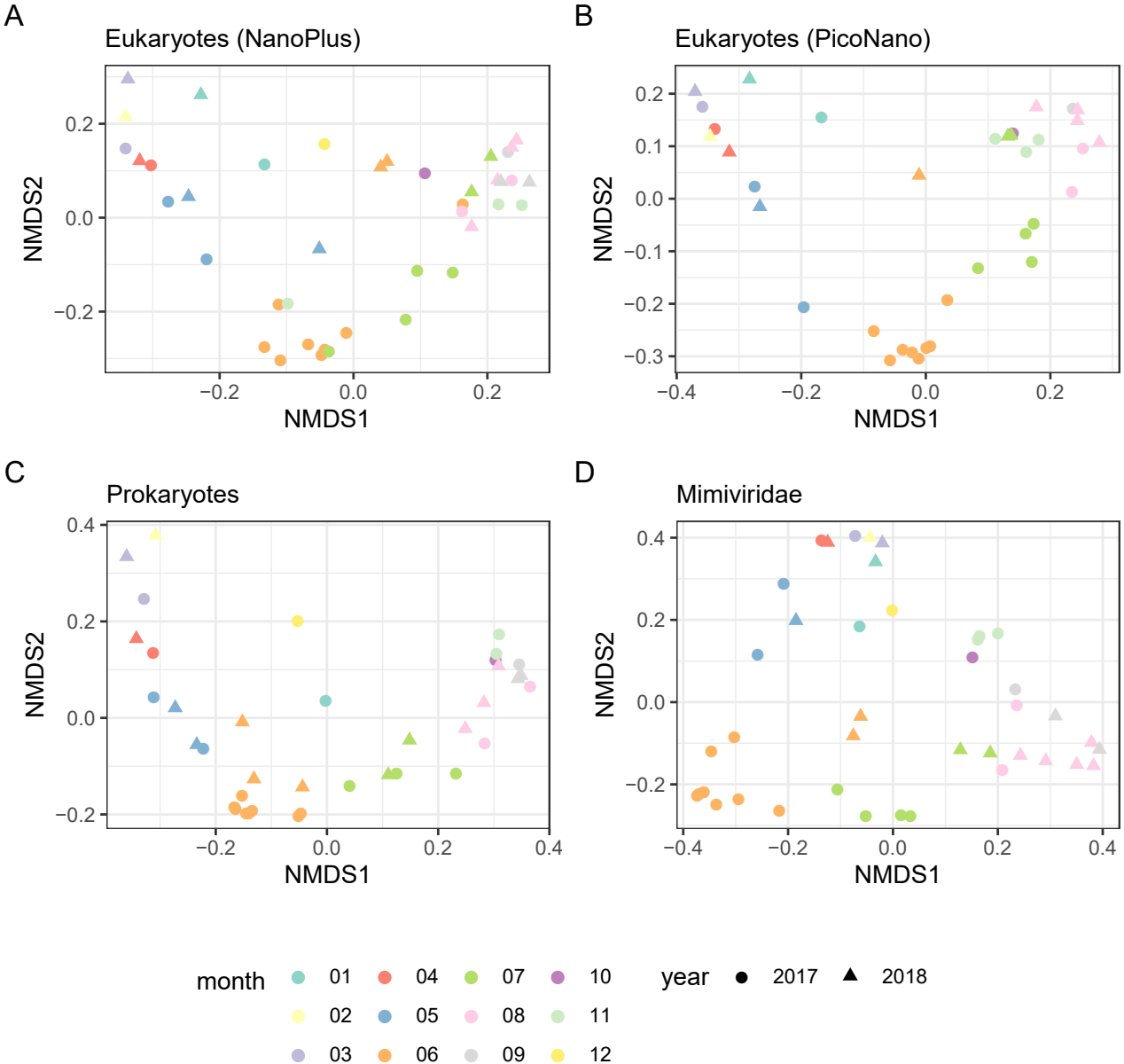
Fig. S5: Phylogenetic tree with *Mimiviridae* ASV placements. Names of leaves correspond to known reference *Mimiviridae*. Unnamed leaves are PolB amino acid sequences assembled from metagenomic data. Red dots indicate placement of sequences generated from *Mimiviridae PolB* amplicon sequencing.

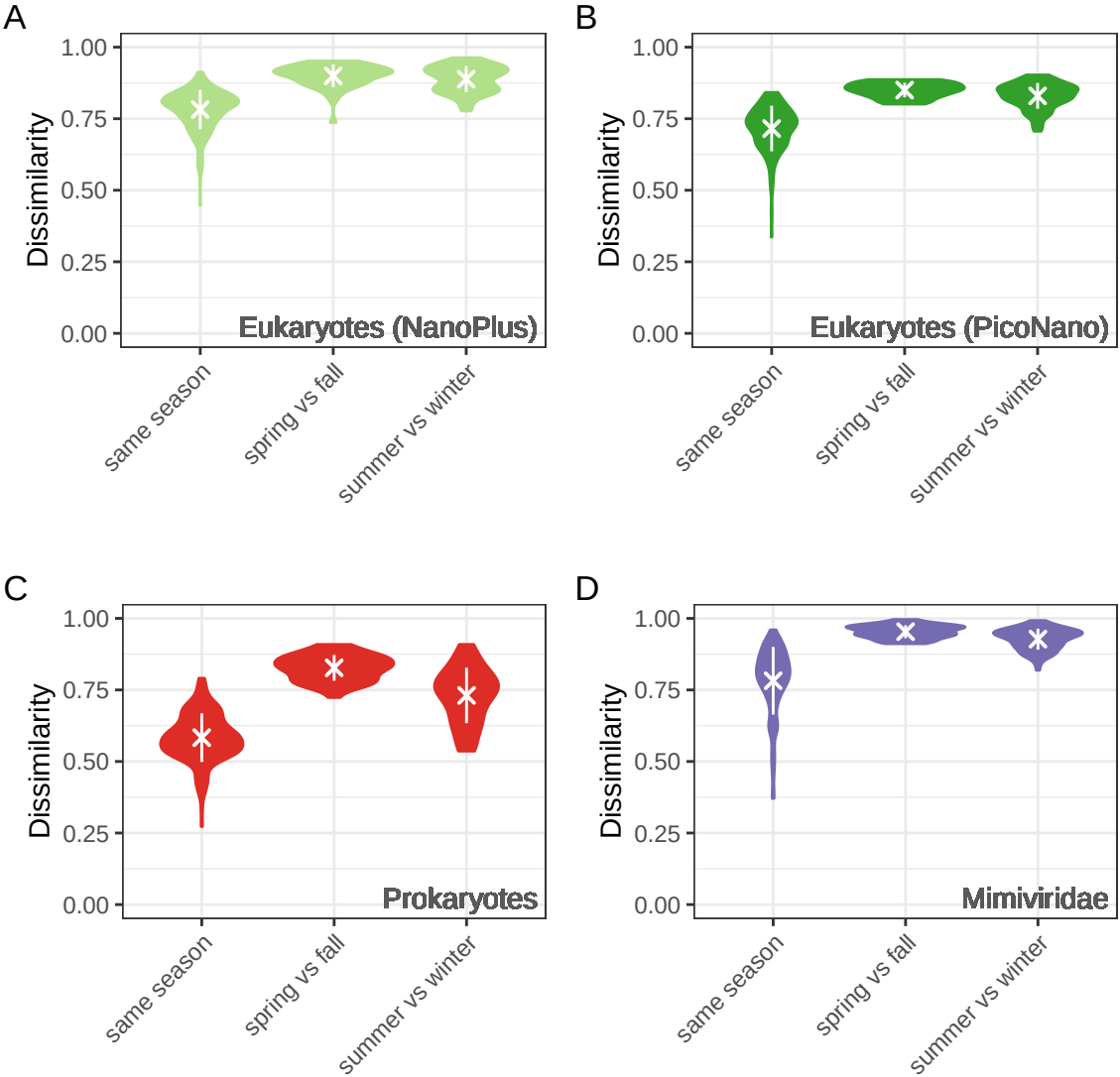
Fig. S6: Correlations of dissimilarity between different communities. All communities are statistically significantly correlated with each other ($p < 10^{-4}$).

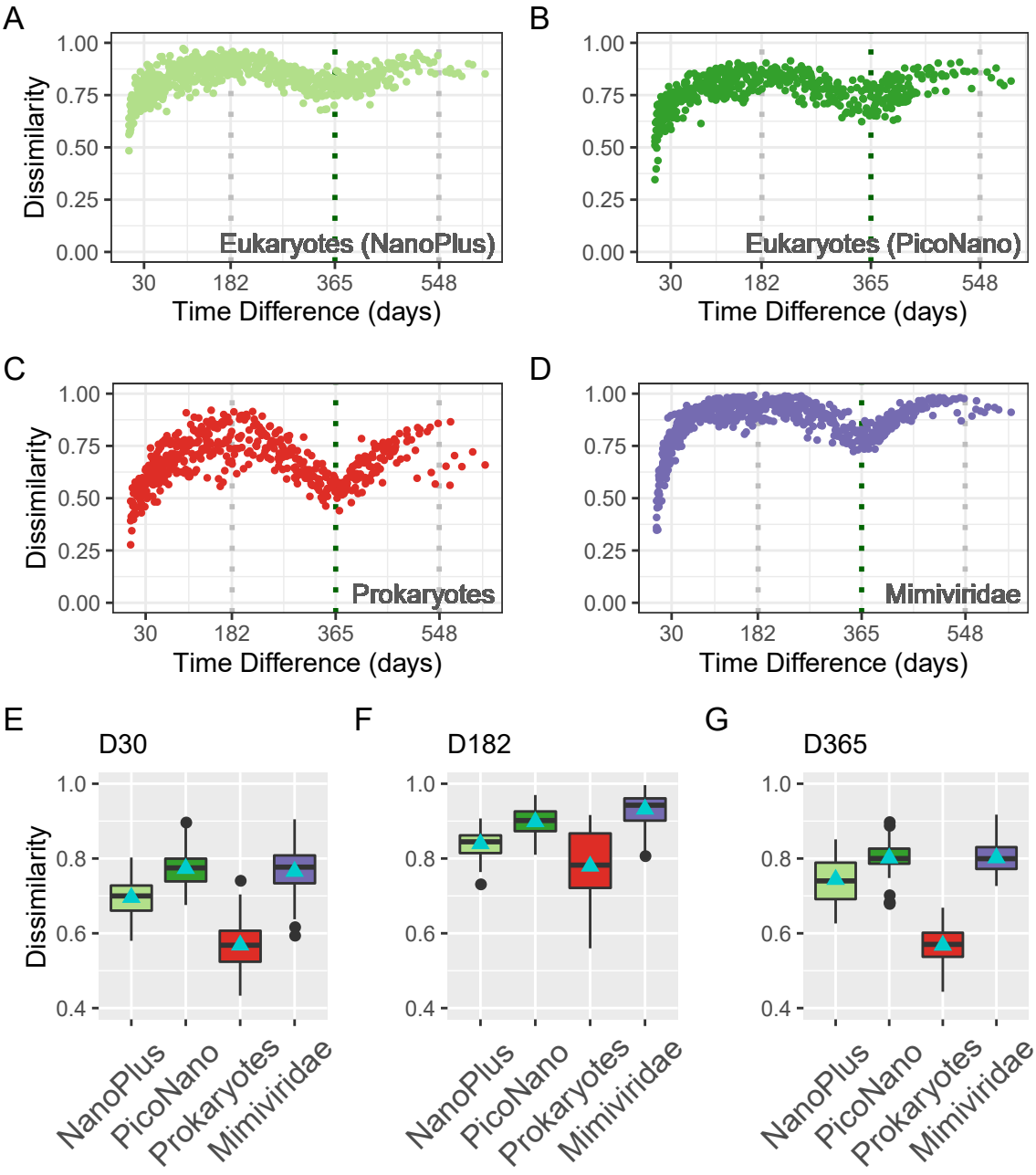
Fig. S7: Persistence of abundant ASVs in different microbial communities. (A) Percentage of abundant ASVs that were persistent (i.e., continually found in a series of samples). (B) Absolute number of abundant ASVs and their persistence.

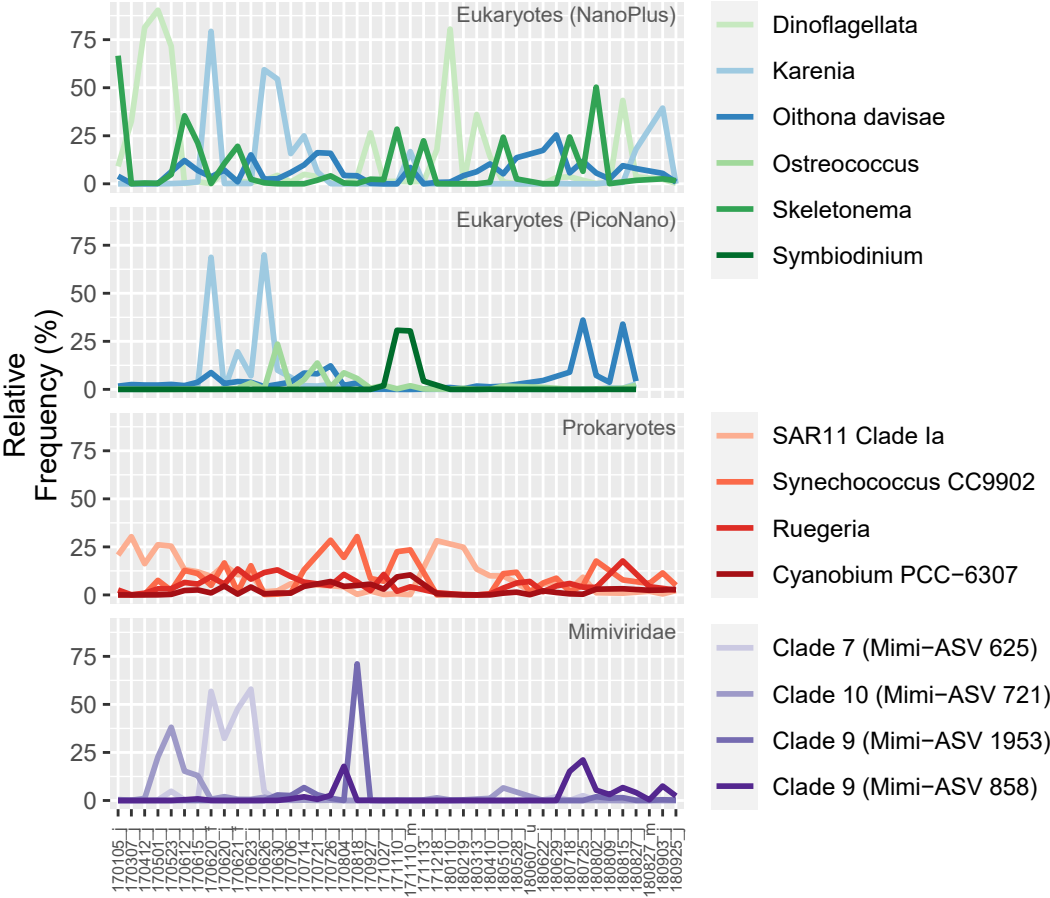
Dataset S1: Microsoft Excel file contains environmental data and cell count data obtained from the Kochi Prefectural Fisheries Research Institute.

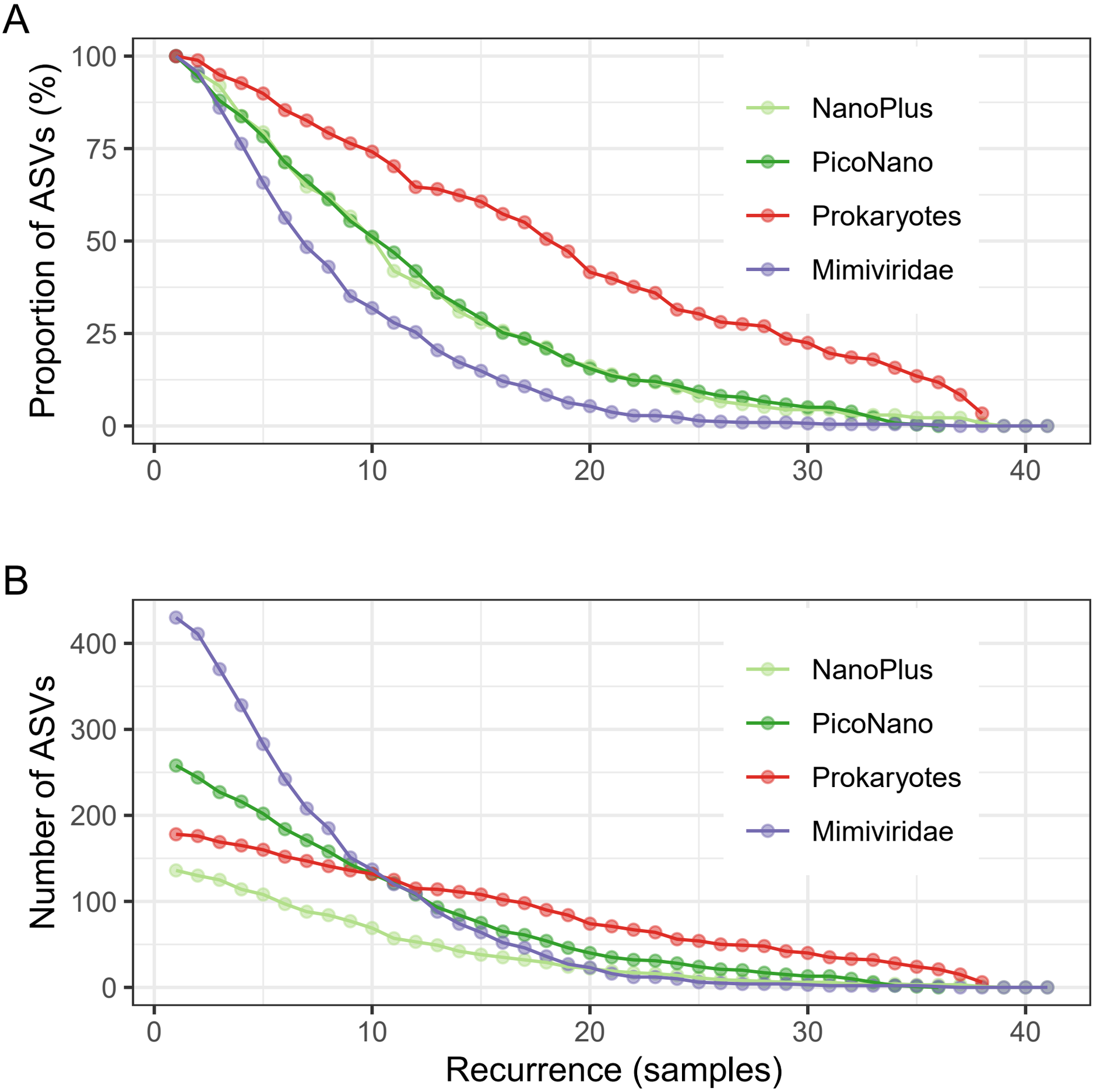

A
Eukaryotes (NanoPlus)
Relative frequency (%)
100
75
50
25
0
Dinoflagellata (Alveolata)
Dinophyceae (Alveolata)
Kareniaceae (Alveolata)
other Alveolata
Diatomea (Stramenopiles)
Heterosigma (Raphidophyceae)
Chattonella (Raphidophyceae)
other Stramenopiles
Trebouxiophyceae (Chlorophyta)
Prasinophytae (Chlorophyta)
Mamiellophyceae (Chlorophyta)
Tetraselmis (Chlorophyta)
other Chlorophyta
Haptophyta
Metazoa
Rhizaria
other Eukaryotes
B
Eukaryotes (PicoNano)
Relative frequency (%)
100
75
50
25
0
C
Microalgae
Cell density (100 cells/mL)
150
100
50
0
Heterocapsa
Karenia
Chattonella
Heterosigma
D
Prokaryotes
Relative frequency (%)
100
75
50
25
0
Alphaproteobacteria (Proteobacteria)
Gammaproteobacteria (Proteobacteria)
other Proteobacteria
Bacteroidia (Bacteroidetes)
Rhodothermia (Bacteroidetes)
other Bacteroidetes
Oxyphotobacteria (Cyanobacteria)
other Cyanobacteria
Verrucomicrobiae (Verrucomicrobia)
Actinobacteria (Actinobacteria)
Acidimicrobiia (Actinobacteria)
other Actinobacteria
Planctomycetacia (Planctomycetes)
other Planctomycetes
other Prokaryotes
E
Mimiviridae
Relative frequency (%)
100
75
50
25
0
Clade 1
Clade 2
Clade 3
Clade 4
Clade 5
Clade 6
Clade 7
Clade 8
Clade 9
Clade 10
Clade 11
Clade 12
Clade 13
others
unclear placement
170105
170307
170412
170501
170523
170612
170615
170620_f
170620
170621_f
170623
170626
170630
170706
170714
170721
170726
170804
170818
170927
171027
171110
171110_m
171113
171218
180110
180219
180313
180410
180510
180528
180607_u
180622
180629
180718
180725
180802
180809
180815
180827
180827_m
180903
180925

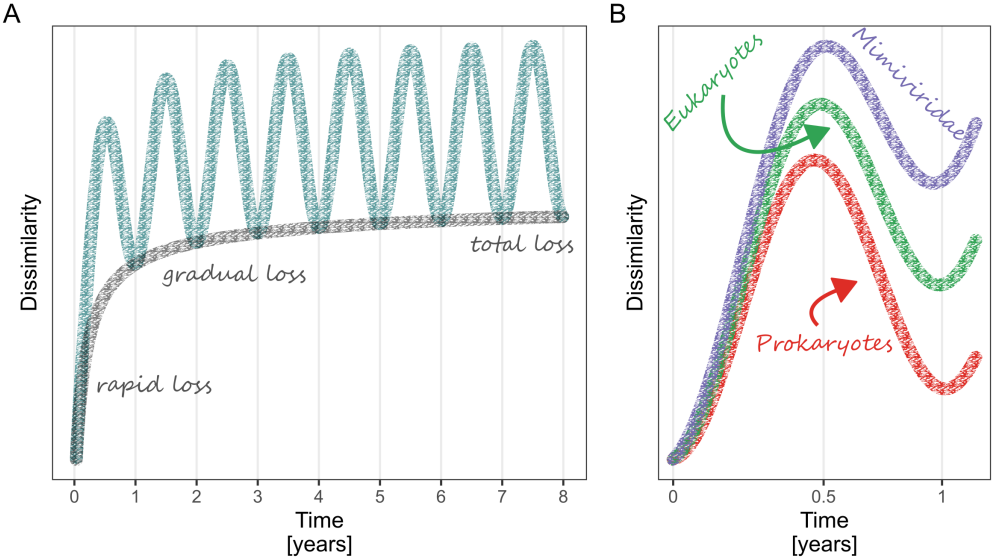

A
Eukaryotes (NanoPlus)
NMDS2
NMDS1
B
Eukaryotes (PicoNano)
NMDS2
NMDS1
C
Prokaryotes
NMDS2
NMDS1
D
Mimiviridae
NMDS2
NMDS1
month
01
02
03
04
05
06
07
08
09
10
11
12
year
2017
2018


A
B
C
D
Dissimilarity
1.00
0.75
0.50
0.25
0.00
Eukaryotes (NanoPlus)
Eukaryotes (PicoNano)
Prokaryotes
Mimiviridae
same season
spring vs fall
summer vs winter


A
B
C
D
E
F
G
Dissimilarity
Time Difference (days)
Eukaryotes (NanoPlus)
Eukaryotes (PicoNano)
Prokaryotes
Mimiviridae
1.00
0.75
0.50
0.25
0.00
30
182
365
548
D30
D182
D365
1.0
0.8
0.6
0.4
NanoPlus
PicoNano
Prokaryotes
Mimiviridae


Eukaryotes (NanoPlus)
Eukaryotes (PicoNano)
Prokaryotes
Mimiviridae
Relative Frequency (%)
0
25
50
75
Dinoflagellata
Karenia
Oithona davisae
Ostreococcus
Skeletonema
Symbiodinium
SAR11 Clade Ia
Synechococcus CC9902
Ruegeria
Cyanobium PCC−6307
Clade 7 (Mimi−ASV 625)
Clade 10 (Mimi−ASV 721)
Clade 9 (Mimi−ASV 1953)
Clade 9 (Mimi−ASV 858)
170105
170307
170412
170501
170523
170612
170615
170620
170621
170623
170626
170630
170706
170714
170721
170726
170804
170818
170927
171027
171110
171110_m
171113
171218
180119
180313
180410
180528
180607_u
180622
180629
180718
180725
180802
180809
180815
180827
180827_m
180903
180925


A
Proportion of ASVs (%)
100
75
50
25
0
0
10
20
30
40
NanoPlus
PicoNano
Prokaryotes
Mimiviridae
B
Number of ASVs
400
300
200
100
0
0
10
20
30
40
NanoPlus
PicoNano
Prokaryotes
Mimiviridae
Recurrence (samples)


A
Dissimilarity
rapid loss
gradual loss
total loss
0
1
2
3
4
5
6
7
8
Time
[years]
B
Dissimilarity
Eukaryotes
Mimiviridae
Prokaryotes
0
0.5
1
Time
[years]