Cyanobacteria are among the most abundant photosynthetic organisms in the oceans; viruses infecting cyanobacteria (cyanophages) can alter cyanobacterial populations, and therefore affect the local food web and global biochemical cycles. These phages carry auxiliary metabolic genes (AMGs), which rewire various metabolic pathways in the infected host cell, resulting in increased phage fitness. Coping with stress resulting from photodamage appears to be a central necessity of cyanophages, yet the overall mechanism is poorly understood. Here we report a novel, widespread cyanophage AMG, encoding a fatty acid desaturase (FAD), found in two genotypes with distinct geographical distribution. FADs are capable of modulating the fluidity of the host’s membrane, a fundamental stress response in living cells. We show that both viral FAD (vFAD) families are Δ9 lipid desaturases, catalyzing the desaturation at carbon 9 in C16 fatty acid chains. In addition, we present a comprehensive fatty acid profiling for marine cyanobacteria, which suggests a unique desaturation pathway of medium- to long-chain fatty acids no longer than C16, in accordance with the vFAD activity. Our findings suggest that cyanophages are capable of fiddling with the infected host’s membranes, possibly leading to increased photoprotection and potentially enhancing viral-encoded photosynthetic proteins, resulting in a new viral metabolic network.

The ISME Journal (2018) 12, 343–355; doi:10.1038/ismej.2017.159; published online 13 October 2017

Introduction

Viruses are the most abundant entity in the oceans, yet the vast majority remains uncultured (Suttle, 2005; Huang et al., 2010; Hurwitz and Sullivan, 2013; Brum et al., 2015). Cells lysed by viruses contribute to energy and nutrient flux in the oceans, while infected cells could also affect global biogeochemical cycles (Fuhrman, 1999; Wilhelm and Suttle, 1999; Hurwitz et al., 2013; Lisle and Robbins, 2016; Roux et al., 2016). Viruses carry in their genomes a wide variety of auxiliary metabolic genes (AMGs), capable of complementing or redirecting the infected host metabolism resulting in increased viral fitness (Breitbart et al., 2007). Cyanophages, phages infecting marine cyanobacteria, display a broad array of AMGs, including photosynthetic light reaction components (Mann et al., 2003; Lindell et al., 2004; Millard et al., 2004, 2009; Zeidner et al., 2005; Sullivan et al., 2006, 2010; Sharon et al., 2009, 2011; Zheng et al., 2013). Photosystem-I (PSI) genes in cyanophages (viral PSI (vPSI)) are found in two main genotypes, arranged in cassettes of seven (psaF, C, A, B, K, E and D) and four (psaD, C, A and B) genes, dubbed vPSI-7 and vPSI-4, respectively (Sharon et al., 2009; Beja et al., 2012; Roitman et al., 2015; Fridman et al., 2017). Since there are no cultured representatives of vPSI-4 phages, little is known regarding their potential influence on the infected host metabolic capacities.

Several AMGs are potentially involved in photoprotection of the infected cyanobacterial cell. For example, high light-inducible proteins enable the dissipation of excess light energy and the correct functioning of the photosynthetic light reactions (Havaux et al., 2003), and are widely found in cyanophages (Lindell et al., 2004; Millard et al., 2004; Sullivan et al., 2005). Photosystem II (PSII) reaction center protein D1 (encoded by the psbA gene) was shown to be constantly damaged during photosynthetic activity and must be repaired and de novo synthesized to maintain active photosynthesis (Adir et al., 2003). The viral psbA gene is expressed upon infection (Lindell et al., 2005, 2007; Clokie et al., 2006) and it was suggested to increase phage
fitness (Bragg and Chisholm, 2008; Hellwegner, 2009). In addition, many cyanophages carry genes for a plastoquinol terminal reductase, potentially involved in photoprotection of PSII (Wegele et al., 2007; Millard et al., 2009; Sullivan et al., 2010). Based on the accumulating data in cyanophage AMG repertoire, it appears that photoprotection of the cell is a central need of the infected cell, the ‘virocell’ (Forterre, 2013) metabolism.

Another, rather unexplored mechanism for coping with photoinhibition in cyanobacteria includes the desaturation of the membranes lipids. Unsaturated fatty acids are critical for growth and for coping with stress in cyanobacterial cells, including photoinhibition, cold adaptation and osmotic stress (Sato and Murata, 1981; Huflejt et al., 1990; Wada et al., 1990, 1992; Tasaka et al., 1996; Gombos et al., 1997). Membrane fluidity affects the assembly and performance of membrane proteins, including the de novo synthesis and activation of D1, leading to a higher recovery rate of PSII activity, and therefore reducing photoinhibition (Gombos et al., 1997). In cyanobacteria, lipid desaturation is performed on fatty acid residues esterified to a glycerolipid by membrane-bound acyl-lipid front-end desaturases (Des proteins), associated with cytoplasmic and thylakoid membranes. Molecular oxygen and an electron donor (ferredoxin) are required for fatty acid desaturation (FAs) activity (Sato and Murata, 1981; Wada et al., 1993; Shanklin and Cahoon, 1998). Four des genes can be found in cyanobacteria, encoding for DesA, DesB, DesC and DesD FADs proteins, catalyzing the desaturation at carbon Δ12, Δ15, Δ9 and Δ6 (counting from the carboxy group), respectively (Wada et al., 1990; Reddy et al., 1993; Sakamoto et al., 1994a, b). Cyanobacteria have been classified into four groups based on their fatty acid composition, depending on the length of their fatty acids (mainly C16 or C18), the amount of the double bonds (zero to four per fatty acid chain) and the sn position of the desaturated fatty acid (sn-1 and/or sn-2 at the glycerol backbone) (Wada and Murata, 1998). However, marine unicellular cyanobacteria, namely Synechococcus and Prochlorococcus, do not fit into any of the four classic groups based on their FAD composition, carrying only desC and desA genes (Chi et al., 2008). DesC performs the first desaturation of fatty acids at position Δ9 and is present in all cyanobacterial strains (Wada and Murata, 1998; Chi et al., 2008). DesC is constitutively expressed (Los et al., 1997; Kis et al., 1998), has the most significant effect on the fluidity of the membrane (Bosse and Martin, 1989; Los et al., 1997) and can respond to environmental changes (for example, temperature) within hours and without de novo synthesis of fatty acids (Sato and Murata, 1981). These monounsaturated fatty acids are essential for growth. Consequently, desC-knockout mutants must be supplemented with unsaturated fatty acids to survive (Resnick and Mortimer, 1966; Tasaka et al., 1996).

Here, we report the identification and characterization of two novel and widespread cyanophage-encoded FAD (vFAD) families. The vFADs were expressed using a heterologous yeast system and were identified as DesC-like FADs, catalyzing the desaturation at carbon Δ9 in C16 fatty acid chains. In addition, we performed a comprehensive fatty acid analysis of marine picocyanobacteria, including Prochlorococcus and Synechococcus strains, and found their lipid composition to be different from other cyanobacteria. Our results suggest that marine cyanobacteria have a rare pathway for fatty acid desaturation, and phages desaturases are well suited to fit in.

Results and discussion

To enrich our knowledge regarding uncultured cyanophages carrying photosynthetic genes, we conducted a metagenomic survey in a reassembled database (Philosof et al., 2017) of the microbiome (Sunagawa et al., 2015) and virome (Brum et al., 2015) data sets from the Tara Oceans expedition, a comprehensive sampling project of oceanic microbial diversity. Using the sequence of a viral PSI psaD gene as query for TBLASTX, we identified a 64 kbp contig containing a vPSI-4 cassette in the assembly of station 70 (South Atlantic Ocean). The contig was extended up to 94 kbp with recruitment of reads from the same station. This contig is predicted to have originated from a cyanophage of the Myoviridae family (T4-like phages), based on RegA (Supplementary Figure 1a) and Transaldolase (Supplementary Figure 1b) maximum-likelihood phylogenetic protein trees, and the presence of three transfer RNA genes (Figure 1 and Supplementary Table 1) widely found among cyanomyophages (Enav et al., 2012). The contig contains structural and DNA replication genes resembling those of cyanomyophages, along with various AMGs common in cyanophages, such as talC (Sullivan et al., 2005; Ignacio-Espinoza and Sullivan, 2012), peptide deformylase (Sharon et al., 2011), psbA and psbD (Mann et al., 2003; Lindell et al., 2005), ferredoxin (Sullivan et al., 2005; Ignacio-Espinoza and Sullivan, 2012), phoH (Goldsmith et al., 2011), among others (Figure 1, Supplementary Figure 2 and Supplementary Table 1). Surprisingly, we also identified a gene coding for a putative vFAD, this being the first report of a cyanophage potentially interfering with fatty acid metabolism in the infected host cell. Using the identified vFAD gene sequence as bait, we were able to retrieve 139 contigs containing vFADs among various viral genes (Supplementary File 1) from publicly available metagenomic data sets (Supplementary Table 2) using the same strategy applied to vPSI-4 genes. The viral origin of the contigs was confirmed by the VirFinder Software (Ren et al., 2017) (Supplementary Table 3). With the exception of 11
contigs encoding solely for a partial DesC (which were not used in further analysis), all contigs were identified as belonging to cyanomyophages based on similarity of the various open reading frames to cultured cyanomyophage isolates using BLAST (Supplementary Table 3).

The vFAD gene encodes for a putative acyl-lipid desaturase, a membrane-bound enzyme that catalyzes the front-end desaturation of fatty acids esterified to glycerolipids. The protein is homologous to membrane-bound DesC \(\Delta 9\) front-end desaturases found in cyanobacteria (and plants) (Figure 2). Moreover, it contains the three characteristic histidine motifs of DesC-like desaturases, two HXXXXH and a HXXXXH, potential ligands of diiron center in the active site of the enzyme (Wada and Murata, 1998). \(\Delta 9\) desaturases from cyanobacteria have been classified phylogenetically into six clades (Chi et al., 2008). Clades \(\Delta 9\)-3 and \(\Delta 9\)-4 (colored in green and blue, respectively, in Figure 2) are composed solely of marine picocyanobacteria, whereas the remaining four clades (shaded in gray in Figure 2) include marine and freshwater cyanobacteria, as well as eukaryotic algae. Interestingly, the estuarine Synechococcus CB0101 (Marson et al., 2014) carries three genes encoding for DesC proteins, one of them clustering separately from the previously described (CB0101 III in Figure 2). Using this protein sequence as a bait, we recruited three new contigs from three Tara Oceans marine stations (137 and 138 in the North Pacific Ocean, and 141 in the North Atlantic Ocean), carrying a similar desaturase (Supplementary File 3). These contigs seem to originate from picocyanobacteria (Supplementary Table 4) and cluster together in a monophyletic group. Moreover, the four encoded proteins share unique motifs in their histidine boxes (Supplementary File 4), thus unveiling a seventh clade of cyanobacterial \(\Delta 9\) proteins (shaded in pink in Figure 2). Other Synechococcus and Prochlorococcus strains (mainly low light adapted, classified as clade IV) have two types of DesC and these proteins cluster separately into two different branches in the DesC phylogenetic tree (Figure 2), indicating a possible specialization for each type. Since the marine picocyanobacterial FAD-specific activity is yet unknown, we will refer to them as \(\Delta 9\)-3 and \(\Delta 9\)-4, according to the classification given by Chi et al. (2008). However, some Prochlorococcus strains carry a single desC gene corresponding to \(\Delta 9\)-4 (shaded blue and marked with an asterisk in Figure 2), whereas some Synechococcus strains contain only one desC \(\Delta 9\)-3 gene (shaded green and marked with an asterisk in Figure 2). Accordingly, cyanophage-encoded vFADs can be found in two genotypes, forming monophyletic branches in the phylogenetic tree. These groups correspond to the unicellular marine picocyanobacterial types, although they share <70% identity on the protein level and have distinct H-box motifs (Supplementary File 4), thus we decided to denominate them vFAD-I (Figure 2, shaded gold) and vFAD-II (Figure 2, shaded purple). We retrieved more vFAD-I contigs than vFAD-II from the metagenomic data sets analyzed; however, the first vFAD discovered, found in the 94 kbp contig, clusters within family II (marked with a black arrow in Figure 2). vFAD families show distinct biogeography (Figure 3a). vFAD-IIs are widespread in the oceans (Figure 3a, golden dots), being found all along the Pacific and Atlantic Oceans, the Indian Ocean and the Mediterranean and Red Seas. In contrast, vFAD-IIs are present only in the Southern Pacific and Southern Atlantic Oceans, as well as in the Indian Ocean (Figure 3a, purple dots). Interestingly, the geographical distribution and abundance of vFAD-II resembles the data found for uncultured phages carrying the vPSI-4 gene cassette (Roitman et al., 2015), which is also found in the 94 kbp contig (Figure 1). To estimate the vFAD relative abundance, we mapped the raw reads from the Tara Ocean metagenomes corresponding to bacterial, giant viruses and viral fractions to the viral desC genes. Based on the recruitment for desC genes of each family, we found that vFAD-I was 46 times more abundant than vFAD-II. The relative abundance of cyanomyophages carrying vFADs of family I among cyanomyophages in positive stations (Figure 3) was estimated to be up to 34% with an average of 7%; vFADs of family II were estimated to be present in up to 3.5% of total cyanomyophages, with an average of 0.1% (Supplementary Table 5). It is worth noting that vFADs were found in all three size fractions, in accordance with Philosof et al. (2017) findings that cyanophages can be widely found in bacterial fractions, probably due to ongoing infections during
Figure 2  Maximum-likelihood phylogenetic tree of DesC. Viral FADs classified as families I and II are shaded in gold and purple, respectively. Picocyanobacterial desaturases are shaded in green and blue for Δ9-3 and Δ9-4 groups, respectively. DesC sequences corresponding to groups Δ9-1, Δ9-2, Δ9-5 and Δ9-6 (Chi et al., 2008) are shaded in grey. Cyanobacterial newly proposed Δ9-7 group is shaded in pink. Black and gray arrows indicate the sequences chosen for expression in yeast. Stars indicate picocyanobacterial strains carrying only one desC gene. The scale bar indicates the average number of amino-acid substitutions per site. Circles represent bootstrap values > 0.9.
sampling (Supplementary Figure 3). To identify possible hosts for the vFAD-carrying phages, the abundance of marine *Synechococcus* and *Prochlorococcus* was evaluated by mapping sequences of the taxonomical marker *petB* reported in Farrant et al. (2016), corresponding to 49 different ‘ecologically significant taxonomic units’, on the same samples used to estimate the abundance of vFADs. We found that the abundance of the viral desC genes of vFAD-I was highly correlated ($R^2 = 0.91$, $P < 0.001$) to the abundance of *petB* originating from *Prochlorococcus* low light clade I (ecologically significant taxonomic unit LLIA) in the North Atlantic Ocean (Supplementary Figure 4). Owing to the low number of samples positive for vFAD-II, we could not detect any significant correlation. Interestingly, the majority of the reads (>90%) for vFAD-II originate from the giant virus fraction (0.45–0.8 μm) (Supplementary Figure 3), which could include whole *Prochlorococcus* cells. This suggests *Prochlorococcus* as the possible host for these phages.

To confirm the vFAD activity, we expressed the viral genes in a heterologous system using the *Saccharomyces cerevisiae* strains INVSc2 and the FAD mutant Ole1 (Stukey et al., 1990). While the INVSc2 strain contains monounsaturated (at position Δ9) and saturated long-chain C16 and C18 fatty acids (Supplementary Figure 5), the Ole1 mutant strain features only saturated fatty acids (Figure 4a) and has to be supplemented with unsaturated fatty acids for normal growth. The lipid profile of INVSc2 cells expressing vFADs could not be distinguished from cells transformed with an empty vector, suggesting for a possible Δ9 desaturation activity (data not shown). This was confirmed by lipid profiles of Ole1 mutant strains expressing vFADs; both vFADs show Δ9 desaturase activity, acting specifically on C16 chains of lipids in yeast (Figures 4b and c). No activity of vFADs on C14 fatty acids could be detected.

**Figure 3** (a) Map of *Tara Oceans* stations analyzed in this project. Gold dots represent stations positive for vFAD-I reads; purple dots mark stations positive for vFAD-II reads. Gray dots stand for stations where no reads for vFADs were found. Latitudes are marked at the left of the map. Oceanic regions are delimited according to the *Tara Oceans* Expedition labeling. (b) Relative abundance of vFADs from families I and II (depicted in gold and purple, respectively), presented in reads per kilobase per million (RPKM), was measured using the *Tara Oceans* metagenomes corresponding to bacterial, giant virus and viral fractions. Box plots were created using a median, 25th percentile, 75th percentile, minimum, maximum and outliers depicted. Whenever the amount of samples was less than five per region per fraction, individual dots are presented.
acids were detected, even when yeast cultures were supplemented with 0.01% myristic acid (data not shown). Marine picocyanobacteria show a potentially unique pathway for acyl-lipid desaturation among cyanobacteria, containing only desA and desC genes for desaturation of carbons Δ9 and Δ12, respectively (Chi et al., 2008), yet their lipid profiles were scarcely determined. Previous work showed fatty acid profiles of two Prochlorococcus strains, Med4 and MIT9313 (Biller et al., 2014). To increase our understanding of marine picocyanobacterial fatty acids, we performed a fatty acid profiling of eight cyanobacterial strains, including both Synechococcus and Prochlorococcus corresponding to the three main picocyanobacterial FAD genotypes. We analyzed strains carrying two desaturases, types Δ9-3 and Δ9-4, Synechococcus WH7803 and WH7805 and Prochlorococcus MIT9313; strains carrying only a Δ9-4, Prochlorococcus Med4 (axenic and non-axenic cultures) and NATL2A; and strains carrying only a Δ9-3, Synechococcus WH8109 and WH8102 (Figures 5a and b and Supplementary Figure 6). We also analyzed Prochlorococcus MIT9312, whose genome is not sequenced yet, and therefore its genotype is unknown, although based on its phylogeny (high light adapted, clade II) we hypothesize it might carry a Δ9-4 (Supplementary Figure 6). All marine picocyanobacterial strains show a distinct fatty acid profile, containing a large amount of C14 fatty acids chains compared with freshwater cyanobacteria (Supplementary Figure 6) (Lang et al., 2011). Interestingly, we could not detect C18:0 fatty acids in any of our cultures and only three strains (Synechococcus WH8109 and WH7805, and 2/5 cultures for Prochlorococcus MED4) showed C18:1 fatty acids. This is in contrast to previous reports, where these fatty acids could add up to 10% of the total fatty acid content of the cells (Biller et al., 2014). Those cultures were all non-axenic, meaning that the C18:1 could have originated from other organisms in the media, although based on Biller et al. (2014) results, who worked with axenic Med4 cultures, this fatty acid could be of picocyanobacterial origin. We speculate that the different growth conditions of the cultures used in the studies had affected their fatty acids composition, leading to the synthesis/absence of C18:0 and C18:1 (light intensity, culture volume, stirring, etc.). This suggestion is supported by the complete absence of C18:1 fatty acids in our axenic cultures while Biller et al. (2014) detected those fatty acids to be up to 10% of the total fatty acids of the same strains (Med4 and MIT9313). We speculate that long fatty acids are not needed under our culturing conditions (see Materials and methods). In addition, previous studies reported C14:0, C16:0 and C16:1 (n-7) to be the most abundant fatty acids in marine phytoplankton (Wakeham and Canuel, 1988) and in marine picocyanobacterial strains (Biller et al., 2014), supporting our results. Although we analyzed strains belonging to three different genotypes regarding the desC gene content, we do not see a distinct desaturation pattern among the picocyanobacterial genotypes, thus we cannot determine a specific activity for Δ9-3 and Δ9-4 cyanobacterial desaturases. However, the fatty acid profiles of the marine picocyanobacteria hint to an unusual substrate specificity of those desaturases (Figure 5c). Picocyanobacterial fatty acid profiles display desaturation at the Δ9 carbon for C14 (n-5) and C16 (n-7) but not in C18 fatty acid chains (n-9) (Figure 5a). In some strains, monounsaturated C18:1 (n-7) could be detected, containing the double bond at position Δ11, thus being the result from elongation of monounsaturated C16 and not of de novo desaturation of saturated C18. We therefore propose that marine picocyanobacterial DesC desaturases have a substrate specificity towards fatty acid chains of C14 and C16 (Figure 5c). However, we cannot discard the possibility that there is little or no synthesis of C18 fatty acid chains in these cyanobacterial strains, thus the lack of substrate could explain their unusual specificity; Biller et al., 2011) did not specify whether the C18:1 detected in their cultures is (n-7) or (n-9).

Based on the vFAD activity assay, acting solely in C16 fatty acids, and the fatty acid profile of marine picocyanobacteria, we propose a model for
Several viruses infecting eukaryotic organisms carry fatty acid metabolism AMGs for lysing the host’s cell (Vardi et al., 2009), to enable the replication of their genome (Lee et al., 2001) or for the biosynthesis of their unique lipids composing the envelope membranes (Ziv et al., 2016). Interestingly, several bacterial-like FADs were recently detected in genomes of Emilia- nia huxleyi viruses (Nissimov et al., 2017). While their activity is yet unknown, it was speculated (Nissimov et al., 2017) that they play a role in the massive remodeling of the fatty acid profiles observed in infected host cells (Evans et al., 2009; Rosenwasser et al., 2014). However, this speculation seems now less favored as this remodeling is characterized by rather higher percentages of saturated fatty acids (Malitsky et al., 2016). Cyanomy- phages, on the other hand, do not contain lipid membrane envelopes and their capsids are composed solely of proteins. We therefore propose that in cyanophages fatty acid metabolism AMGs, that is, vFADs, are carried out to modulate the fluidity of the cytoplasmic or thylakoid membranes of the infected cell. Modulating the cytoplasmic membrane could lead to better lysis, whereas modulating the thylakoid membranes could improve the stress response of the infected cell reducing photodamage and oxidative stress, among other stresses, resulting in better physiological conditions for the ongoing infection. In the 94 kbp contig, we found along with the vFAD, vPSII and vPSI genes, whose activity might benefit from modifications in the thylakoid membrane fluidity, and a gene encoding for ferredoxin, which could potentially act as the electron donor to the vFAD (Figure 1).

Marine Synechococcus and Prochlorococcus are among the most abundant photosynthetic organisms on Earth, and it was estimated that cyanophages lyse between 0.005 and 10% of cyanobacteria daily (Waterbury and Valois, 1993; Suttle and Chan, 1994). During infection, the virocell’s physiology is remarkably different from the original, uninfected cyanobacteria, as phages bring new metabolic capabilities with the potential to rewire the host’s metabolism. Here we report a novel pathway in cyanophages, that is, fatty acid metabolism that could have an overall impact on the virocell’s performance. This might lead to a higher fitness of the phage and to a change in the quality of the debris left after burst, which becomes part of the dissolved organic matter used by heterotrophs and it is shunted back into the food web (Wilhelm and Suttle, 1999). As we keep unveiling rare phage
Photoprotection.

Better stress response, including cold adaptation and the desaturation in the membranes leading to improved lysis and stability of the infected cells. Additionally, phages could increase the desaturation degree in the membranes, leading to higher desaturating C16:0 fatty acids. This might lead to the maintenance of the cytoplasmatic and/or the thylakoid membranes by carrying vFAD genes can increase or maintain the desaturation degree. The scaffold using Bowtie2 (Langmead and Salzberg, 2012). ORFs were identified in the 94 kbp contig using GeneMark (Besemer and Borodovsky, 1999; Zhu et al., 2010) and manually annotated using BLASTX (default parameters) and transfer RNAscan-SE (Lowe and Eddy, 1997). The vFAD protein sequence was used as query for a TBLASTN search (e-value 0.1) against metagenomic data sets (Supplementary Table 2). All retrieved contigs were screened using BLASTX (e-value 10e−10) against the NCBI non-redundant (nr) protein database to identify all putative proteins in the contigs. FADs from cyanophage origin were selected based on top hits with <70% identity to picocyanobacteria.

Relative abundance of vFADs was calculated using Salmon (version 0.8.2) (Patro et al., 2017). A collection of 1150 DNA sequences (Supplementary Table 6) composed of cyanobacterial FADs, the BLASTX identified vFADs, cytochrome b6 (petB) from photosynthetic microorganisms (chloroplasts, freshwater and marine cyanobacteria) and viral marker genes (gp20, gp23, DNAPol, MCP and psaA) were used to create a Salmon index. The index was used for the quantification of the DNA collection in the 399 metagenomes from the Tara Oceans microbial, giant viruses and viral fractions with Salmon in the quasimapping mode with the following parameters ‘—meta—incompPrior 0.0 —libType A —gcBias—seqBias—numBootstraps 100’. Quantification results were processed by tximport (version 1.4.0) (Soneson et al., 2015), followed by the filtering of sequences with <20 mapped reads and normalization with edgeR (version 3.18.1) (Robinson et al., 2010). Reads per kilobase per million were calculated from the normalization results by the edgeR function reads per kilobase per million. Abundance plots were generated in Python (version 3.6.0) using the visualization package Seaborn (version 0.8.0) (Waskom et al., 2016) after grouping and summarization using pandas (version 0.20.1) (McKinney, 2010).

**Materials and methods**

**Metagenomic data analysis**

Metagenomic data sets from the Tara Oceans biome (Sunagawa et al., 2015) and virome (Brum et al., 2015) were reassembled using IDBA-UD (Peng et al., 2012) assembler as described elsewhere (Philosof et al., 2017) providing higher quantity of longer scaffolds than previously reported (Sunagawa et al., 2015). Errors in the assembly were corrected using two read-mapping-based in-house tools as described elsewhere (Philosof et al., 2017). Viral psaD sequences obtained in a previous study of vPSI-4 genes (Roitman et al., 2015) were used as query to recruit scaffolds in the reassembled Tara Ocean data set using TBLASTX (Altschul et al., 1990; Camacho et al., 2009) with the default parameters. One of the identified scaffolds, SAMEA2621085 (station 70, depth 5 m, 0–0.22 filter), contains the four genes of vPSI-4 (psaD, C, A and B). The scaffold carrying the vPSI-4 genes was extended using the miniassembly technique described elsewhere (Sharon et al., 2013). This process leads to the recruitment of other fragments of the same genome until no further elongation could be reached. The resulting 94 kbp fragment went through QC, and consistency of the extended scaffold was confirmed by mapping the sample reads to the scaffold using Bowtie2 (Langmead and Salzberg, 2012).

**vFAD–cyanobacteria correlation analysis**

The positive samples for vFADs were used to perform a linear regression between the normalized and summarized counts of viral desC and cyanobacterial petB from different ecologically significant taxonomic units (Farrant et al., 2016) (Supplementary Table 5), using Python (version 3.6.0) and the ‘ols’ function of the package statsmodels (version 0.8.0) (Skipper and Perktold, 2010). Detection of outliers in the different linear regression analysis was based on the Cook’s distance (D), discarding those with D > 1.

**Geographical distribution of vFADs**

The map was plotted using a custom R script (version 3.4.0) (R Core Team, 2017) and the packages: maps (version 3.2.0) (Becker et al., 2017), ggplot2 (Wickham, 2009) and ggalt (version 0.4.0).
(Rudis et al., 2017). Minor aesthetic adjustments were performed in Inkscape (version 0.92).

Data availability and bioinfomatic analysis
The R scripts and Jupyter (Kluvyer et al., 2016) notebooks used for normalization, abundance estimation, correlation analysis and map plotting are available at: https://github.com/BejaLab/vFADs.

Phylogenetic construction and analysis
Newly identified FADs, and talIC and regA gene sequences were translated to proteins according to the correct open reading frame and aligned along with sequences from picocyanobacteria and cyanophages retrieved from GenBank. Multiple sequence alignments were created using ClustalX v.2.1 (Larkin et al., 2007). Maximum-likelihood phylogenetic trees were constructed using the phylogeny.fr pipeline (Dereeper et al., 2008), including the PhyML v.3.0 (Guindon et al., 2010) and the WAG substitution model for amino acids (Whelan and Goldman, 2001). One hundred bootstrap replicates were performed for each analysis. See Supplementary Files 5–7 for the alignments used to construct the trees.

Expression of vFADs
One representative from each of the vFAD families (SAMEA2621033_16500 for vFAD-I and SAMEA2621085_722 for vFAD-II, marked with a gray and a black arrow, respectively in Figure 2) were chosen for expression. We performed codon usage adaptation for optimal expression in yeast using Integrated DNA Technologies (IDT) tool for codon optimization to vectors (Thermo Fisher Scientific, Waltham, MA, USA) using EcoRI and NotI sites in frame so that the gene is fused to the vector's His-tag at the N terminus of the protein, and sequenced to confirm their identity. The plasmids were transformed into Saccharomyces cerevisiae strain INVSc2 and Ole1 (Ole1) following a modified protocol from Xiao (2006). Individual colonies were grown overnight at 30 °C in SD media with glucose, lacking uracil. To cultivate Ole1 cells, the media were supplemented with 0.02% linoleic acid (18:2 (n-6)) and 0.2% Tween-60. To induce expression a 0.5 ml overnight culture were transferred to 20 ml medium containing galactose and the appropriate supplements. Cells were cultured for 4 days at 30 °C, harvested by centrifugation at 3000 g for 10 min, frozen at −20 °C and lyophilized for 48 h.

Picocyanobacterial cultivation
Prochlorococcus strains were grown in a seawater-based medium Pro99 medium (Moore et al., 2007) based on Mediterranean seawater. Synechococcus strains were grown in an artificial seawater-based medium (Wyman et al., 1985) with modifications as described previously (Lindell et al., 1998). All strains were grown in 30 ml cultures at 21 °C under cool white light under a 14:10 h light–dark cycle, at a 10–15 μmol photon m−2 s−1. Synechococcus strains WH7803 and WH8102 and Prochlorococcus strains Med4, NATL2A, MIT9312 and MIT9313 were grown as axenic strains, whereas Synechococcus strains WH8109 and WH7805, Prochlorococcus strain Med4 and freshwater Synechococcus strain PCC7942 were non-axenic cultures. Three cultures were grown for every strain and analyzed separately, except Med4, for which we grew three axenic cultures and two non-axenic cultures. (The non-axenic cultures were used for identification of the gas chromatography/flame ionization detection (GC/FID) of fatty acid methyl esters (FAMEs), as they have all fatty acids identified, and the axenic cultures were for the fatty acid abundance analysis.) The bacteria were harvested at the beginning of the stationary phase by centrifugation at 6000 g for 15 min, and then again at 9000 g for 10 min. Pellets were flash frozen and stored at −80 °C until they were lyophilized for 24 h.

Lipid extraction and analysis
For analysis of esterified fatty acids in yeast, lyophilized cell pellets were submitted to transesterification using sodium methoxide (Hornung et al., 2002): Cells were homogenized in 0.5 ml 0.5 M sodium methoxide and 1.4 ml methanol by vortexing. After shaking for 1 h, FAMEs were extracted by adding 2 ml saturated sodium chloride and 4 ml hexane. The hexane phase was dried under streaming nitrogen and dissolved in 30 μl acetonitrile.

For analysis of fatty acid profiles from cyanobacteria, lyophilized bacteria cells were submitted to acidic hydrolysis (Miquel and Browse, 1992). One milliliter of a methanolic solution containing 2.75% (v v−1) sulfuric acid (95–97%) and 2% (v v−1) dimethoxypropan was added to the sample. The sample was incubated for 1 h at 80 °C. To extract the resulting FAME, 200 μl of saturated sodium chloride solution and 2 ml of hexane were added. The hexane phase was dried under streaming nitrogen and dissolved in 100 μl acetonitrile for GC analysis.

For determination of the position of double bonds in fatty acids, FAMEs were converted into their 4,4-dimethyloxazoline (DMOX) derivatives according to Christie (1998). Ninety microliters of FAME resulting from acidic hydrolysis was dried under streaming nitrogen, 200 μl 2-amino-2-methyl-1-propanol was added and the sample was incubated at 180 °C for at least 14 h. Fatty acid derivatives were extracted by adding 1 ml of dichloromethane to the sample, followed by 2.5 ml hexane and 1 ml water. The hexane phase was washed once with 1 ml water and then dried under streaming nitrogen. DMOX derivatives were separated from remaining FAME by thin layer chromatography, using petrol ether/diethyl ether (2:1, v v−1) as running solvent. DMOX
derivatives were extracted from the plate, dissolved in 10 μl acetonitrile and subjected to GC/mass spectrometry.

GC/FID analysis was performed with an Agilent 6890 gas chromatograph (Agilent Technologies, Waldbronn, Germany) fitted with a capillary DB-23 column (30 mx0.25 mm; 0.25 μm coating thickness; J&W Scientific, Agilent). Helium was used as carrier gas at a flow rate of 1 ml min⁻¹. The temperature gradient was 150 °C for 1 min, 150–200 °C at 8 K min⁻¹, 200–250 °C at 25 K min⁻¹ and 250 °C for 6 min. FAMEs were identified according to the retention time of the corresponding peaks in the external standard (Supelco 37 component FAME Mix; Sigma, Munich, Germany). GC/mass spectrometry analysis for DMOX derivatives was carried out using a ThermoFinnigan Polaris Q mass selective detector connected to ThermoFinnigan Trace gas chromatograph (Austin, TX, USA) equipped with a capillary DB-23 column. GC was performed using the same conditions as for GC/FID. Electron energy of 70 eV, an ion source temperature of 230 °C, and a temperature of 260 °C for the transfer line is used. See Supplementary Figure 7 for the DMOX derivatives analysis.

Conflict of Interest
The authors declare no conflict of interest.

Acknowledgements
We thank Debbie Lindell and Gazalah Sabehi for kindly providing cyanobacterial strains, Andrea Nickel and Sabine Freitag for expert technical assistance and Cornelia Herrfurth for her advice with fatty acid analysis. This work was supported by funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/under REA Grant Agreement No. 317184, a European Commission ERC Advanced Grant No. 321647, the Louis and Lyra Richmond Agreement No. 317184, a European Commission ERC Framework Programme FP7/2007-2013/under REA Grant (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/under REA Grant. This work was supported by funding from the People Programme (Marie Curie Actions) of the European Union’s Seventh Framework Programme FP7/2007-2013/under REA Grant Agreement No. 317184, a European Commission ERC Advanced Grant No. 321647, the Louis and Lyra Richmond Memorial Chair in Life Sciences (to OB) and the Edmond de Rothschild Caesarea Foundation (to SR).

Author contributions
SR and OB designed the project. SR conducted the molecular biology experiments, JF-U and IS performed bioinformatics and EH and IF performed lipidomics; SR and OB wrote the manuscript with contributions from all authors to data analysis, figure generation and the final manuscript.

References
Adir N, Zer H, Shochat S, Ohad I. (2003). Photoinhibition—a historical perspective. Photosynth Res 76: 343–370.
Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. (1990). Basic local alignment search tool. J Mol Biol 215: 403–410.
Becker RA, Wilkis AR, Brownrigg R, Minka TP, Deckmyn A. (2017). Maps: Draw Geographical Maps. R package.3.2.0. Available at: https://CRAN.R-project.org/package=maps.
Beja O, Fridman S, Glaser F. (2012). Viral clones from the GOS expedition with an unusual photosystem-I gene cassette organization. ISME J 6: 1617–1620.
Besemer J, Borodovsky M. (1999). Heuristic approach to deriving models for gene finding. Nucleic Acids Res 27: 3911–3920.
Biller SJ, Schubotz F, Rognensack SE, Thompson AW, Summons RE, Chisholm SW. (2014). Bacterial vesicles in marine ecosystems. Science 343: 183–186.
Bossie MA, Martin CE. (1989). Nutritional regulation of yeast delta-9 fatty acid desaturase activity. J Bacteriol 171: 6409–6413.
Bragg JC, Chisholm SW. (2008). Modeling the fitness consequences of a cyanophage-encoded photosynthesis gene. PLoS One 3: e3550.
Breitbart M, Thompson LR, Suttle CA, Sullivan MB. (2007). Exploring the vast diversity of marine viruses. Oceanography 20: 135–139.
Brum JR, Ignacio-Espinoza JC, Roux S, Doulcier G, Acinas SG, Alberti A et al. (2015). Ocean plankton. Patterns and ecological drivers of ocean viral communities. Science 348: 1261498.
Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K et al. (2009). BLAST plus: architecture and applications. BMC Bioinformatics 10: 421.
Chi X, Yang Q, Zhao F, Qin S, Yang Y, Shen J et al. (2008). Comparative analysis of fatty acid desaturases in cyanobacterial genomes. Comp Funct Genomics 2008: 284508.
Christie W. (1998). Gas chromatography-mass spectrometry methods for structural analysis of fatty acids. Lipids 33: 343–353.
Clokie MRJ, Shan JY, Bailey S, Jia Y, Krisch HM, West S et al. (2006). Transcription of a ‘photosynthetic’ T4-like phage during infection of a marine cyanobacterium. Environ Microbiol 8: 827–835.
Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36: W465–W469.
Enav H, Beja O, Mandel-Gutierrez Y. (2012). Cynophages tRNAs may have a role in cross-infectivity of oceanic Prochlorococcus and Synechococcus hosts. ISME J 6: 619–628.
Evans C, Pond DW, Wilson WH. (2009). Changes in Emiliania huxleyi fatty acid profiles during infection with E-huxleyi virus 86: physiological and ecological implications. Aquat Microb Ecol 55: 219–228.
Farrant GK, Dore H, Cornejo-Castillo FM, Partensky F, Ratin M, Ostrowski M et al. (2016). Delineating ecologically significant taxonomic units from global patterns of marine picocyanobacteria. Proc Natl Acad Sci USA 113: E3365–E3374.
Forterre P. (2013). The virocell concept and environmental microbiology. ISME J 7: 233–236.
Fridman S, Flores-Uribe J, Larom S, Alalouf O, Liran O, Yacoby I et al. (2017). A myovirus encoding both photosystem I and II proteins enhances cyclic electron flow in infected Prochlorococcus cells. Nat Microbiol;
Langmead B, Salzberg SL. (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9: 357–U354.
Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H et al. (2007). Clustal W and Clustal X version 2.0. Bioinformatics 23: 2947–2948.
Lee WM, Ishikawa M, Ahlquist P. (2001). Mutation of host delta9 fatty acid desaturase inhibits brome mosaic virus RNA replication between template recognition and RNA synthesis. J Virol 75: 2097–2106.
Lindell D, Jaffe JD, Coleman ML, Putschik ME, Axmann IM, Rector T et al. (2007). Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. Nature 449: 83–86.
Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW. (2005). Photosynthesis genes in marine viruses yield proteins during host infection. Nature 438: 86–89.
Lindell D, Padan E, Post AF. (1998). Regulation of ntcA expression and nitrite uptake in the marine Synechococcus sp. strain WH 7003. J Bacteriol 180: 1878–1886.
Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. (2004). Transfer of photosynthesis genes to and from Prochlorococcus viruses. Proc Natl Acad Sci USA 101: 11013–11018.
Lisle JT, Robbins LL. (2016). Viral lysis of photosynthesizing microbes as a mechanism for calcium carbonate nucleation in seawater. Front Microbiol 7: 1938.
Los DA, Ray MK, Murata N. (1997). Differences in the control of the temperature-dependent expression of four genes for desaturases in Synechocystis sp. PCC 6803. Mol Microbiol 25: 1167–1175.
Lowe TM, Eddy SR. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25: 955–964.
Malitsky S, Ziv C, Rosenwasser S, Zheng S, Schatz D, Porat Z et al. (2016). Viral infection of the marine alga Emiliania huxleyi triggers lipidome remodeling and induces the production of highly saturated triacylglycerol. New Phyto 210: 88–96.
Mann NH, Cook A, Millard A, Bailey S, Clokie M. (2003). Marine ecosystems: bacterial photosynthesis genes in a virus. Nature 424: 741.
Marsan D, Wommack KE, Ravel J, Chen F. (2014). Draft genome sequence of Synechococcus sp. strain CB0101, isolated from the Chesapeake Bay Estuary. Genome Announc 2: e01111–13.
McKinney W. (2010). Data structures for statistical computing in python. In: van der Walt S, Millman J (eds). Proceedings of the 9th Python in Science Conference. SciPy 2010: Austin, pp. 51–56.
Millard A, Clokie MRJ, Shub DA, Mann NH. (2004). Genetic organization of the psbAD region in phages infecting marine Synechococcus strains. Proc Natl Acad Sci USA 101: 11007–11012.
Millard AD, Zwingler K, Downey MJ, Mann NH, Scanlan DJ. (2009). Comparative genomics of marine cyanomyoviruses reveals host genes localized to a hyperplastic region: implications for mechanisms of cyanophage evolution. Environ Microbiol 11: 2370–2387.
Miquel M, Browse J. (1992). Arabidopsis mutants deficient in polysaturated fatty acid synthesis. J Biol Chem 267: 1502–1509.
Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, Lindell D et al. (2007). Cultivating the marine cyanobacterium Prochlorococcus. Limnol Oceanogr Methods 5: 353–362.
Nissimov Ji, Pagarete A, Ma F, Cody S, Dunigan DD, Kimmance SA et al. (2017). Coccolithoviruses: a review of cross-kingdom genomic thievery and metabolic thuggery. *Viruses* 9: 52.

Patro R, Duggal G, Love MI, Irazarry RA, Kingsford C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14: 417–419.

Peng Y, Leung HC, Yiu SM, Chin FY. (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics* 28: 1420–1428.

Philosof A, Yutin N, Flores-UrIBE J, Sharon I, Koonin EV, Battchikova N, Aro EM, Giglione C, Meinnel T, Glaser F, Haynes M, Glaser F, Sharon I, Alperovitch A, Rohwer F, Haynes M, Glaser F, Atamna-Ismaeel N et al. (2009). Photosystem I gene cassettes are present in marine virus genomes. *Nature* 461: 258–262.

Sharon I, Battachikova N, Aro EM, Giglione C, Meinnel T, Glaser F et al. (2011). Comparative metagenomics of microbial traits within oceanic viral communities. *ISME J* 5: 1178–1190.

Shankey S, Perktold J. (2010). Statsmodels: econometric and statistical modeling with python. In: van der Walt S, Millman J (eds). *Proceedings of the 9th Python in Science Conference*. SciPy 2010: Austin, pp 57–61.

Soneson C, Love MI, Robinson MD. (2015). Deseq2: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 31: 2069–2077.

Soutar JD, Mortimer KK. (1966). Unsaturated fatty acid mutants of *Saccharomyces cerevisiae*. *J Bacteriol* 92: 597–600.

Robinson MD, McCarthy DJ, Smyth GK. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139–149.

Roitman S, Flores-UrIBE J, Philosof A, Knowles B, Rohwer F, Ignacio-Espinoza JC et al. (2013). Closing the gaps on the viral photosystem-I psaDCAB gene organization. *Environ Microbiol* 17: 5100–5108.

Rosenwasser S, Mausz MA, Schatz D, Sheyn U, Malitsky S, Aharoni A et al. (2014). Rewiring host lipid metabolism by large viruses determines the fate of *Emiliania huxleyi*, a bloom-forming alga in the ocean. *Plant Cell* 26: 2689–2707.

Roux S, Brum JR, Dutilh BE, Sunagawa S, Duhaime MB, Loy A et al. (2016). Ecobiocenomes and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* 537: 689–693.

Rudis B, Bolker B, Schulz J. (2017). ggalt: Extra Coordinate Systems, ‘Geoms’, Statistical Transformations, Scales and Fonts for ‘ggplot2’.0.4.0. Available at: https://CRAN.R-project.org/package=ggalt.

Sakamoto T, Wada H, Nishida I, Ohmori M, Murata N. (1994a). Delta-9 acyl-lipid desaturases of cyanobacteria —molecular-cloning and substrate specificities in terms of fatty-acids, sn-positions, and polar head groups. *J Biol Chem* 269: 25576–25580.

Sakamoto T, Los DA, Higashi S, Wada H, Nishida I, Ohmori M et al. (1994b). Cloning of omega-3 desaturase from cyanobacteria and its use in altering the degree of membrane-lipid unsaturation. *Plant Mol Biol* 24: 249–263.

Sato N, Murata N. (1981). Studies on the temperature shift-induced desaturation of fatty-acids in monogalactosyl diacylglycerol in the blue-green-alga (Cyanobacterium) *Anabaena variabilis*. *Plant Cell Physiol* 22: 1043–1050.

Shanklin J, Cahoon EB. (1998). Desaturation and related modifications of fatty acids. *Annu Rev Plant Physiol Plant Mol Biol* 49: 611–641.

Sharon I, Morowitz MJ, Thomas BC, Costello EK, Rolman DA, Banfield JF. (2013). Time series community genomics analysis reveals rapid shifts in bacterial species, strains, and phage during infant gut colonization. *Genome Res* 23: 111–120.

Sharon I, Alperovitch A, Rohwer F, Haynes M, Glaser F, Atamna-Ismaeel N et al. (2009). Photosystem I gene cassettes are present in marine virus genomes. *Nature* 461: 258–262.

Sullivan MB, Coleman ML, Weigele P, Rohwer F, Chisholm SW. (2005). *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biol* 3: 790–806.

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

Sullivan MB, Lindell D, Lee JA, Thompson LR, Bielawski JP, Chisholm SW. (2006). Prevalence and evolution of core photosystem II genes in marine cyanobacterial viruses and their hosts. *PLoS Biol* 4: 1344–1357.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G et al. (2015). Ocean plankton. Structure and function of the global ocean microbiome. *Science* 348: 1261359.

Suttle CA. (2005). Viruses in the sea. *Nature* 437: 356–361.

Suttle CA, Chan AM. (1994). Dynamics and distribution of cyanophages and their effect on marine cyanobacterial viruses and their hosts. *PLoS Biol* 4: 1344–1357.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G et al. (2015). Ocean plankton. Structure and function of the global ocean microbiome. *Science* 348: 1261359.

Suttle CA. (2005). Viruses in the sea. *Nature* 437: 356–361.

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

Tasaka Y, Gombos Z, Nishiya Y, Mohanty P, Ohba T, Okhi K et al. (1996). Targeted mutagenesis of acyl-lipid desaturases in *Synechococcus*: Evidence for the important roles of polyunsaturated membrane lipids in growth, respiration and photosynthesis. *EMBO J* 15: 6416–6425.

Vardi A, Van Mooy BA, Fredricks HF, Popendorf KJ, Ososilinski JE, Haramaty L et al. (2009). Viral glyco-sphingolipids induce lytic infection and cell death in marine phytoplankton. *Science* 326: 861–865.

Wada H, Murata N. (1998). Membrane lipids in cyanobacteria. In: *Lipids in Photosynthesis: Structure, Function and Genetics* Paul-André S, Norio M (eds). Springer: Dordrecht, Netherlands, pp 65–81.

Wada H, Gombos Z, Murata N. (1990). Enhancement of chilling tolerance of a cyanobacterium by genetic
manipulation of fatty acid desaturation. *Nature* **347**: 200–203.

Wada H, Gombos Z, Sakamoto T, Murata N. (1992). Genetic manipulation of the extent of desaturation of fatty-acids in membrane-lipids in the cyanobacterium *Synechocystis* PCC6803. *Plant Cell Physiol* **33**: 535–540.

Wada H, Schmidt H, Heinz E, Murata N. (1993). *In vitro* ferredoxin-dependent desaturation of fatty acids in cyanobacterial thylakoid membranes. *J Bacteriol* **175**: 544–547.

Wakeham SG, Canuel EA. (1988). Organic geochemistry of particulate matter in the eastern Tropical North Pacific-Ocean—implications for particle dynamics. *J Mar Res* **46**: 183–213.

Waskom M, Botvinnik O, drewokane, Hobson P, David, Halchenko Y, Lee A. (2016). *Seaborn*: v0.7.1. Available at: http://doi.org/10.5281/zenodo.54844.

Waterbury JB, Valois FW. (1993). Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. *Appl Environ Microbiol* **59**: 3393–3399.

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

Whelan S, Goldman N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol* **18**: 691–699.

Wickham H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag: New York, NY, USA.

Wilhelm SW, Suttle CA. (1999). Viruses and nutrient cycles in the sea—viruses play critical roles in the structure and function of aquatic food webs. *Bioscience* **49**: 781–788.

Wyman M, Gregory RPF, Carr NG. (1985). Novel role for phycocerythrin in a marine cyanobacterium, *Synechococcus* Strain Dc2. *Science* **230**: 818–820.

Xiao W. (2006). *Yeast Protocols*. Humana Press: New York City, NY, USA.

Zeidner G, Bielawski JP, Shmoish M, Scanlan DJ, Sabehi G, Beja O. (2005). Potential photosynthesis gene recombination between *Prochlorococcus* and *Synechococcus* via viral intermediates. *Environ Microbiol* **7**: 1505–1513.

Zheng Q, Jiao N, Zhang R, Chen F, Suttle CA. (2013). Prevalence of *psbA*-containing cyanobacterial podo-viruses in the ocean. *Sci Rep* **3**: 3207.

Zhu W, Lomsadze A, Borodovsky M. (2010). *Ab initio* gene identification in metagenomic sequences. *Nucleic Acids Res* **38**: e132.

Ziv C, Malitsky S, Othman A, Ben-Dor S, Wei Y, Zheng S et al. (2016). Viral serine palmitoyltransferase induces metabolic switch in sphingolipid biosynthesis and is required for infection of a marine alga. *Proc Natl Acad Sci USA* **113**: E1907–E1916.