Intact polar diacylglycerol biomarker lipids isolated from suspended particulate organic matter accumulating in an ultraoligotrophic water column

Jenan J. Kharbush^{*a,1}, Andrew E. Allen^{a,b}, Ahmed Moustafa^{b,c}, Pieter C. Dorrestein^d, Lihini I Aluwihare^a

^a *Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Dr., La Jolla, CA 92093, USA*

^b *Dept. of Microbial and Environmental Genomics, J. Craig Venter Institute, 4120 Capricorn Lane, La Jolla, CA 92037, USA*

^c *Dept. of Biology and Biotechnology, American University in Cairo, New Cairo, Egypt*

^d *Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA*

^{*} Corresponding author. Tel.: (608) 843-8517.

E mail address: jenanjk@gmail (Jenan Kharbush).

¹ Present address: Dept. of Earth and Planetary Sciences, Harvard University, 20 Oxford St., Cambridge, MA 02138, USA

1 ABSTRACT

2 Intact polar diacylglycerols (IP-DAGs) are essential components of cell membranes. Because
3 they are structurally diverse and hypothesized to represent primarily living cells, they are
4 potential molecular markers for a recent contribution by microbial communities to various
5 carbon reservoirs. This study employed a novel molecular networking approach to investigate
6 the evolution of IP-DAG structural diversity with depth in an ultraoligotrophic environment
7 of the western South Pacific Ocean to test the hypothesis that particle transport to depth is
8 rapid enough to preserve the IP-DAG biomarker signature of the photosynthetic community.
9 IP-DAG profiles for several cultured cyanobacteria and photosynthetic picoeukaryotes were
10 used as templates for constructing molecular networks to compare and interpret IP-DAG
11 signatures of suspended particles isolated from a water column depth profile. Analysis of
12 corresponding genetic community composition data for the same field samples was used to
13 connect IP-DAG structures with their likely biological sources. Our data show that, although
14 most IP-DAG classes associated with photosynthetic organisms were not observed below the
15 euphotic zone, several other IP-DAG classes in deep samples might provide interesting targets
16 for future studies seeking to examine the in situ contribution of deep sea microbes to
17 suspended particulate organic matter (POM). Overall, the results represent the deepest water
18 column IP-DAG dataset to date and demonstrate the utility of molecular networking for
19 analyzing and visualizing complex environmental datasets.

20
21 **KEYWORDS:** Tonga Trench; South Pacific; nutrient limitation; molecular networking; fatty
22 acids; picophytoplankton; dark ocean; betaine lipids; microbial communities

1. Introduction

Fatty acids (FAs) have a long history as tracers of biological activity in aquatic environments. Their sources in the environment are, however, often ambiguous due to the production of common structures by multiple organisms, including different groups of algae (Lang et al., 2011). For this reason, the use of intact polar diacylglycerol lipids (IP-DAGs), the parent compounds of FAs, has been proposed as a means of achieving an additional layer of metabolic, and in some cases even taxonomic, specificity as compared with examination of total FAs alone (Sturt et al., 2004). The IP-DAG structure, defined by both the head group and associated FAs, can reveal not only the potential phylogenetic associations of source organisms (Schubotz et al., 2009) but also the conditions under which they were produced, such as nutrient limitation (Van Mooy et al., 2009). In addition, because the polar head group is thought to be hydrolyzed relatively quickly after cell death, particularly under oxygenated conditions (White et al. 1979; Harvey et al. 1986), IP-DAGs can be assumed to represent living or recently living biomass. This makes them potentially useful biological tracers of ecosystem processes, including rapid exchange between living and detrital pools of organic matter (OM).

Planktonic IP-DAGs have only recently gained prominence as biomarkers for studying marine biogeochemical processes. Three major groups are thought to be commonly produced by marine plankton in open ocean environments (Van Mooy et al., 2009; Van Mooy and Fredricks, 2010; Pendorf et al., 2011): (i) three types of phospholipids, phosphatidylglycerol (PG), phosphatidylethanolamine (PE) and phosphatidylcholine (PC), (ii) three types of glycolipids, monoglycosyl-diacylglycerol (MGDG), diglycosyldiacylglycerol (DGDG) and sulfoquinovosyl-diacylglycerol (SQDG) and (iii) three types of betaine lipids,

diacylglyceryl trimethylhomoserine (DGTS), diacylglyceryl hydroxymethyltrimethylalanine (DGTA) and diacylglyceryl carboxyhydroxymethylcholine (DGCC). A fourth glycolipid variation is by way of glucuronosyldiacylglycerol (GADG), recently observed in strains of the heterotrophic SAR11 clade (Carini et al., 2015) and in the Alphaproteobacterium *Erythrobacter* sp. NAP1 (Sebastián et al., 2015) when each was grown under P starved conditions. Structures are shown in Fig. 1.

Studies have also attempted to establish associations among particular head group structures and marine microbial groups, and have suggested that the IP-DAG composition varies among phytoplankton and heterotrophic bacteria in surface waters (Schubotz et al., 2009; Van Mooy and Fredricks, 2010; Brandsma et al., 2012). As a group, the glycolipids are commonly referred to as “chloroplast lipids” because they are frequently associated with the thylakoid membranes of photosynthetic organisms, including both cyanobacteria and eukaryotic phytoplankton (Wada and Murata, 1998; Sato, 2004;). Furthermore, a recent study using ¹³C-labeled substrates proposed that SQDG in particular appears to be exclusively produced by photosynthetic organisms in the surface ocean (Pependorf et al., 2011), but the sources of others, such as MGDG, are more difficult to define and may be produced by photosynthetic organisms and heterotrophic bacteria (Sebastián et al., 2015). Experiments with these ¹³C-labeled substrates also supported previous indications that the betaine lipids are likely produced mainly by eukaryotic phytoplankton (Kato et al., 1996; Pependorf et al., 2011), at least in marine surface waters, although some bacterial production has been observed in cultures of heterotrophic bacteria under P-limited conditions (Benning et al., 1995; Sebastián et al., 2015) and in anoxic regions of the water column (Schubotz et al., 2009). As for the phospholipids, PG and PE are more commonly attributed to heterotrophic

bacteria, but PG in particular is also found in cyanobacteria, as it has been shown to be an essential component of the thylakoid membrane (Sato et al., 2000). Similarly, PC likely can have both a bacterial source and a eukaryotic source (Kato et al., 1996; Pependorf et al., 2011), although it has not been found in cyanobacteria (Van Mooy et al., 2006).

Typically, IP-DAGs are separated using high performance liquid chromatography (HPLC) and detected and identified using electrospray ionization mass spectrometry (ESI-MS) techniques (e.g. Sturt et al., 2004). The polar lipid extracts of environmental samples can be extremely complex and IP-DAGs are often present at relatively low concentration. Thus, manually extracting and analyzing individual mass spectra to definitively assign IP-DAG structures from complex samples can be time consuming.

Molecular networking is an analytical strategy that can greatly facilitate this task. Molecular networks are visual representations of structurally similar compounds, constructed on the basis of the similarity in their tandem MS fragmentation patterns (Watrous et al., 2012; Yang et al., 2013). Similar structures therefore cluster together and can be assigned on the basis of their proximity to the spectra of standard compounds. Each of nine major IP-DAG classes commonly observed in marine samples (Fig. 1) each has a characteristic fragmentation pattern (Sturt et al., 2004; Van Mooy and Fredricks, 2010; Schubotz et al., 2009) that makes each of them an ideal candidate for analysis using this approach. In addition to aiding in identification, the resulting structural diversity, as well as the relative abundance of individual structures, can be visually compared between samples.

In the first application of molecular networking to marine particulate samples, we sought to (i) investigate the evolution of IP-DAG structural diversity in suspended particulate OM (POM) with depth in an oligotrophic marine water column, and (ii) demonstrate the

93 utility of molecular networking to rapidly identify, classify and visualize differences in IP-
94 DAG diversity in a complex dataset where individual compounds of interest are in very low
95 abundance. An additional motivation was to test the feasibility of utilizing IP-DAGs in
96 suspended POM as a tracer for the biomass of small photosynthetic organisms from the
97 surface through the water column into the deep ocean. As noted above we chose to focus on
98 IP-DAGs over other commonly employed biomarkers like FAs because of the greater
99 taxonomic specificity afforded by examining head group structures (Sturt et al., 2004), their
100 relative lability and because there was precedent for finding IP-DAGs in deeper waters (Close
101 et al., 2014). In addition, the head groups of IP-DAGs can provide information about
102 environmental conditions experienced by producing organisms, such as P limitation (Van
103 Mooy et al., 2006; Sebastián et al., 2015). To initially seed the molecular network with
104 relevant classes of IP-DAGs we used available laboratory cultures of phytoplankton expected
105 to be present in the environment of interest. The cultures further served as a tool to connect
106 the IP-DAG dataset to the major phytoplankton groups identified in the same field samples
107 from 16S, 18S and plastid 16S phylogenetic affiliations, in order to assign tentative sources to
108 some IP-DAGs found in suspended POM.

109 Our data show that SQDG, previously implicated as a biomarker for photosynthetic
110 organisms, does not persist to depth in the suspended POM reservoir. Other glycolipids,
111 which could have mixed heterotrophic and autotrophic sources, show a similar pattern.
112 However, some MGDG and betaine lipids were detected in deeper samples, which probably
113 identifies the overall shift to a bacterial dominated IP-DAG contribution to the suspended
114 POM reservoir in the dark ocean at this site. In contrast, the phospholipids PE and PG were
115 present throughout the water column, as would be expected if heterotrophic bacteria were the

major contributors of these lipids to the suspended particulate reservoir. Paired analysis of 16S and 18S based microbial diversity on the same particles confirmed the important photosynthetic role of *Prochlorococcus* spp. and the picoeukaryote *Pelagomonas* spp., potentially identifying the major contributors to SQDG and betaine lipids found in shallow POM samples.

2. Material and Methods

2.1. Sample collection

All environmental samples were collected on the Microbial Oceanography of the Tonga Trench (MOTT) cruise on the R/V Revelle in September 2012, from a water column station in the northern part of the Tonga Trench (16°38.5' S, 172°12.0'W, water depth 9100 m). Depth profiles of the water column were obtained via conductivity, temperature, and depth (CTD) measurements and dissolved O₂ concentration was measured at both stations with a Sea-Bird Electronics oxygen probe validated by shipboard Winkler titrations taken at 500 m depth intervals. Nutrient samples were collected by way of Niskin bottles and analyzed for NO₃⁻, PO₄³⁻ and NH₃ concentrations through the analytical facilities at UC Santa Barbara's Marine Science Institute (<http://msi.ucsb.edu/services/analytical-lab>). POM samples were collected on 0.7 µm glass fiber filters (GF/Fs) preceded by a 45 µm Nitex pre-filter, using situ filtration with two battery powered, large volume pumps (WTS-LV; McLane Research Laboratories, Falmouth, MA). GF/Fs were pre-combusted for 8 h at 450 °C and Nitex was acid-cleaned before use. Multiple samples were collected from several depths: the surface, deep chlorophyll maximum at 140 m, and the the local O₂ minimum at 400, 3000 and 5000 m. Sample volumes (Supplementary material Table A.1) were determined with both the pump

139 control software and a mechanical flow meter attached to the pump outlet, which generally
140 agreed within 10% of each other. Surface POM samples were collected with a diaphragm
141 pump fitted with a 0.7 μm GF/F filter but no pre-filter, and the volume filtered was
142 determined with a mechanical flow meter. All GF/F filters collected were folded in half to
143 protect the sample face, wrapped in pre-combusted Al foil and immediately frozen to -80 °C
144 until extraction.

145 2.2. *Organisms and culture conditions*

146 IP-DAG structural diversity was qualitatively analyzed in several cultures of both
147 eukaryotic and prokaryotic picophytoplankton, comprising three major groups:
148 Cyanobacteria, Chlorophytes and Pelagophytes (Table 1). Because Chlorophyta is a very
149 large algal division, we selected organisms from several classes within this division, including
150 Prasinophyceae (*Ostreococcus* and *Micromonas*), Trebouxiophyceae (*Picochlorum*) and
151 Chlorophyceae (*Picocystis*). *Ostreococcus* and *Micromonas* are reported to be coastal
152 organisms, but with distributions that can extend into oligotrophic water (Palenik et al., 2007;
153 Worden et al., 2009). The *Picochlorum* representative was isolated from a coastal marine
154 environment (Foflonker et al., 2014), while *Picocystis* appears to be associated with
155 hypersaline environments (Wang et al., 2014). We would not expect the latter two organisms
156 to be present in the ultraoligotrophic South Pacific water sampled. However, the goal of this
157 component of the study was to examine the IP-DAG diversity across available organisms to
158 establish whether unique signatures could be attributed to picoeukaryotes such as
159 *Pelagomonas*, which appears widespread in oligotrophic marine environments (Dupont et al.,
160 2014; Worden et al., 2012).

In addition, it must be stated that cultures were not grown specifically for this study. Samples relevant to the study were obtained from labs where the organisms are maintained on a regular basis, so growth conditions and media were not standardized across the various experiments. All cultures were grown under conditions considered to be nutrient-replete and were harvested just before stationary phase. Another important consideration is that current knowledge suggests that culture conditions and growth phase can alter the IP-DAG composition fingerprint. P availability in particular can affect the IP-DAG composition of both phytoplankton (Van Mooy et al., 2006, 2009) and heterotrophic bacteria (Sebastián et al., 2015). On the other hand, as a test we had the opportunity to analyze an Fe replete and Fe starved culture of *Pelagomonas*, and observed similar IP-DAG composition and structural diversity in each culture (Fig. A.2). Thus, clearly the effects of culture conditions on IP-DAG composition should be more thoroughly investigated in the future. For the purposes of this study, however, the resulting distributions of head groups and associated FAs in each culture are useful not only in estimating the range of IP-DAG structural diversity across different phytoplankton classes, but also in “training” the molecular network and establishing some context within which to interpret the lipid and genetic data from field samples.

A *Prochlorococcus marinus* (str. MIT 9215) biomass sample was obtained from M. Saito (Woods Hole Oceanographic Institution) and was grown in modified a Pro99 medium (Moore et al., 2007) without added Zn but with an additional 6 mM NaHCO₃ to delay onset of stationary phase due to CO₂ limitation. *Synechococcus* sp. CC9311 was obtained from B. Palenik (Scripps Institution of Oceanography) and was grown in f/4 medium (Guillard, 1975). *Ostreococcus marinus*, *Micromonas pusilla* sp. RCC299, *Picocystis* sp. SENEW1 and *Picochlorum* sp. SENEW3 were also obtained from B. Palenik but were grown in f/2 algal

medium (Guillard, 1975). *Pelagomonas* sp. CCMP1756 was obtained from T. Coale in the lab of A. Allen (J. Craig Venter Institute, San Diego) and was grown in Aquil synthetic seawater medium amended with 880 μM NO_3^- , 100 μM PO_4^{3-} , 100 μM EDTA and 100 nM Fe (Sunda WG, 2005). The *Picochlorum* culture was axenic; all other organisms comprised $\geq 95\%$ of the biomass in each culture.

2.3. Extraction of IP-DAGs

IP-DAGs were extracted from using sonication (Huguet et al., 2010). Briefly, either glass fiber filters (Whatman GF/F) or centrifuged culture samples were sonicated in an ultrasonic bath for 5 min, 3x each with MeOH, 1:1 (v/v) MeOH:dichloromethane (DCM) and lastly with DCM. Extracts were combined and vacuum-filtered through a pre-combusted Buchner funnel to remove any small pieces of filter; the solvent was removed using a rotary evaporator followed by a N_2 stream and the extracts re-dissolved in a small amount of DCM. Culture samples were also centrifuged at 4000 g (5 min) in between each sonication step to remove cell debris. A DCM/MilliQ water extraction was performed on the resulting total lipid extract (TLE) to remove residual salts and the organic phase was passed through a small pipette filled with anhydrous Na_2SO_4 to remove any remaining water. Each environmental sample was fractionated into an apolar and a polar fraction using a 5 ml pipette packed with pre-combusted activated silica and eluted sequentially with hexane:DCM (9:1; v/v) and DCM:MeOH (1:1; v/v). The polar fraction (or the total TLE in the case of culture samples) was dried down, weighed, and re-dissolved in DCM before further analysis.

2.4. HPLC-ESI-tandem MS (MS^n)

IP-DAGs were analyzed using HPLC-ESI- MS^n according to Sturt et al. (2004), with a few modifications. IP-DAGs were injected onto a LiChrospher diol column (250 mm x 4 mm,

5 μm , Phenomenex) at a concentration of ca. 1.3-4 $\mu\text{g}/\text{ul}$, to give a total on-column TLE weight of 20 μg . Analysis was performed at room temperature with a flow rate of 0.2 ml/min, using an Agilent 1200 HPLC system coupled to a Thermo-Finnigan LTQ-XL linear ion trap mass spectrometer equipped with an ESI source. MS parameters: capillary 200 $^{\circ}\text{C}$ and 41V; spray 5 kV; sheath gas 15; auxiliary gas 5; sweep gas 5 (arbitrary units). Remaining parameters were determined by manual tuning using a solution of 1,2-dipalmitoyl sn-glycero-3-phosphocholine directly infused into the ESI source. Data were acquired in three scan events: the first (MS^1) was a full scan from m/z 400- 2000 and the second (MS^2) used a dynamic exclusion procedure to isolate ions from the MS^1 scan for collision induced dissociation (CID) fragmentation. The final scan (MS^3) isolated the most abundant ion from the MS^2 scan for CID. The normalized collision energy for CID was at an intensity of 35 (arbitrary units). All samples were analyzed in separate runs in both positive and negative modes to obtain additional structural information, especially for those IP-DAGs that ionize more efficiently in one mode vs. the other.

Commercially available standards for some IP-DAG head groups were also analyzed to obtain retention times and fragmentation patterns for comparison with environmental samples (Table A.2). Linear standard curves across several orders of magnitude were constructed for the standards where possible. However, the structural differences in IP-DAG head groups result in significant variation in detector response factor (Sturt et al., 2004). Therefore, because of a lack of authentic standards for several of the major classes of IP-DAGs in our samples, we treat quantification of all classes as relative, normalizing to the total lipid weight injected on to the column.

2.5. *Molecular networking of IP-DAGs*

The Global Natural Products Social Molecular Networking online workflow (GNPS; <http://gnps.ucsd.edu>) was used to create molecular networks of related compounds based on the similarity of their tandem MS fragmentation patterns (Watrous et al., 2012; Wang et al., 2016). These networks are generated through computational alignment and comparison of MS² fragmentation patterns. Multiple identical MS² spectra (minimum of two) are combined to give a consensus spectrum of a particular lipid structure, which is represented by an individual node in the network. The consensus spectra are then compared with each other, and related structures are connected by edges to generate spectral families with similar fragmentation patterns and therefore structural similarities, resulting in a molecular map of structural diversity and “relatedness” among samples. These molecular maps are often quite large and contain many unidentifiable compounds so, for ease of visualization, we present only the spectral families containing IP-DAG structures (Fig. 1).

Published fragmentation patterns (Sturt et al., 2004; Schubotz et al., 2009; Van Mooy and Fredricks, 2010), as well as authentic standards when possible, were used to assign representative structures of each IP-DAG class. Several representative spectra from each class were then uploaded to the community-curated GNPS spectral library as bronze-level standard spectra; this allowed the rapid computational identification of IP-DAG structures with minimal manual verification needed. In some cases it was, however, not possible to determine the exact structure of certain odd numbered FAs. For example, we could not distinguish straight chain 15:0 and 17:0 from *iso* or *anteiso* branched counterparts. Similarly, a FA assigned as 19:1 may be a chain with one double bond, but is more likely a cyclopropyl-containing FA, as the latter are much more common. Although the network analysis is based on fragmentation pattern and therefore cannot distinguish between DGTS and DGTA,

comparing the relative retention time of our DGTS standard with that of the betaine lipids in the culture samples demonstrated the presence of DGTS and another distinct set of peaks with similar mass spectra. The latter peaks eluted several minutes later, so we tentatively assigned them to DGTA, based on previous reports indicating that DGTA is more polar than DGTS (Vogel et al., 1990) and would therefore be expected to elute after DGTS under our chromatographic conditions. In addition, the putative DGTA lipids tend to contain a greater abundance of longer chain FAs than DGTS, a trend previously noted by Dembitsky (1996). Nevertheless, to avoid any misassignment, particularly for environmental samples where the retention time separation was less obvious, we report both DGTS and DGTA as one IP-DAG structural group, the betaine lipids.

Specific parameters for the GNPS online workflow are as follows. First, the data were clustered with MS-Cluster using a parent mass tolerance of 1.5 Da and an MS² fragment ion tolerance of 0.5 Da to create consensus spectra. Consensus spectra that contained less than two spectra were discarded, and a network was created where edges were filtered to have a cosine score ≥ 0.7 and more than 5 matched peaks. Furthermore, edges between two nodes were retained only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra were then searched against the GNPS spectral libraries, which now include standard IP-DAG spectra, with all matches required to have a score ≥ 0.7 and at least three matched peaks. A lipid extract from a clean GF/F was used as a process blank to subtract any background signals arising from the solvent or extraction procedure. The network was then imported into Cytoscape version 3.2.0 (Shannon et al., 2003) for visualization. Individual parent masses are represented by circular nodes and the similarity between them as determined by comparison of their MS² fragmentation patterns (yielding a cosine score) is

represented by edges between nodes. All data files used in the generation of the network are publically available through the online GNPS analysis infrastructure, including the environmental samples (Accession number MSV000078592), *Synechococcus* (MSV000079430), *Prochlorococcus* (MSV000079431) and eukaryotic phytoplankton cultures (MSV000079432).

2.6. DNA extraction, 16S PCR (polymerase chain reaction) amplification and sequencing

DNA was extracted from a portion of each GF/F filter based on an existing phenol-CHCl₃ extraction protocol described by Kharbush et al. (2013) and Rusch et al. (2007). 16S rDNA PCR was carried out using the Accuprime High Fidelity PCR kit (Invitrogen) and primers were 357F (5' CCTACGGGAGGCAGCAG3') and 926R (5' CCGTCAATTCMTTTRAGT 3'). The reverse primer was complemented with 454 adaptors and sample-specific 10 bp barcodes. Samples were cleaned using Agencourt AMPure XP (Becker Coulter), quantified with the Picogreen assay (Molecular Probes), mixed in approximately equimolar amount, and sequenced using Roche/454 GS FLX titanium technology.

Environmental raw sequences were processed to remove primers and barcodes. SSU-ALIGN (Nawrocki, 2009) was used to identify rRNA according to the predefined SSU HMM models. Sequences < 200 nucleotides were discarded. Potential chimeric PCR amplifications were checked using UCHIME (Edgar et al., 2011) and removed. BLAST (Altschul et al., 1992) search was performed against SILVA (release 121; Quast et al., 2013) for taxonomic classification. The search was executed with a TimeLogic® DeCypher® system (Active Motif Inc., Carlsbad, CA) with *e* value threshold $\leq 1\text{E-}10$. Heat maps of the most abundant organisms were constructed in R, with sequence abundances normalized to the total number

of reads in each sample. Absolute counts for sequences obtained from each sample can be found in Fig. A.1.

3. Results

3.1. IP-DAG analysis of cultured picophytoplankton

A summary of IP-DAG classes and the number of associated structures in each culture is provided in Fig. 2 and a full list of all IP-DAGs in each culture can be found in Table A.3.

Analysis of the cyanobacteria *Prochlorococcus* and *Synechococcus* demonstrated that both produce the same major IP-DAG classes (PG, MGDG, DGDG and SQDG). *Prochlorococcus*, however, produced FAs such as 19:1, whereas *Synechococcus* appears to use only combinations of C₁₄, C₁₆ and C₁₈ FAs, either saturated or mono-unsaturated (Table A.3). As observed in previous studies (e.g. Van Mooy et al., 2006; Galloway and Winder, 2015), *Prochlorococcus* and *Synechococcus* cultures contained no polyunsaturated FAs (PUFAs).

The picoeukaryotic cultures produced a greater diversity of IP-DAG structures than the cyanobacteria, including the betaine lipids DGTS and DGTA. Consistent with the hypothesis that betaine lipids may replace PC in cell membranes of organisms that frequently contend with P limited conditions (Benning et al., 1995; Geiger et al., 1999), the organisms typically found in oligotrophic environments produced the most betaine lipid structures (*Ostreococcus*, *Micromonas*, and *Pelagomonas*) and also lacked PC. *Picocystis* and *Picochlorum*, on the other hand, produce fewer betaine lipid structures and many more PC structures, consistent with their distribution in coastal environments where PO₄³⁻ limitation is not as common.

Molecular networking analysis of the LC-MS² spectra of IP-DAGs from the cultures allowed visualization of the structural separation and/or overlap within classes of IP-DAGs. While the three glycolipid head groups (MGDG, DGDG, SQDG) were produced by all three major picophytoplankton groups examined (Cyanobacteria, Chlorophytes, and Pelagophytes; Fig. 3A, B), betaine lipids were produced only by eukaryotic picophytoplankton (Fig. 3C). The phospholipids (PE and PG, Fig. 3D, E) were present in all cultures. However, because the cultures here were not axenic and heterotrophic bacteria are known to produce PE and PG, we refrain from assigning these phospholipids to any particular organism. Unlike PG, PE in particular is not known to be present in chloroplast membranes. Similarly, GADG was detected in the *Prochlorococcus* culture but at much lower concentrations than other head group structures. It is therefore likely that both PE and GADG in the non-axenic *Prochlorococcus* culture were produced by contaminating heterotrophic bacteria.

Examination of the FA composition of glycolipids found in each culture showed a clear distinction between the cyanobacteria and eukaryotes. While all picophytoplankton cultures produced SQDG, MGDG and DGDG structures, the majority of eukaryotic IP-DAGs contained at least one and often two PUFA structures, such as 18:3, 18:4, 20:4 and 22:6, while the cyanobacterial IP-DAGs contained shorter chain, less saturated FAs (Fig. 3). The clearest example was the MGDG/DGDG family (Fig. 3A), where there was very little overlap between the IP-DAGs produced by cyanobacteria and those produced by picoeukaryotes. The SQDG family, on the other hand, contained more structures produced by both the cyanobacteria and picoeukaryotes (Fig. 3B), consistent with the reported association of SQDG with the thylakoid membrane in these organisms (Wada and Murata, 1998; Sato, 2004; Popenorf et al., 2011).

3.2. IP-DAG profiles of Tonga Trench POM

Water column features of the Tonga Trench site were characteristic of an oligotrophic open ocean environment, with low concentrations of NO_3^- and PO_4^{3-} in surface water, and a deep euphotic zone, resulting in a deep chlorophyll maximum at 140 m. A local O_2 minimum was also present at 400 m (Fig. 4A, B). The total lipid yield/l sampled decreased with depth (Fig. 4C), with very little material recovered from deeper samples.

Several major classes of IP-DAGs were found in water column POM samples, including PE, PG, MGDG, DGDG and SQDG, and the betaine lipids DGTS/DGTA. DGCC was not detected in any of the samples; neither was PC or GADG. Quantification of IP-DAG classes showed that all decreased with depth (Fig. 5). SQDG and DGDG disappeared entirely below 400 m.

Network analysis of the IP-DAG structures showed that the structural diversity of IP-DAGs in surface samples was distinct from that below the euphotic zone at 400 m and below (Fig. 6). With a few exceptions, IP-DAG structures containing PUFAs were confined to surface samples, consistent with a eukaryotic phytoplankton source. For example, the most common MGDG IP-DAG structures in surface samples contained combinations of 18:4 and 18:5 FAs (Fig. 6A), which were not found in deeper samples. In contrast, structures present throughout the water column contained less unsaturated FAs (e.g. MGDG 14:0/16:1, Fig. 6A). Similarly, most betaine structures were identified exclusively in surface samples, consistent with expected production by eukaryotic phytoplankton. Some betaine lipids were, however, present in both surface and deep samples, and still others were only found in samples below 400 m (e.g. 16:0/18:1 and 16:1/18:1, respectively, Fig. 6B). Like the pattern within the

MGDG head group class, these betaine IP-DAGs contained less unsaturated or odd numbered FA chains, distinguishing them from those containing PUFAs produced by eukaryotic phytoplankton in surface water. However, the network analysis did identify trace amounts of a few PUFA-containing betaine structures at depth (e.g. 16:0/20:5, Fig. 6B). The FA composition of the SQDG IP-DAG structures was limited to FAs with no or few sites of unsaturation, and only one structure was found anywhere below the euphotic zone (Fig. 6C). Finally, the phospholipids PG and PE were found throughout the water column, with several structures containing odd chain FAs, consistent with a bacterial source (Fig. 6D, E). A complete list of all structures identified is provided in Table A.4.

3.3. Bacterial and eukaryotic community composition in Tonga Trench POM

16S rDNA analysis demonstrated that the surface community was dominated by *Prochlorococcus*, although heterotrophic genera from the SAR11 clade are also present (Fig. 7A). Other major bacterial groups, such as Planctomycetes, the SAR 324 Deltaproteobacteria clade and Nitrospinae increased in abundance below the euphotic zone. Nitrospinae, a phylum of NO₂⁻ oxidizing bacteria known to prefer low O₂ conditions, were most abundant at 400 m, corresponding to the local O₂ minimum and potentially also to the depth of enhanced nitrification (NH₄⁺ concentration showed a minimum at this depth and macronutrient profiles were consistent with enhanced regeneration). The potentially chemolithoautotrophic SAR 324 Marine Group B clade was at maximum abundance at 5000 m, consistent with its known distribution in deep colder water in low latitude regions (Wright et al., 1997; Brown and Donachie, 2007; Sheik et al., 2014).

18S rDNA analysis showed that surface waters contained various groups of dinoflagellates, including several groups in the order Syndiniales, consisting of endosymbiotic and parasitic dinoflagellates (Fig. 7B). Radiolarians (Class Polycystinea) were also present in surface water but were more abundant in the deeper samples. Radiolarians in the order Spumellaria were most abundant at 5000 m, while those in the order Nassellaria were most abundant at the local O₂ minimum at 400 m. Finally, sequences matching various species of copepod were found mainly in the surface water, which is likely a result of the larger pore size pre-filter used during sample collection, but also could be from disaggregated fecal pellets that contributed to suspended POM.

Because the large eukaryotic organisms contain proportionately greater amounts of DNA relative to picoplankton, very few 18S sequences belonging to the latter could be detected. Therefore, the presence and diversity of eukaryotic picophytoplankton was examined through the 16S rDNA of plastids rather than 18S rDNA (Fig. 7C). As observed in other oligotrophic environments (Worden et al., 2012; Dupont et al., 2014;), *Pelagomonas* was most abundant in the deep chlorophyll maximum at 140 m. Also present were several types of nanoflagellate prymnesiophytes, (*Chrysochromulina* and *Imantonia* sp.), and dictyochophytes (*Mesopedinella* and *Dictyochophyte* sp. *RCC332*). Chlorophytes were, however, absent from all samples.

4. Discussion

The structural diversity of intact polar lipids, such as the IP-DAGs, and their potential in identifying taxonomic diversity, activity and nutrient status, is becoming more commonly examined in a variety of marine environments ranging from the surface ocean (e.g. Van Mooy

et al., 2006; Brandsma et al., 2012) to deep subsurface sediments (Lipp and Hinrichs, 2009). However, few studies of planktonic IP-DAGs have examined depth profiles of lipid diversity extending into the dark ocean (Close et al., 2014). The low concentration of suspended POM (Druffel et al., 1992) and associated lipids, such as FAs (e.g. Wakeham, 1995) makes a non-targeted study particularly challenging, because the complexity of the dataset may mask the true diversity of IP-DAGs that can be detected via manual identification of mass spectra. Here we employed a mass spectrum-based molecular networking tool to assign IP-DAG structures in the complex total ion chromatographs generated from polar lipid extracts of suspended POM in an effort to comprehensively describe IP-DAG diversity in an oligotrophic water column. The samples represent a single depth profile of suspended POM collected using an in situ pump from depths as great as 5000 m. Using laboratory cultures of relevant organisms, the molecular network was first seeded with example spectra of the major IP-DAG classes expected to be present in the marine water column, which allowed us to quickly detect and visualize the diversity of IP-DAGs in both the cultures and ultimately the field samples (Figs. 3 and 6). Furthermore, we combined DNA sequence information from the same suspended POM samples with culture-based IP-DAG taxonomy to tentatively assign the major phytoplanktonic sources for IP-DAGs detected in field samples. A summary of IP-DAG distribution and likely source organisms is shown in Table 2.

All picophytoplankton cultures examined here produced SQDG, MGDG and DGDG structures, with the majority of eukaryotic IP-DAGs containing at least one and often two PUFA structures. Although not unexpected from previous studies (e.g. Biller et al., 2014), this result validates a potential means of separating prokaryotic and eukaryotic contributions within IP-DAG classes produced by both types of picophytoplankton, such as the glycolipids

435 MGDG and DGDG. This is something the trained molecular network can do quickly and
436 efficiently within complex datasets. In the current application the network clearly
437 differentiated the environmental MGDG/DGDG classes based on FA composition, enabling
438 us to infer contributions from both eukaryotic and prokaryotic producers. Furthermore, the
439 likely source organisms could be inferred from the genetic data, which showed that the most
440 abundant picophytoplankton in surface waters were *Prochlorococcus* and *Pelagomonas*,
441 while the larger phytoplankton class was dominated by dinoflagellates of the order
442 Dinophyceae. Whether these dinoflagellates were acting as phototrophs, heterotrophs or both
443 in this environment is not known, but Dinophyceae have been observed to contain mostly
444 18:4 and 18:5 FAs associated with glycolipid IP-DAGs in their plastid membranes (Leblond
445 and Chapman, 2000; Gray et al., 2009). Thus it seems likely that a significant portion of the
446 PUFA-containing glycolipids (MGDG and DGDG, Fig. 6A) originated from these
447 dinoflagellates in addition to *Pelagomonas*, which also contained some 18:4 glycolipids in
448 culture (Fig. 3A).

449 In contrast, the SQDG class contains more overlap in FA content between
450 picoeukaryotic and cyanobacterial cultures. Still, the lack of any PUFA FAs in the SQDG
451 structures found in environmental samples, and the fact that *Prochlorococcus* has been
452 observed to synthesize primarily SQDG as its major membrane lipid (Van Mooy et al., 2006),
453 together suggest that *Prochlorococcus* is the major contributor to the SQDG lipid pool in the
454 Tonga Trench environment. Interestingly, although all of the eukaryotic picophytoplankton
455 cultures contained SQDG, only *Micromonas* and *Ostreococcus* contained possible homologs
456 of known genes from the SQDG biosynthesis pathway (*sqdB/SQD1* and/or *sqdX/SQD2* for
457 cyanobacteria/plants; Benning, 1998; Yu et al., 2002), suggesting there may be another

458 unidentified SQDG biosynthesis pathway present in certain eukaryotic phytoplankton. Other
459 examples of microalgae found to produce SQDG but apparently lacking the known
460 biosynthetic genes include the Chlorophytes *Tetraselmis chuii* (Li et al., 2008) and *Chlorella*
461 *kessleri* (Sato et al., 2008) and the Raphidophyte *Heterosigma carterae* (Keusgen et al.,
462 1997).

463 The method employed here to isolate suspended POM captured detrital particles
464 presumably originating ultimately in surface water as well as attached and free living
465 heterotrophic bacteria. The imprint of surface ocean communities and in situ communities on
466 suspended POM recovered from the dark ocean is supported by the genetic data, where
467 sequences from both *Prochlorococcus* and SAR324, for example, were recovered. However,
468 IP-DAGs isolated from dark ocean POM did not contain signatures that could be clearly tied
469 to phytoplankton producers. In particular, the IP-DAG structures considered primarily
470 associated with thylakoid membranes and therefore with phototrophic organisms (SQDG,
471 DGDG, and MGDG), nearly or completely disappear beyond the euphotic zone. The
472 persistence of a surface ocean signal in the DNA dataset but not the IP-DAG dataset is at least
473 in part due to the different analytical approaches utilized for genetic (PCR amplification) vs.
474 geochemical (MS) investigations. However, a shorter residence time for the polar head groups
475 on IP-DAGs as compared with DNA cannot be ruled out as contributing to this trend as well.
476 Isotope studies have suggested that suspended POM isolated with in situ pumps is delivered
477 to the water column from sinking particles with replacement times < 10 yr (Bacon and
478 Anderson, 1982; Druffel et al., 2003), whereas IP-DAGs are anticipated to have short
479 lifetimes, of the order of days to weeks, in oxygenated water (Harvey et al., 1986). The
480 absence of SQDG, DGDG, and to a large extent MGDG, from the deep suspended POM

reservoir implies that this reservoir had not interacted recently (weeks/days) with rapidly sinking particles. Episodic alteration of the deep suspended POC reservoir is certainly likely but could not be addressed with the current dataset. However, it was noted that the 3000 m sample had higher FACS cell counts (data not shown), TLE concentration (Fig. 5C), MGDG lipids (Fig. 6A), and higher relative numbers of *Prochlorococcus* 16S rDNA genes (Fig. 7A), than the sample above or below it, which could indicate a temporal offset between events that influenced suspended POM at 400 m vs. 3000 m.

The absence of GADG IP-DAGs is noteworthy, given that this lipid type was shown to be the most abundant class produced by some heterotrophic bacteria under P starvation (Carini et al., 2015; Sebastián et al., 2015), including members of the SAR11 clade, which are abundant in the Tonga Trench. This would support suggestions that this oligotrophic region of the Pacific Ocean is not severely P limited even in surface waters (Behrenfeld et al., 2006; Moutin et al., 2007), at least for heterotrophic organisms.

Unlike the glycolipids, numerous phospholipid structures were identified throughout the water column, even in the deepest samples, and likely originate from an in situ bacterial source. Numerous studies have documented that PE and PG are the most common components of heterotrophic bacterial membranes (e.g. Popendorf et al., 2011; Van Mooy et al., 2006; Wada and Murata, 1998) and the FA composition of the identified PE and PG structures (containing odd FA chains) also suggests bacterial production (Fig. 6E). Similarly, although the MGDG structures at 3000 m contain a FA composition similar to that of *Prochlorococcus* cultures, an in situ bacterial source is probable given previous reports of MGDG in heterotrophic bacteria in oligotrophic environments, such as the Sargasso Sea

(Popendorf et al., 2011), as well as in culture experiments (Carini et al., 2015; Sebastián et al., 2015).

On the other hand, the presence of betaine lipids at depth is somewhat surprising. Their production in surface water has been attributed mainly to eukaryotic phytoplankton (Kato et al., 1996; Popendorf et al., 2011; Van Mooy and Fredricks, 2010). Certain structures were present in both surface and deep samples (e.g. 14:0/18:4, 16:0/18:1, 18:1/22:0, Fig. 6B), but others were found only below the euphotic zone (e.g. 16:1/18:1, 19:1/22:0, Fig. 6B). As with the phospholipids, it is possible that the structures at depth represent in situ heterotrophic bacterial production rather than surface export. Retention time data suggest that most of what we observe in our environmental samples is consistent with DGTS. These lipids are not commonly observed in bacteria, but were recently reported to accumulate as the major ones under P starvation in some cultures of heterotrophic bacteria (Sebastián et al., 2015). In the dark ocean inorganic P starvation would not be expected, so the presence of betaine lipids at deep depths requires a different explanation. Furthermore, the known genetic potential for betaine lipid biosynthesis in bacteria is confined mainly to members of the Alphaproteobacteria (López-Lara et al., 2003; Sebastián et al., 2015) and a BLAST search for the two key enzymes in bacterial DGTS synthesis (BtaA and BtaB, Riekhof et al., 2005) in the available genomes of SAR11, Nitrospinae, and SAR324 - our most commonly recovered heterotrophic 16S sequences - did not yield any positive hits, despite their prevalence in the 16S community data at multiple depths.

A close examination of the FA composition of betaine lipid structures at each depth revealed that the majority present in deep samples contain common even FAs with a low degree of unsaturation, such as 16:1 and 18:1, with the exception of a few structures that

contain PUFAs (e.g. 16:0/20:5, 16:0/18:4, 14:0/18:4, Fig. 6B). The presence and structural diversity of betaine lipids throughout the water column as identified by the network is intriguing and begs for more information regarding possible sources of these compounds. Some piezophilic bacteria in the deep sea also produce PUFAs (DeLong and Yayanos, 1986; Fang et al., 2003) and could be the source of these lipids. Alternatively, the presence of PUFA-containing betaine lipids throughout the water column could result from selective preservation over other IP-DAG head group classes; however, no study has compared relative degradation rates of the various IP-DAG classes. A final mystery is the unusual FA, 22:0, which was only detected in the betaine lipid class here, and is of unknown biological source

5. Conclusions and future directions

This study examined a single depth profile of suspended POM collected from the Tonga Trench region of the western South Pacific Ocean, so the observations must be interpreted in this limited context. This is the first time, however, that this environment has been sampled in this manner. The analysis describes the deepest water column IP-DAG dataset to date, and the application of the molecular networking tool to identify IP-DAGs in this complex environmental dataset highlights the utility of such informatics tools for advancing environmental metabolomics.

In terms of food web dynamics it is interesting to consider the limited penetration of glycolipid structures in the water column. Are these lipids essentially consumed by the surface ocean microbial loop, whose activity is evident in the relative abundance of 18S sequences from mixotrophic dinoflagellates? Alternatively, is it the efficient degradation of these lipids as particles sink through the water column that limits their delivery to depth? In

our samples, the detection of *Prochlorococcus* and dinoflagellate 16S and 18S rDNA, respectively, suggests that nucleic acids persist at a concentration high enough to be effectively amplified as surface-derived OM is delivered to depth. Our analysis of IP-DAGs, on the other hand, relies on intrinsically detectable concentrations of lipids to be present in the suspended POM. However, these lipids potentially have the advantage of providing some time constraints on the export time of particles to certain depths. Perhaps in more productive regions, or in times of episodic high OM flux events, we could expect to see more OM escaping degradation and sinking below the euphotic zone at a rate fast enough to preserve an IP-DAG surface signature. For example, in areas of the ocean where seasonal phytoplankton blooms result in larger, faster sinking particles, more OM escapes the euphotic zone (Karl et al., 2011), with semi-intact cells containing chloroplasts even settling on the deep sea floor in some locations (Smith et al., 1996). In addition, recent work with single amplified genomes from deep trench environments showed that cyanobacterial genomes were present even in very deep samples, indicating that cyanobacterial cells can remain intact enough to be sorted and sequenced using this technique (R. Leon Zayas, personal communication).

Future studies could attempt to collect more material by increasing pumping time, or, alternatively, by decreasing filter pore size to 0.2 μm instead of 0.7 μm . In this way, more of the submicron community and “XPOM” (Close et al., 2014) might be isolated and potentially available for compound-specific isotope analysis (CSIA) studies. Unlike culture independent techniques such as metatranscriptomics and single cell genomics, lipid biomarker-CSIA approaches have the potential to provide quantitative information on the relative contributions of distinct members of the microbial community to deep ocean food webs. CSIA has been successfully applied to FAs from two general IP-DAG classes, glyco- and phospholipids, to

obtain valuable information on the contribution of the heterotrophic bacterial community to exported POM in the mesopelagic water column (Close et al., 2014, 2013). Applying CSIA to more specific lipid head group classes would allow greater specificity in tracing carbon flow in food webs; for example, to distinguish the contributions of photosynthetic picoeukaryotes from other members of the microbial community. The betaine lipids, in particular, which have been primarily linked to eukaryotic phytoplankton, showed a unique distribution in the water column at this site vs. other IP-DAGs derived primarily from phytoplankton, and so would provide interesting targets for future studies using stable isotope composition to better constrain their potential biological sources.

The results also raise questions about the relative degradation rates of various IP-DAG structures in different environments. Degradation studies of various IP-DAG classes to examine potential differences in both the abiotic and biotic rates at which particular structures are degraded are needed to provide insight regarding the likelihood of certain structures being delivered to depth. Such studies could also help establish the residence time of suspended particles at depth. For instance, while the lipid classes found mainly in the euphotic zone (SQDG, MGDG, DGDG) are primarily found in the plastid or thylakoid membranes of photosynthetic organisms, those found deeper in the water column are usually associated with cell membranes (PE, betaine lipids). Whether a difference in cellular localization could influence the selective preservation of these groups over the others has not been explored, nor has the possibility that glyco- and sulfolipids may be more susceptible to bacterial degradation. In the future, correctly interpreting the origins and fate of these biomarkers in the marine water column will require studies that specifically address these unanswered questions.

Acknowledgements

We thank B. Palenik (SIO), M. Saito (MIT-WHOI) and T. Coale (UCSD/JCVI) for contribution of cultures, and A. Rabines (JCVI) for assistance with DNA and sample processing. We thank A. Sessions, H.G. Close and an anonymous reviewer for comments that greatly helped improve the manuscript. The MOTT 2012 cruise to the Tonga Trench was funded by the University of California Ship-Funds Program.

Associate editor – M. Elvert

Figure and Table Captions

Table 1. Cultured organisms (N/A, not available).

Table 2. Summary of the dominant IP-DAG structures in surface and deep POM and likely source organisms, based on genetic data and IP-DAG compositions in phytoplankton cultures.

Fig. 1. Structures of IP-DAG head groups.

Fig. 2. Number of structures assigned in each picophytoplankton culture for each IPL head group class. Circles represent the presence of a particular IPL head group in each culture while numbers within circles represent the number of unique structures identified.

Fig. 3. Molecular networking of IPL structures in picophytoplankton cultures, separated into head group classes MGDG/DGDG (A), SQDG (B), Betaine (C), PG (D), and PE (E). Each

circular node corresponds to a unique MS² fragmentation pattern and therefore a unique IPL structure. Node sizes are scaled relative to the number of MS² spectra matching each IPL structure in order to highlight the most commonly observed structures within each IPL class, and these selected structures are labeled with the identified FA chains. Colored pie charts on each node indicate in which culture(s) the IPL structure was identified and in which relative proportion, according to the legend provided. For example, therefore, in **A** the most commonly observed MGDG species was (14:0/16:1), identified in both *Prochlorococcus* and *Synechococcus* but not in any of the other cultures. Note that, because this depiction is based on the number of MS² spectra collected the information is only semi-quantitative and does not necessarily reflect actual lipid concentrations. Also note that any 19:1 FAs are likely cyclopropyl-containing FAs rather than linear chains with one double bond, but this could not be determined from the mass spectra alone.

Fig. 4. Water column profiles of **(A)** temperature and oxygen **(B)** major nutrients and **(C)** TLE yield and chlorophyll concentrations with depth.

Fig. 5. Relative quantification of IPL classes with depth, normalized to the total µg of TLE on column. PE, PG, and betaine lipids are shown in **A**, while MGDG, DGDG, and SQDG are shown in **B**.

Fig. 6. Molecular networking of IPL structures in Tonga Trench samples, separated into head group classes MGDG/DGDG **(A)**, Betaine **(B)**, SQDG **(C)**, PG **(D)**, and PE **(E)**. Each circular node corresponds to a unique MS² fragmentation pattern and therefore a unique IPL structure.

Node sizes are scaled relative to the number of MS² spectra matching each IPL structure in order to highlight the most commonly observed structures within each IPL class, and these select structures are labeled with the identified FA chains. Colored pie charts on each node indicate at which depth(s) the IPL structure was identified and in which relative proportion, according to the legend provided. For example, therefore, in **A**, the most commonly observed MGDG structure was (18:4/18:5), found only in the surface and deep chlorophyll maximum samples. Note that, because this depiction is based on the number of MS² spectra collected, the information is only semi-quantitative and does not necessarily reflect actual lipid concentrations. Note also that FAs denoted 19:1 are likely cyclopropyl-containing FAs rather than linear chains with one double bond, but this could not be determined from the mass spectra alone. Similarly, 15:0 and 17:0 FAs may represent either linear chains or *iso/anteiso* branched chains.

Fig. 7. Community analysis of environmental samples from the Tonga Trench. Heat map shows abundances normalized to the number of reads in each sample for (A) 16S, (B) 18S, and (C) 16S plastid hits to the specified taxa.

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1992. Basic local alignment search tool. *Journal of Molecular Biology* 215, 403–410. doi:10.1016/S0022-2836(05)80360-2
- Bacon, M.P., Anderson, R.F., 1982. Distribution of thorium isotopes between dissolved and particulate forms in the deep sea. *Journal of Geophysical Research* 87, 2045–2056. doi:10.1029/JC087iC03p02045
- Behrenfeld, M.J., Worthington, K., Sherrell, R.M., Chavez, F.P., Strutton, P., McPhaden, M., Shea, D.M., 2006. Controls on tropical Pacific Ocean productivity revealed through nutrient stress diagnostics. *Nature* 442, 1025–1028. doi:10.1038/nature05083
- Benning, C., 1998. Biosynthesis and function of the sulfolipid sulfoquinovosyl diacylglycerol. *Annual Review of Plant Physiology and Plant Molecular Biology* 49, 53–75.

doi:10.1146/annurev.arplant.49.1.53

Benning, C., Huang, Z.-H., Gage, D.A., 1995. Accumulation of a novel glycolipid and a betaine lipid in cells of *Rhodobacter sphaeroides* grown under phosphate limitation. Archives of Biochemistry and Biophysics 317, 103–111.

Biller, S.J., Schubotz, F., Roggensack, S.E., Thompson, A.W., Summons, R.E., Chisholm, S.W., 2014. Bacterial vesicles in marine ecosystems. Science 343, 183–186.

Brandsma, J., Hopmans, E.C., Brussaard, C.P.D., Witte, H.J., Schouten, S., Sinninghe Damsté, J.S., 2012. Spatial distribution of intact polar lipids in North Sea surface waters: Relationship with environmental conditions and microbial community composition. Limnology and Oceanography 57, 959–973. doi:10.4319/lo.2012.57.4.0959

Brown, M. V., Donachie, S.P., 2007. Evidence for tropical endemism in the Deltaproteobacteria Marine Group B/SAR324 bacterioplankton clade. Aquatic Microbial Ecology 46, 107–115. doi:10.3354/ame046107

Carini, P., Van Mooy, B.A.S., Thrash, J.C., White, A., Zhao, Y., Campbell, E.O., Fredricks, H.F., Giovannoni, S.J., 2015. SAR11 lipid renovation in response to phosphate starvation. Proceedings of the National Academy of Sciences USA 112, 201505034. doi:10.1073/pnas.1505034112

Close, H.G., Shah, S.R., Ingalls, A.E., Diefendorf, A.F., Brodie, E.L., Hansman, R.L., Freeman, K.H., Aluwihare, L.I., Pearson, A., 2013. Export of submicron particulate organic matter to mesopelagic depth in an oligotrophic gyre. Proceedings of the National Academy of Sciences USA 110, 12565–70. doi:10.1073/pnas.1217514110

Close, H.G., Wakeham, S.G., Pearson, A., 2014. Lipid and ¹³C signatures of submicron and suspended particulate organic matter in the Eastern Tropical North Pacific: Implications for the contribution of Bacteria. Deep-Sea Research Part I: Oceanographic Research Papers 85, 15–34. doi:10.1016/j.dsr.2013.11.005

DeLong, E., Yayanos, A.A., 1986. Biochemical function and ecological significance of novel bacterial lipids in deep-sea prokaryotes. Applied and Environmental Microbiology 51, 730–737.

Dembitsky, V.M., 1996. Betaine ether-linked glycerolipids: Chemistry and biology. Progress in Lipid Research 35, 1–51. doi:10.1016/0163-7827(95)00009-7

Druffel, E.R.M., Bauer, J.E., Griffin, S., Hwang, J., 2003. Penetration of anthropogenic carbon into organic particles of the deep ocean. Geophysical Research Letters 30, 3–6. doi:10.1029/2003GL017423

Druffel, E.R.M., Williams, Peter, M., Bauer, J.E., Ertel, J.R., 1992. Cycling of dissolved and particulate organic matter in the open ocean. Journal of Geophysical Research 97, 15,639–15,659. doi:10.1029/2008GM000741/summary

Dupont, C.L., McCrow, J.P., Valas, R., Moustafa, A., Walworth, N., Goodenough, U., Roth, R., Hogle, S.L., Bai, J., Johnson, Z.I., Mann, E., Palenik, B., Barbeau, K.A., Craig Venter, J., Allen, A.E., 2014. Genomes and gene expression across light and productivity gradients in eastern subtropical Pacific microbial communities. The ISME Journal 9,

1076–1092. doi:10.1038/ismej.2014.198

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi:10.1093/bioinformatics/btr381

Fang, J., Chan, O., Kato, C., Sato, T., Peebles, T., Niggemeyer, K., 2003. Phospholipid FA of piezophilic bacteria from the deep sea. *Lipids* 38, 885–887. doi:10.1007/s11745-003-1140-7

Foflonker, F., Price, D.C., Qiu, H., Palenik, B., Wang, S., Bhattacharya, D., 2014. Genome of the halotolerant green alga *Picochlorum* sp. reveals strategies for thriving under fluctuating environmental conditions. *Environmental Microbiology* 17, 412–426. doi:10.1111/1462-2920.12541

Galloway, A.W.E., Winder, M., 2015. Partitioning the relative importance of phylogeny and environmental conditions on phytoplankton fatty acids. *Plos One* 10, e0130053. doi:10.1371/journal.pone.0130053

Geiger, O., Röhrs, V., Weissenmayer, B., Finan, T.M., Thomas-Oates, J.E., 1999. The regulator gene *phoB* mediates phosphate stress-controlled synthesis of the membrane lipid diacylglyceryl-N,N,N-trimethylhomoserine in *Rhizobium* (*Sinorhizobium*) *meliloti*. *Molecular Microbiology* 32, 63–73. doi:10.1046/j.1365-2958.1999.01325.x

Gray, C.G., Lasiter, A.D., Leblond, J.D., 2009. Mono- and digalactosyldiacylglycerol composition of dinoflagellates. III. Four cold-adapted, peridinin-containing taxa and the presence of trigalactosyldiacylglycerol as an additional glycolipid. *European Journal of Phycology* 44, 439–445. doi:10.1080/09670260902787977

Guillard, R.R., 1975. Culture of Phytoplankton for Feeding Marine Invertebrates In: *Culture of Marine Invertebrate Animals*. Springer US, pp. 29–60.

Harvey, H.R., Fallon, R.D., Patton, J.S., 1986. The effect of organic matter and oxygen on the degradation of bacterial membrane lipids in marine sediments. *Geochimica et Cosmochimica Acta* 50, 795–804. doi:10.1016/0016-7037(86)90355-8

Huguet, C., Martens-Habbena, W., Urakawa, H., Stahl, D.A., Ingalls, A.E., 2010. Comparison of extraction methods for quantitative analysis of core and intact polar glycerol dialkyl glycerol tetraethers (GDGTs) in environmental samples. *Limnology and Oceanography: Methods* 8, 127–145. doi:10.4319/lom.2010.8.127

Karl, D.M., Church, M.J., Dore, J.E., Letelier, R.M., Mahaffey, C., 2011. Predictable and efficient carbon sequestration in the North Pacific Ocean supported by symbiotic nitrogen fixation. *Proceedings of the National Academy of Sciences USA* 109, 1842–1849. doi:10.1073/pnas.1120312109/-/DCSupplemental. www.pnas.org/cgi/doi/10.1073/pnas.1120312109

Kato, M., Sakai, M., Adachi, K., Ikemoto, H., Sano, H., 1996. Distribution of betaine lipids in marine algae. *Phytochemistry* 42, 1341–1345.

Keusgen, M., Curtis, J.M., Thibault, P., Walter, J.A., Windust, A., Ayer, S.W., 1997. Sulfoquinovosyl diacylglycerols from the alga *Heterosigma carterae*. *Lipids* 32, 1101–

749 1112. doi:10.1007/s11745-997-0142-9

750 Kharbush, J.J., Ugalde, J.A., Hogle, S.L., Allen, E.E., Aluwihare, L.I., 2013. Composite
751 bacterial hopanoids and their microbial producers across oxygen gradients in the water
752 column of the California Current. *Applied and Environmental Microbiology* 79, 7491–
753 7501. doi:10.1128/AEM.02367-13

754 Lang, I., Hodac, L., Friedl, T., Feussner, I., 2011. Fatty acid profiles and their distribution
755 patterns in microalgae: a comprehensive analysis of more than 2000 strains from the
756 SAG culture collection. *BMC Plant Biology* 11, 124. doi:10.1186/1471-2229-11-124

757 Leblond, J.D., Chapman, P.J., 2000. Lipid class distribution of highly unsaturated long chain
758 fatty acids in marine diatoms. *Journal of Phycology* 1108, 1103–1108.
759 doi:10.1046/j.1529-8817.2000.00018.x

760 Li, H., Yan, X., Xu, J., Zhou, C., 2008. Precise identification of photosynthetic glycerolipids
761 in microalga *Tetraselmis chuii* by UPLC-ESI-Q-TOF-MS. *Science in China, Series C:*
762 *Life Sciences* 51, 1101–1107. doi:10.1007/s11427-008-0138-1

763 Lipp, J.S., Hinrichs, K.U., 2009. Structural diversity and fate of intact polar lipids in marine
764 sediments. *Geochimica et Cosmochimica Acta* 73, 6816–6833.
765 doi:10.1016/j.gca.2009.08.003

766 López-Lara, I.M., Sohlenkamp, C., Geiger, O., 2003. Membrane lipids in plant-associated
767 bacteria: their biosyntheses and possible functions. *Molecular plant-microbe interactions*
768 16, 567–579. doi:10.1094/MPMI.2003.16.7.567

769 Moore, L.R., Coe, A., Zinser, E.R., Saito, M. A., Sullivan, M.B., Lindell, D., Frois-Moniz, K.,
770 Waterbury, J., Chisholm, S.W., 2007. Culturing the marine cyanobacterium
771 *Prochlorococcus*. *Limnology and Oceanography: Methods* 5, 353–362.
772 doi:10.4319/lom.2007.5.353

773 Moutin, T., Karl, D.M., Duhamel, S., Rimmelin, P., Raimbault, P., Van Mooy, B.A.S.,
774 Claustre, H., 2007. Phosphate availability and the ultimate control of new nitrogen input
775 by nitrogen fixation in the tropical Pacific Ocean. *Biogeosciences Discussions* 4, 2407–
776 2440. doi:10.5194/bgd-4-2407-2007

777 Nawrocki, E.P., 2009. Structural RNA Homology Search and Alignment using Covariance
778 Models. Washington University in Saint Louis, School of Medicine.

779 Palenik, B., Grimwood, J., Aerts, A., Rouzé, P., Salamov, A., Putnam, N., Dupont, C.,
780 Jorgensen, R., Derelle, E., Rombauts, S., Zhou, K., Otiillar, R., Merchant, S.S., Podell, S.,
781 Gaasterland, T., Napoli, C., Gendler, K., Manuell, A., Tai, V., Vallon, O., Piganeau, G.,
782 Jancek, S., Heijde, M., Jabbari, K., Bowler, C., Lohr, M., Robbens, S., Werner, G.,
783 Dubchak, I., Pazour, G.J., Ren, Q., Paulsen, I., Delwiche, C., Schmutz, J., Rokhsar, D.,
784 Van de Peer, Y., Moreau, H., Grigoriev, I. V., 2007. The tiny eukaryote *Ostreococcus*
785 provides genomic insights into the paradox of plankton speciation. *Proceedings of the*
786 *National Academy of Sciences USA* 104, 7705–7710. doi:10.1073/pnas.0611046104

787 Popenorf, K.J., Lomas, M.W., Van Mooy, B.A.S., 2011. Microbial sources of intact polar
788 diacylglycerolipids in the Western North Atlantic Ocean. *Organic Geochemistry* 42,
789 803–811. doi:10.1016/j.orggeochem.2011.05.003

790 Popendorf, K.J., Tanaka, T., Pujo-Pay, M., Lagaria, A., Courties, C., Conan, P., Oriol, L.,
791 Sofen, L.E., Moutin, T., Van Mooy, B.A.S., 2011. Gradients in intact polar
792 diacylglycerolipids across the Mediterranean Sea are related to phosphate availability.
793 *Biogeosciences* 8, 3733–3745. doi:10.5194/bg-8-3733-2011

794 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner,
795 F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data
796 processing and web-based tools. *Nucleic Acids Research* 41, 590–596.
797 doi:10.1093/nar/gks1219

798 Riekhof, W.R., Andre, C., Benning, C., 2005. Two enzymes, BtaA and BtaB, are sufficient
799 for betaine lipid biosynthesis in bacteria. *Archives of Biochemistry and Biophysics* 441,
800 96–105. doi:10.1016/j.abb.2005.07.001

801 Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu,
802 D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-
803 Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E.,
804 Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.-H., Falcón, L.I., Souza, V.,
805 Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T.,
806 Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L.,
807 Nealson, K., Friedman, R., Frazier, M., Venter, J.C., 2007. The Sorcerer II Global Ocean
808 Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS biology*
809 5, e77. doi:10.1371/journal.pbio.0050077

810 Sato, N., 2004. Roles of the acidic lipids sulfoquinovosyl diacylglycerol and
811 phosphatidylglycerol in photosynthesis: Their specificity and evolution. *Journal of Plant*
812 *Research* 117, 495–505. doi:10.1007/s10265-004-0183-1

813 Sato, N., Hagio, M., Wada, H., Tsuzuki, M., 2000. Requirement of phosphatidylglycerol for
814 photosynthetic function in thylakoid membranes. *Proceedings of the National Academy*
815 *of Sciences USA* 97, 10655–10660. doi:10.1073/pnas.97.19.10655

816 Sato, N., Kamimura, R., Sugimoto, K., Tsuzuki, M., 2008. Difference in SQDG metabolism
817 between green algae and cyanobacteria under the sulfur-starved condition. In: Allen, J.F.,
818 Gantt, E., Golbeck, J.H., Osmond, B. (Eds.), *Photosynthesis. Energy from the Sun: 14th*
819 *International Congress on Photosynthesis*. Springer, pp. 795–798.

820 Schubotz, F., Wakeham, S.G., Lipp, J.S., Fredricks, H.F., Hinrichs, K.-U., 2009. Detection of
821 microbial biomass by intact polar membrane lipid analysis in the water column and
822 surface sediments of the Black Sea. *Environmental Microbiology* 11, 2720–34.
823 doi:10.1111/j.1462-2920.2009.01999.x

824 Sebastián, M., Smith, A.F., González, J.M., Fredricks, H.F., Van Mooy, B., Koblížek, M.,
825 Brandsma, J., Koster, G., Mestre, M., Mostajir, B., Pitta, P., Postle, A.D., Sánchez, P.,
826 Gasol, J.M., Scanlan, D.J., Chen, Y., 2015. Lipid remodelling is a widespread strategy in
827 marine heterotrophic bacteria upon phosphorus deficiency. *The ISME Journal* 1–11.
828 doi:10.1038/ismej.2015.172

829 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N.,
830 Schwikowski, B., Ideker, T., 2003. Cytoscape : A Software Environment for Integrated
831 Models of Biomolecular Interaction Networks. *Genome Research* 13, 2498–2504.

832 doi:10.1101/gr.1239303

833 Sheik, C.S., Jain, S., Dick, G.J., 2014. Metabolic flexibility of enigmatic SAR324 revealed
834 through metagenomics and metatranscriptomics. *Environmental Microbiology* 16, 304–
835 317. doi:10.1111/1462-2920.12165

836 Smith, C.R., Hoover, D.J., Doan, S.E., Pope, R.H., Demaster, D.J., Dobbs, F.C., Altabet,
837 M.A., 1996. Phytodetritus at the abyssal seafloor across 10° of latitude in the central
838 equatorial Pacific. *Deep Sea Research Part II: Topical Studies in Oceanography* 43,
839 1309–1338. doi:10.1016/0967-0645(96)00015-X

840 Sturt, H.F., Summons, R.E., Smith, K., Elvert, M., Hinrichs, K.-U., 2004. Intact polar
841 membrane lipids in prokaryotes and sediments deciphered by high-performance liquid
842 chromatography/electrospray ionization multistage mass spectrometry--new biomarkers
843 for biogeochemistry and microbial ecology. *Rapid Communications in Mass*
844 Spectrometry 18, 617–28. doi:10.1002/rcm.1378

845 Sunda WG, Price NM, M.F., 2005. Trace metal ion buffers and their use in culture studies. In:
846 Anderson, R. (Ed.), *Algal Culturing Techniques*. Elsevier Academic Press, Burlington,
847 MA, USA, pp. 35–63.

848 Van Mooy, B.A.S., Fredricks, H.F., 2010. Bacterial and eukaryotic intact polar lipids in the
849 eastern subtropical South Pacific: Water-column distribution, planktonic sources, and
850 fatty acid composition. *Geochimica et Cosmochimica Acta* 74, 6499–6516.
851 doi:10.1016/j.gca.2010.08.026

852 Van Mooy, B.A.S., Fredricks, H.F., Pedler, B.E., Dyhrman, S.T., Karl, D.M., Koblížek, M.,
853 Lomas, M.W., Mincer, T.J., Moore, L.R., Moutin, T., Rappé, M.S., Webb, E.A., 2009.
854 Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus
855 scarcity. *Nature* 458, 69–72. doi:10.1038/nature07659

856 Van Mooy, B.A.S., Rocap, G., Fredricks, H.F., Evans, C.T., Devol, A.H., 2006. Sulfolipids
857 dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine
858 environments. *Proceedings of the National Academy of Sciences USA* 103, 8607–12.
859 doi:10.1073/pnas.0600540103

860 Vogel, G., Woznica, M., Gfeller, H., Muller, C., Stumpfli, A.A., Jenny, T.A., Eichenberger,
861 W., 1990. 1(3),2-Diacylglycerol-3(1)-O-2'-(hydroxymethyl)(N,N,N-trimethyl)-beta-
862 alanine (DGTA): A novel betaine lipid from *Ochoromonas danica* (Chrysophyceae).
863 *Chemistry and Physics of Lipids* 52, 99–109. doi:10.1016/0009-3084(90)90154-J

864 Wada, H., Murata, N., 1998. Membrane lipids in cyanobacteria. In: Siegenthaler, P.-A.,
865 Murata, N. (Eds.), *Lipids in Photosynthesis: Structure, Function and Genetics*. Kluwer
866 Academic Publishers, Netherlands, pp. 65–81.

867 Wakeham, S.G., 1995. Lipid biomarkers for heterotrophic alteration of suspended particulate
868 organic matter in oxygenated and anoxic water columns of the ocean. *Deep-Sea*
869 *Research Part I* 42, 1749–1771. doi:10.1016/0967-0637(95)00074-G

870 Wang, M., Carver, J.J., Phelan, V. V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D.,
871 Watrous, J., Kapon, C.A., Luzzatto-Knaan, T., Porto, C., Bouslimani, A., Melnik, A. V.,
872 Meehan, M.J., Liu, W.-T., Crüsemann, M., Boudreau, P.D., Esquenazi, E., Sandoval-

873 Calderón, M., Kersten, R.D., Pace, L.A., Quinn, R.A., Duncan, K.R., Hsu, C.-C., Floros,
874 D.J., Gavilan, R.G., Kleigrew, K., Northen, T., Dutton, R.J., Parrot, D., Carlson, E.E.,
875 Aigle, B., Michelsen, C.F., Jelsbak, L., Sohlenkamp, C., Pevzner, P., Edlund, A.,
876 McLean, J., Piel, J., Murphy, B.T., Gerwick, L., Liaw, C.-C., Yang, Y.-L., Humpf, H.-
877 U., Maansson, M., Keyzers, R.A., Sims, A.C., Johnson, A.R., Sidebottom, A.M., Sedio,
878 B.E., Klitgaard, A., Larson, C.B., Boya, C., Torres-Mendoza, D., Gonzalez, D.J., Silva,
879 D.B., Marques, L.M., Demarque, D.P., Pociute, E., O'Neill, E.C., Briand, E., Helfrich,
880 E.J.N., Granatosky, E.A., Glukhov, E., Ryffel, F., Houson, H., Mohimani, H., Kharbush,
881 J.J., Zeng, Y., Vorholt, J.A., Kurita, K.L., Charusanti, P., McPhail, K.L., Nielsen, K.F.,
882 Vuong, L., Elfeki, M., Traxler, M.F., Engene, N., Koyama, N., Vining, O.B., Baric, R.,
883 Silva, R.R., Mascuch, S.J., Tomasi, S., Jenkins, S., Macherla, V., Hoffman, T., Agarwal,
884 V., Williams, P.G., Dai, J., Neupane, R., Gurr, J., Rodríguez, A.M.C., Lamsa, A., Zhang,
885 C., Dorrestein, K., Duggan, B.M., Almaliti, J., Allard, P.-M., Phapale, P., Nothias, L.-F.,
886 Alexandrov, T., Litaudon, M., Wolfender, J.-L., Kyle, J.E., Metz, T.O., Peryea, T.,
887 Nguyen, D.-T., VanLeer, D., Shinn, P., Jadhav, A., Müller, R., Waters, K.M., Shi, W.,
888 Liu, X., Zhang, L., Knight, R., Jensen, P.R., Palsson, B.O., Pogliano, K., Linington,
889 R.G., Gutiérrez, M., Lopes, N.P., Gerwick, W.H., Moore, B.S., Dorrestein, P.C.,
890 Bandeira, N., n.d. Sharing and community curation of mass spectra by GNPS. *Nature*
891 *biotechnology*.

892 Wang, S., Lambert, W., Giang, S., Goericke, R., Palenik, B., 2014. Microalgal assemblages in
893 a poikilohaline pond. *Journal of Phycology* 50, 303–309. doi:10.1111/jpy.12158

894 Watrous, J., Roach, P., Alexandrov, T., Heath, B.S., Yang, J.Y., Kersten, R.D., van der Voort,
895 M., Pogliano, K., Gross, H., Raaijmakers, J.M., Moore, B.S., Laskin, J., Bandeira, N.,
896 Dorrestein, P.C., 2012. Mass spectral molecular networking of living microbial colonies.
897 *Proceedings of the National Academy of Sciences USA* 109, E1743–52.
898 doi:10.1073/pnas.1203689109

899 White, D.C., Davis, W.M., Nickels, J.S., King, J.D., Bobbie, R.J., 1979. Determination of the
900 Sedimentary Microbial Biomass by Extractible Lipid Phosphate. *Oecologia* 40, 51–62.

901 Worden, A.Z., Janouskovec, J., McRose, D., Engman, A., Welsh, R.M., Malfatti, S., Tringe,
902 S.G., Keeling, P.J., 2012. Global distribution of a wild alga revealed by targeted
903 metagenomics. *Current Biology* 22, R675–R677. doi:10.1016/j.cub.2012.07.054

904 Worden, A.Z., Lee, J.-H., Mock, T., Rouzé, P., Simmons, M.P., Aerts, A.L., Allen, A.E.,
905 Cuvelier, M.L., Derelle, E., Everett, M. V, Foulon, E., Grimwood, J., Gundlach, H.,
906 Henrissat, B., Napoli, C., McDonald, S.M., Parker, M.S., Rombauts, S., Salamov, A.,
907 Von Dassow, P., Badger, J.H., Coutinho, P.M., Demir, E., Dubchak, I., Gentemann, C.,
908 Eikrem, W., Gready, J.E., John, U., Lanier, W., Lindquist, E.A., Lucas, S., Mayer,
909 K.F.X., Moreau, H., Not, F., Otillar, R., Panaud, O., Pangilinan, J., Paulsen, I., Piegu, B.,
910 Poliakov, A., Robbens, S., Schmutz, J., Toulza, E., Wyss, T., Zelensky, A., Zhou, K.,
911 Armbrust, E.V., Bhattacharya, D., Goodenough, U.W., Van de Peer, Y., Grigoriev, I. V.,
912 2009. Green evolution and dynamic adaptations revealed by genomes of the marine
913 picoeukaryotes *Micromonas*. *Science* 324, 268–272. doi:10.1126/science.1167222

914 Wright, T.D., Vergin, K.L., Boyd, P.W., Giovannoni, S.J., 1997. A novel delta-subdivision
915 proteobacterial lineage from the lower ocean surface layer. *Applied and Environmental*

916 Microbiology 63, 1441–1448.

917 Yang, J.Y., Sanchez, L.M., Rath, C.M., Liu, X., Boudreau, P.D., Bruns, N., Glukhov, E.,
918 Wodtke, A., De Felicio, R., Fenner, A., Wong, W.R., Linington, R.G., Zhang, L.,
919 Debonsi, H.M., Gerwick, W.H., Dorrestein, P.C., 2013. Molecular networking as a
920 dereplication strategy. *Journal of Natural Products* 76, 1686–1699.
921 doi:10.1021/np400413s

922 Yu, B., Xu, C., Benning, C., 2002. Arabidopsis disrupted in SQD2 encoding sulfolipid
923 synthase is impaired in phosphate-limited growth. *Proceedings of the National Academy*
924 *of Sciences USA* 99, 5732–5737. doi:10.1073/pnas.082696499

925

Figure 1

Glycolipids

SQDG
Sulfoquinovyldiacylglycerol

MGDG
Monoglycosyldiacylglycerol

GADG
Glucuronosyldiacylglycerol

DGDG
Diglycosyldiacylglycerol

Betaine lipids

DGTS
Diacylglyceryl trimethyl-homoserine

DGTA
Diacylglyceryl-hydroxymethyl-trimethylalanine

DGCC
Diacylglyceryl carboxyhydroxymethylcholine

Phospholipids

PG
Phosphatidylglycerol

PE
Phosphatidylethanolamine

PC
Phosphatidylcholine

Figure 2

	Synechococcus	Prochlorococcus	Ostreococcus	Micromonas	Picocystis	Picochlorum	Pelagomonas
PG	2	5	6	2	3	1	1
PE		8	1		3	1	11
PS			1			2	
PC					15	9	
SQDG	10	10	6	3	5	4	7
MGDG	6	9	4	9	6	1	5
DGDG	3	3	5	3	8	5	2
GADG		4					
Betaine			19	21	3	11	35

Figure 3

Figure 4

Figure 5

Figure 6

Figure 7

A. 16S

B. 18S

C. 16S plastid

Table 1

Genus	Phytoplankton group	Reported distribution	Reference
<i>Prochlorococcus marinus</i> str. MIT 9215	Cyanobacteria	Open ocean, oligotrophic	N/A
<i>Synechococcus</i> sp. CC9311	Cyanobacteria	Coastal ocean	(Palenik et al., 2006)
<i>Ostreococcus marinus</i>	Chlorophyte (Prasinophytes-Mamiellales)	Coastal, extending into oligotrophic	(Palenik et al., 2007)
<i>Micromonas pusilla</i> sp. RCC299	Chlorophyte (Prasinophytes-Mamiellales)	Coastal, extending into oligotrophic	(Worden et al., 2009)
<i>Picocystis</i> sp. SENEW1	Chlorophyte	Hypersaline	(Wang et al., 2014)
<i>Picochlorum</i> sp. SENEW3	Chlorophyte (Trebouxiophyceae)	Marine, halotolerant	(Foflonker et al., 2014)
<i>Pelagomonas</i> sp. CCMP1756	Pelagophyte	Open ocean, oligotrophic	(Dupont et al., 2014)

Table 2

Sample	Dominant lipids	FA composition	Likely sources (based on genetic data)
Surface (0 and 140 m)	MGDG/DGDG	PUFA	Dinoflagellates, <i>Pelagomonas</i>
		Non-PUFA	<i>Prochlorococcus</i>
	SQDG	Non-PUFA	<i>Prochlorococcus</i> , <i>Pelagomonas</i>
	Betaine	PUFA	<i>Pelagomonas</i> , other eukaryotic phytoplankton
Deep (400, 3000 and 5000 m)		Non-PUFA	Unknown, possibly heterotrophic bacteria
	PE/PG	Non-PUFA	Heterotrophic bacteria, <i>Prochlorococcus</i>
	MGDG	Non-PUFA	Heterotrophic bacteria
	Betaine	PUFA	Unknown, possibly piezophilic bacteria
		Non-PUFA	Heterotrophic bacteria
	PE/PG	Non-PUFA	Heterotrophic bacteria