



FEATURE ARTICLE

Bacterial and archaeal biogeography of the deep chlorophyll maximum in the South Pacific Gyre

Emily A. Walsh^{1,*}, David C. Smith¹, Mitchell L. Sogin², Steven D'Hondt¹

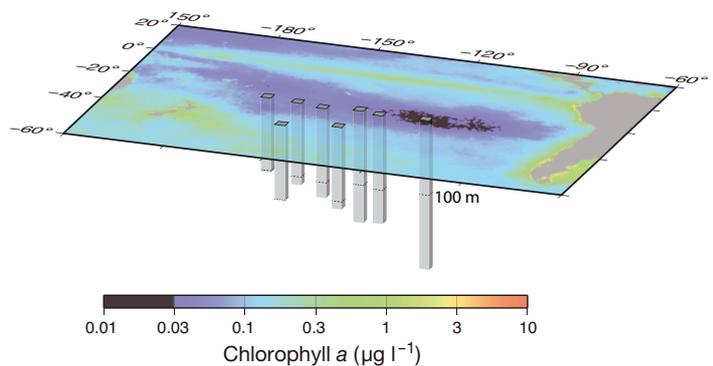
¹Graduate School of Oceanography, University of Rhode Island, Narragansett Bay Campus, 215 South Ferry Road, Narragansett, RI 02882, USA

²Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543, USA

ABSTRACT: We used 16S rRNA gene tag pyrosequencing to examine the biogeography of bacterial and archaeal community composition in the deep chlorophyll maximum (DCM) of the South Pacific Gyre (SPG), the largest and most oligotrophic region of the world ocean. Dominant DCM bacterial taxa, including *Prochlorococcus*, SAR11, SAR406, and SAR86, were present at each sampled site in similar proportions, although the sites are separated by thousands of kilometers and up to 100 m in water depth. Marine Group II (MGII) and MGIII *Euryarcheota* dominated the archaeal assemblages of the DCM at these sites. Bray-Curtis indices show that assemblage composition of these sites is >70% similar for *Bacteria* and >80% similar for *Archaea*. Despite these similarities, communities of the central SPG, the western SPG margin, and the southern SPG margin are distinguishable from each other. Comparison of our bacterial results to samples from the DCM of the North Pacific Gyre (NPG) and the relatively nutrient- and chlorophyll-rich Equatorial Pacific (EQP) shows that DCM bacterial assemblage composition is >50% similar throughout all 3 regions. Nonetheless, the SPG, NPG, and EQP assemblages are statistically distinct from each other (ANOSIM, $p = 0.001$), with the communities of the 2 gyres resembling each other more closely than either resembles the EQP community (which lives geographically between them). Variation in assemblage composition correlates with sea-surface chlorophyll concentration ($r^2 = 0.71$, $p < 0.003$). This study demonstrates that the DCM horizons of different oceanic regions harbor statistically distinct communities that are consistent within regions for thousands of kilometers.

KEY WORDS: South Pacific Gyre · Microbial biogeography · Microbial diversity · Deep chlorophyll maximum · Pyrosequencing

Resale or republication not permitted without written consent of the publisher



Prokaryotic communities of the deep chlorophyll maximum (DCM) of the South Pacific Gyre vary significantly with sea-surface chlorophyll (colours) and DCM depth (histograms).

Image: Robert Pockalny, University of Rhode Island

INTRODUCTION

Oceanic gyres constitute the largest nutrient-limited ecosystems on earth. They encompass 75% of the open ocean, with a total area of 51×10^6 km² worldwide (Polovina et al. 2008). The largest gyre, the South Pacific Gyre (SPG), lies far from the influence of terrestrial run-off and nutrient-rich upwelling zones (D'Hondt et al. 2009). A clear gradient in sea-surface chlorophyll *a* (chl *a*) that increases from a hyper-oligotrophic center to its oligotrophic edges marks the boundaries of the SPG (Fig. 1). The small size of prokaryotes has been inferred to provide a competitive advantage under these nutrient-limited conditions (Dufour et al. 1999).

The hyper-oligotrophic center of the SPG represents the most nutrient-limited region of the global ocean, and its deep chlorophyll maximum (DCM)

*Corresponding author: ewalsh@gso.uri.edu

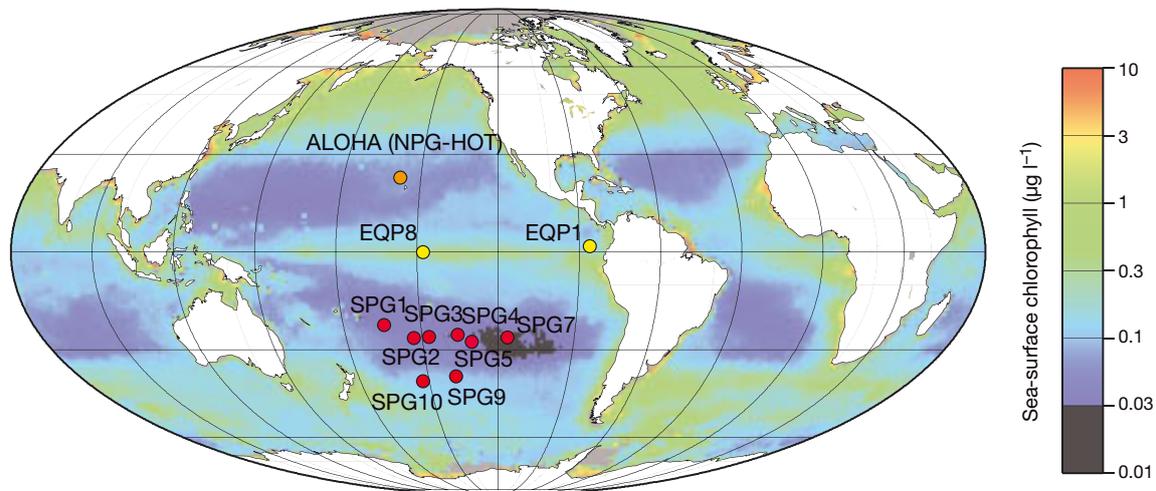


Fig. 1. Locations of sampling stations. Superimposed chlorophyll *a* content is the collected annual average (mean from September 1994 to December 2004) from SeaWiFS satellite data that were uploaded into GeoMapApp (Behrenfeld & Falkowski 1997, Gregg et al. 2005). South Pacific Gyre (SPG) sites are indicated by red circles. Equatorial Pacific (EQP) and North Pacific Gyre (NPG) stations are represented by yellow and orange circles, respectively

extends to greater depth than anywhere else on the planet (Morel et al. 2007). A permanent feature in both subtropical and tropical environments, DCMs account for a large fraction of net primary productivity in the open ocean (Weston et al. 2005, Huisman et al. 2006). However, little is known about the biogeography and diversity of their resident microbial communities, or how they respond to changes in environmental properties.

The idea that environmental characteristics of a habitat dictate the composition of a microbial assemblage prompts the often debated adage, 'everything is everywhere, but the environment selects' (Baas-Becking 1934). This theory assumes that microorganisms ubiquitously disperse and microbial distributions are solely explained by habitat preferences (Fenchel & Finlay 2004). In contrast, some recent studies hypothesize that local historical contingencies, such as evolutionary drift and dispersal limitations, cause geographic distance to strongly affect community structure in various marine and continental environments (Papke et al. 2003, Whitaker et al. 2003, Schauer et al. 2010, Martiny et al. 2011). In this study, we examined how microbial community composition of the DCM varies with environmental properties and geographic distance in the SPG, as well as between the SPG and other major Pacific regions.

To address this objective, we (1) investigated the continuity of microbial assemblages throughout the SPG, and between the SPG and other regions, and (2) tested how assemblage composition varies with environmental properties, including temperature,

dissolved nutrient concentrations, DCM depth, and sea-surface chl *a* concentration. We sampled the DCM at 8 stations along a 7500 km transect across the western, central, and southern portions of the gyre (Fig. 1). We used samples from the Equatorial Pacific and the North Pacific Gyre as out-groups to examine regional variability within the DCM.

To accomplish our goals, we used massively parallel sequencing (MPS) in marker gene analyses of short, rapidly evolving regions in ribosomal RNA (rRNA) genes. We selected domain-specific primers to target the V6 hypervariable region of the 16S rRNA gene. MPS analyses of these short regions can produce 60–65 bp pyrotags with sufficient taxonomic resolution to assess bacterial and archaeal community diversity (Sogin et al. 2006). These pyrotag data allow us to trace changes in microbial composition throughout Earth's largest oceanographic province.

MATERIALS AND METHODS

Field measurements

We collected our samples during expedition KNOX-02RR in southern-hemisphere summer (December 2006/January 2007) on the RV 'Roger Revelle.' The samples span a 7500 km transect through the western, central, and southern portions of the South Pacific Gyre (SPG) (Fig. 1). At each of the 8 sites, we cast a CTD-rosette system (Sea-Bird Electronics) and

fluorometer (Seapoint) and collected water using 10 l Niskin bottles. We used fluorescence to quantify chl *a* concentrations for all casts except at Stn SPG1. In total, 8 chl *a* samples, taken for DNA, were filtered through sterile polycarbonate membrane filters (0.22 μm pore size GTTP, Millipore) on board and stored at -80°C for shore-based analysis. Nutrient data were analyzed post-cruise using standard methods (Halm et al. 2012). To investigate the potential occurrence of phytoplankton blooms at times of sampling, we compared *in situ* chlorophyll measurements and 8 d time-averaged data to 1 mo time-averaged sea-surface chlorophyll data (<http://disc.sci.gsfc.nasa.gov/giovanni/overview/index.html>). We found no evidence of phytoplankton blooms (in the form of increased sea-surface chlorophyll) at any of our SPG stations at their times of sampling.

Pyrosequencing and community richness

We analyzed 9900 to 32 600 bacterial V6-tag sequences and 9200 to 35 500 archaeal V6-tag sequences per site according to procedures documented on the Visualization and Analysis of Microbial Population Structures (VAMPS) website (<http://vamps.mbl.edu>; Sogin et al. 2006, Huber et al. 2007, Huse et al. 2007, 2010). The total numbers of tags for all 8 sites are $\sim 178\,000$ bacterial V6-tag sequences and $\sim 168\,000$ archaeal V6-tag sequences. We used the Global Alignment for Sequence Taxonomy (GAST) system for tag sequence identification (Huse et al. 2008). To avoid bias introduced during the filtering and clustering process, we clustered all of the analyzed V6 sequencing data according to the single linkage pre-clustering (SLP) method (Huse et al. 2010). We clustered our sequencing reads into operational taxonomic units (OTUs) based upon the 97% similarity threshold.

We compared our SPG data to bacterial data from the DCM of 2 other distinct Pacific regions, the North Pacific Gyre (NPG) and the Equatorial Pacific upwelling zone (EQP). The NPG data are publicly available data from 3 samples of the Hawaiian Ocean Time-series (HOT, $22^{\circ}45' \text{N}$, $158^{\circ}00' \text{W}$), taken at 3 time points throughout 2006 and 2007 (October – HOT 186, February – HOT 189, and August – HOT 194; <http://hahana.soest.hawaii.edu/hot/>). These 3 samples provide a broad perspective on the influence of seasonality at this horizon in this region. We collected the EQP samples during RV 'Knorr' 195(III) expedition in northern hemisphere winter (January 2009); they are separated from each other by nearly

60° of longitude (Stn 1: $1^{\circ}48.21' \text{N}$, $86^{\circ}11.29' \text{W}$; Stn 8: $0^{\circ}00.36' \text{N}$, $147^{\circ}47.50' \text{W}$).

The SPG, NPG, and EQP samples were sequenced using the same reverse primers (1064r). The forward primers (967f) were also the same for the SPG and NPG samples. They differed for the EQP samples, but the EQP forward primers (518f) also targeted the V6 region (V4–V6) hypervariable region of the 16S rRNA gene and were analyzed on the same sequencer at the Marine Biological Laboratory (MBL, Woods Hole, MA, USA). To correct for the differences in read length, we trimmed the EQP sequences to match those of the NPG and SPG stations and re-entered them into the GAST system for sequence identification. Because we could not find DCM archaeal datasets for other regions, we did not include archaeal data in our comparison of SPG microbial communities to communities from other regions.

Accession numbers and data availability

The sequencing results and supporting data are publicly available on the VAMPS website (<http://vamps.mbl.edu>) under the following project ID: SPG (KCK_KNX_Bv6, KCK_KNX_Av6), EQP (KCK_EQP_Bv6_CUT), and NPG (KCK_HOT_Bv6). The VAMPS site also provides project metadata including information about sampling conditions and environmental factors.

Statistical analyses

We performed statistical analysis of all samples with the Primer-e and vegan software packages (Clarke & Gorley 2006, Oksanen et al. 2011), in order to examine relationships between environmental variables and microbial community composition. We performed canonical correspondence analysis (CCA), agglomerative hierarchical clustering algorithm (CLUSTER, group-average), similarity profiles (SIMPROF), analysis of similarity (ANOSIM), Spearman rank correlations (RELATE), similarity percentages (SIMPER), and an Envfit test on species data sets. We generated resemblance matrices using Bray-Curtis similarities for biological data. We subsampled each dataset down to the lowest number of reads prior to estimating values for richness (DIVERSE). We performed rarefaction analysis using the single linkage pre-clustering algorithm to cluster at 3% dissimilarity.

RESULTS

Environmental properties of the SPG, EQP, and NPG

The DCM varies from 100 to 200 m below the sea surface throughout the SPG, with its greatest depth in the heart of the gyre (Stn SPG7). The DCM is a measure of a maximum in chlorophyll concentration, although due to photoacclimation, it is not necessarily a maximum in biomass. The DCM is much shallower at the western edge (Stns SPG1 and SPG2) and southern edge of the gyre (Stns SPG9 and SPG10; Table 1). Despite being shallower at the western and southern edges, the DCM formed well below the depth of the pycnocline at every sampled SPG station (see Fig. S2 in the Supplement at www.int-res.com/articles/suppl/a075p001_supp.pdf). The temperature of the sampled horizon varied from 11°C at the southern gyre edge to 20°C along our northern sample transect. Salinity ranged from 34.3 to 35.6 psu, and *in situ* chl *a* concentration ranged from 0.22 to 0.40 mg m⁻³ (Halm et al. 2012). Average annual sea-surface chl *a* concentrations (mean SeaWiFS data from September 1994 to December 2004) at each site ranged from 0.029 to 0.117 mg m⁻³, with the lowest values in the central gyre (Stn SPG7) (Behrenfeld & Falkowski 1997, Gregg et al. 2005). This ultra-oligotrophic status agrees with a complementary study which revealed that concentrations of dissolved phosphate and nitrogen remained below detection limits from the sea surface to about 100 m below the sea surface at Stns SPG1 to SPG7 (Halm et al. 2012).

To consider the possibility that blooms affected our data, we examined satellite chlorophyll data for the 8 d and 1 mo windows that preceded our sampling date at each site (<http://disc.sci.gsfc.nasa.gov/giovanni/overview/index.html>). There is no evidence of a bloom preceding our sampling at any site (data not shown).

The sampling sites outside the SPG varied significantly in their sea-surface chlorophyll content, ranging from 0.87 to 1.18 mg m⁻³ in the EQP to 0.082 mg m⁻³ at Stn ALOHA in the NPG (Behrenfeld & Falkowski 1997, Gregg et al. 2005). While the DCM was relatively deep in the NPG (100 m), it was shallow in the eastern equatorial Pacific (41 m at Stn EQP1, 34 m at Stn EQP8) (Fig. S3). All DCM samples from the NPG and the sample from Stn EQP1 resided below the pycnocline, while the DCM of Stn EQP8 resided above (Fig. S2). Sea-surface temperature (SST) remained between 22.67 and 23.63°C at Stn ALOHA (<http://hahana.soest.hawaii.edu/hot/>; Karl & Lukas 1996) during the 2006–2007 timeframe. At the time of sample collection, SST was 17.45°C at Stn EQP1 and 24.27°C at Stn EQP8.

Community composition of the SPG

According to the GAST algorithm (Huse et al. 2008), the most abundant bacterial pyrotags at all SPG sites are assigned to the *Alphaproteobacteria*, *Gamma-proteobacteria*, *Bacteroidetes*, *Cyanobacteria*, and *Deferribacteres* (Fig. 2). A majority of the archaeal

Table 1. Environmental characteristics of the samples from the deep chlorophyll maximum of the South Pacific Gyre (SPG); na: data not available

Stn	Latitude (S) Longitude (W)	Sampling time (h)	Depth (m)	Phos- phate (µM)	Nitrate (µM)	Ammo- nium (µM)	Nitrite (µM)	Sigma theta (kg m ⁻³)	Salinity (psu)	Temper- ature (°C)	Conduc- tivity (S cm ⁻¹)	<i>In situ</i> chl <i>a</i> (µg l ⁻¹)	Sea sur- face chl <i>a</i> (µg l ⁻¹)
SPG1	23° 51.0' 165° 38.6'	20:25	100	0	0.08	0.04	0.07	25.76	35.61	19.64	4.83	na	0.05
SPG2	26° 03.1' 156° 53.6'	20:44	110	0.02	0.25	0.03	0.06	26.4	35.89	18.16	4.68	0.39	0.05
SPG3	27° 56.5' 148° 35.4'	16:58	120	0.06	0.21	0.03	0.05	26.04	35.51	18.62	4.72	0.29	0.04
SPG4	26° 28.9' 137° 56.4'	19:29	150	0.07	0.25	0.03	0.04	25.97	35.51	19.39	na	0.21	0.03
SPG5	28° 26.8' 131° 23.4'	06:19	145	0.08	0.34	b/d	0.1	26.19	35.39	18.1	4.65	0.34	0.04
SPG7	27° 44.5' 117° 37.2'	02:27	200	0.03	0.14	0.06	0.07	26.32	35.49	18.82	4.74	0.23	0.03
SPG9	39° 18.6' 139° 48.0'	21:04	110	0.93	5.39	0.17	0.46	26.67	34.35	10.8	3.83	0.36	0.11
SPG10	41° 51.1' 153° 06.3'	06:25	100	0.93	4.24	0.12	0.51	26.5	34.47	10.5	3.81	0.34	0.12

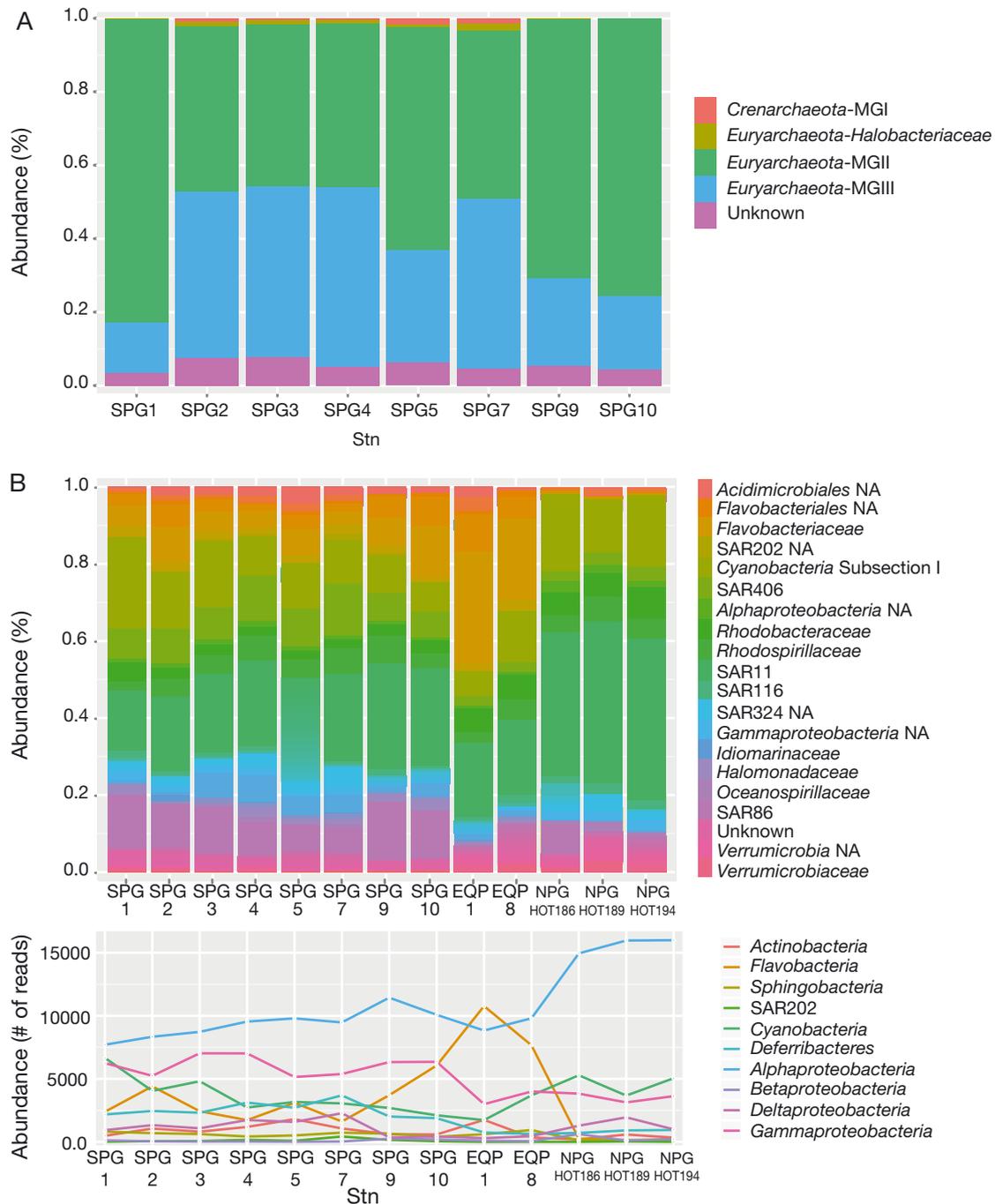


Fig. 2. (A) Archaeal and (B) bacterial community composition profiles. Relative abundance of dominant bacterial taxa (*Alphaproteobacteria*, *Gammaproteobacteria*, *Bacteroidetes*, and *Cyanobacteria*) and archaeal taxa (Marine Group II [MGII] and Marine Group III [MGIII] *Euryarchaeota*) as determined using the Global Assignment of Sequence Taxonomy (GAST) system for tag identification. SPG: South Pacific Gyre; EQP: Equatorial Pacific; NPG: North Pacific Gyre (Stn ALOHA); Stn ALOHA was sampled at 3 timepoints: October 2006 (HOT 186), February 2007 (HOT 189) and August 2007 (HOT 194). NA: no data for this taxonomic rank

sequences (95–99%) at each site showed best matches to Marine Group II (MGII) (~64%) and MGIII (~34%) *Euryarchaeota*. *Alphaproteobacteria* comprised the most abundant bacterial group at each station and were relatively most abundant in the central and

southern portions of the gyre. *Gammaproteobacteria* composed the second most abundant bacterial group at each station, except at Stn SPG1, which contained the highest proportion of *Cyanobacteria* in the gyre. While a high percentage of cyanobacterial tags

occurred at every station, their relative abundance diminished along the path of the SPG transect (from western gyre edge to gyre center to southern gyre edge). In contrast, the proportion of *Firmicutes* did not substantially change in the western and central portions of the gyre, but increased in the southern region of the gyre (Stns SPG9 and SPG10). Other dominant groups such as *Bacteroidetes*, *Deltaproteobacteria*, and *Actinobacteria* did not show variation in their relative abundance from station to station (Fig. 2).

The most common bacterial tag at each station matched *Pelagibacter ubique* of the SAR11 clade, which comprised 57% of the total identified alphaproteobacterial tags and 16.3% of all identified bacterial tags. Tags from the SAR11 clade collectively comprised 76.5% of the alphaproteobacterial tags and 21.9% of all the identified bacterial tags in all of our samples. Tags attributed to *Rhodobacteraceae* and *Rhodospirillaceae* compose the remaining *Alphaproteobacteria*, comprising 2.4 and 4.2% of the bacterial tags, respectively. The most abundant representatives of the gammaproteobacterial group included SAR86 (9.9%), *Alteromonas* (2.4%), *Halomonas* (1.4%), and an unknown *Gammaproteobacteria* species (1.4%). *Bacteroidetes* tags constituted 12% of the total identified tags in our SPG samples; they included unknown *Flavobacteriaceae* (6%), *Owenweeksia* (2.1%), *Marinoscillum* (1.9%), and unknown *Cryomorphaceae* (1.3%). Cyanobacterial tags (11.9%) were all identified as *Prochlorococcus*. The *Defferibacter* tags (9.2%) all aligned to the SAR406 clade.

The archaeal tags display far less taxonomic diversity. MGII dominates the archaeal community, comprising 64% of the archaeal tags. MGIII constitutes 34% of the total archaeal tags. MGI and *Halobacteriaceae* each comprised less than 1% of the total archaeal tags.

Diversity within the SPG DCM

Abundance-weighted (Bray-Curtis) similarity indices indicate that the composition of relatively abundant taxa in the DCM remains strikingly similar throughout the SPG. In contrast, the Jaccard index (presence/absence) differs greatly from sample to sample, indicating that rare taxa differ greatly from sample to sample (Boyle et al. 1990; Fig. 3). Although bacterial richness is high, the long-tailed rank abundance curve for each sample, along with the Jaccard results, suggests that very low-abundance taxa account for most of the bacterial richness (Fig. S1).

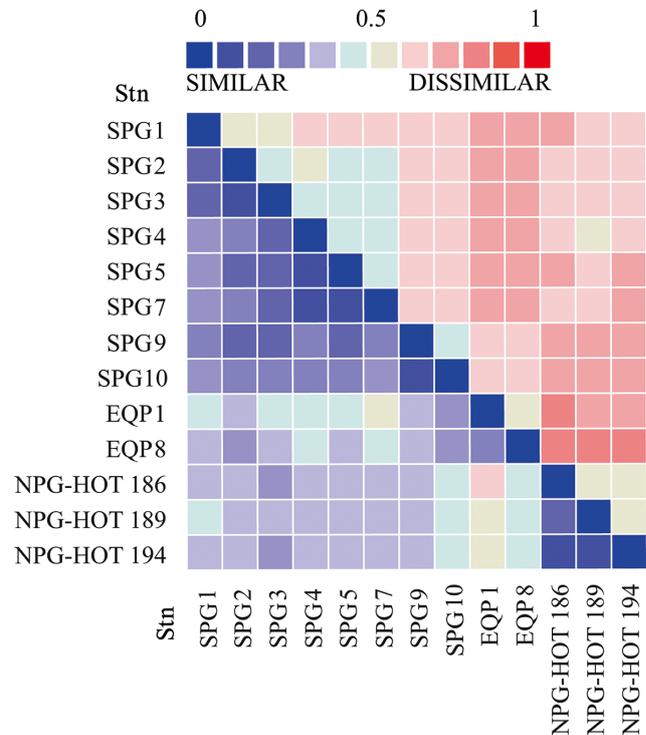


Fig. 3. Split heat map featuring the Bray-Curtis Index (lower left triangle) and Jaccard Index (upper right triangle) as calculated for the pairwise similarity of species-level bacterial taxonomic assignments between the Eastern and Central Equatorial Pacific (EQP), North Pacific Gyre (NPG), and South Pacific Gyre (SPG) stations. A score of 0 (blue) indicates that the communities are identical, and a score of 1 (red) indicates that they are completely dissimilar

Taxon-independent analyses using SLP clustering reveal that the richness of bacterial OTUs ($\leq 97\%$ similarity) exceeds richness of archaeal OTUs at all 8 sites. Rarefaction analysis of the 3% clusters reveals that over the range of sequences generated for each site (9200 to 35 500), bacterial richness increases with increasing number of sequences 2 to 4 times faster than archaeal richness (Fig. S1). Despite the large number of amplicons analyzed, deeper sequencing would be required to fully capture the bacterial richness of each sample. In contrast, archaeal communities appear to plateau at about 500 OTUs, suggesting that total archaeal richness is relatively well represented (Fig. S1).

Variation in DCM community composition within the SPG

We combined our SPG molecular results with environmental data to test the effects of distance and environmental properties on community composition

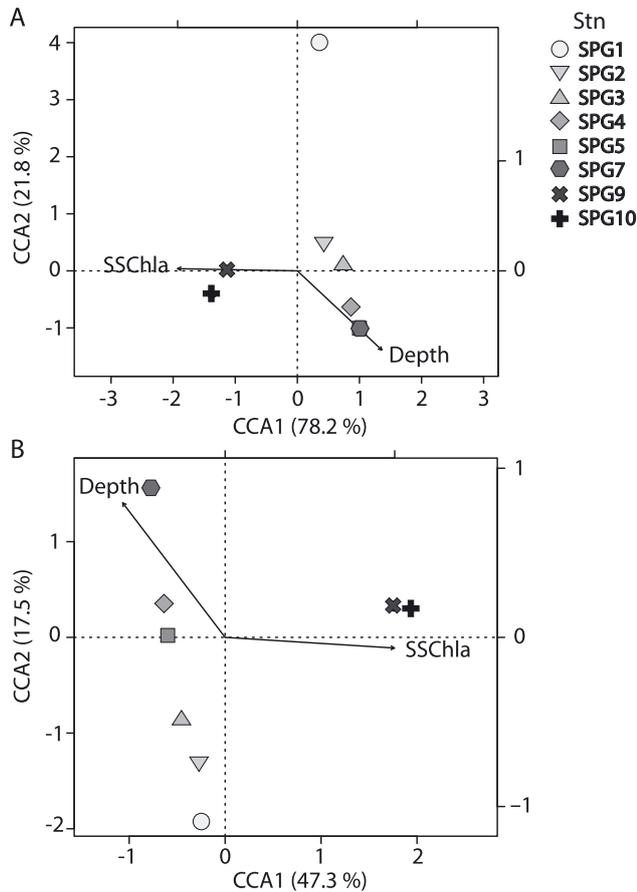


Fig. 4. Canonical correspondence analysis (CCA) plots of the South Pacific Gyre (SPG) stations generated from (A) archaeal and (B) bacterial operational taxonomic units (OTUs, 97% similarity) and environmental data matrices (permutations = 499) using the ‘vegan’ package in R. A triplot of the significant ($p < 0.05$) environmental factors is represented by the black arrows

and diversity. Bray-Curtis similarity indices, based on OTUs (97%), indicate that, from station to station, there was a strikingly similar bacterial community (>70% similarity) within the SPG. We examined the influence of geography by using CCA, SIMPROF, and agglomerative hierarchical clustering analysis (Figs. 4 & 5). The clustering analysis demonstrates a clear spatial gradient in microbial community composition, as western Stns SPG1–SPG3, central Stns SPG4, SPG5 and SPG7, and southern Stns SPG9 and SPG10 define 3 separate groupings. We performed a SIMPROF test to determine whether the samples within each grouping are statistically indistinguishable within a 95% confidence interval (Fig. 5). These groupings correspond to those that result from the CCA, as the bacterial communities of the deepest stations (Stns SPG4 and SPG7) more closely resemble each other than Stn SPG5, which resides geo-

graphically between them (Fig. 4B). Archaeal samples also remained true to the CCA, with the stations of the central gyre clustering together; this likely results from their similar proportions of MGII and MGIII *Euryarchaeota* (Figs. 2 & 4A).

Comparison to DCM community composition in other regions

To test whether microbial communities in the DCM of the SPG are distinct from those of other regions, we used bacterial tag populations from the DCM of the NPG and EQP as out-groups. Our analysis of similarity (ANOSIM) analysis ($p = 1$, significance = 0.1%) shows that samples from the SPG partition

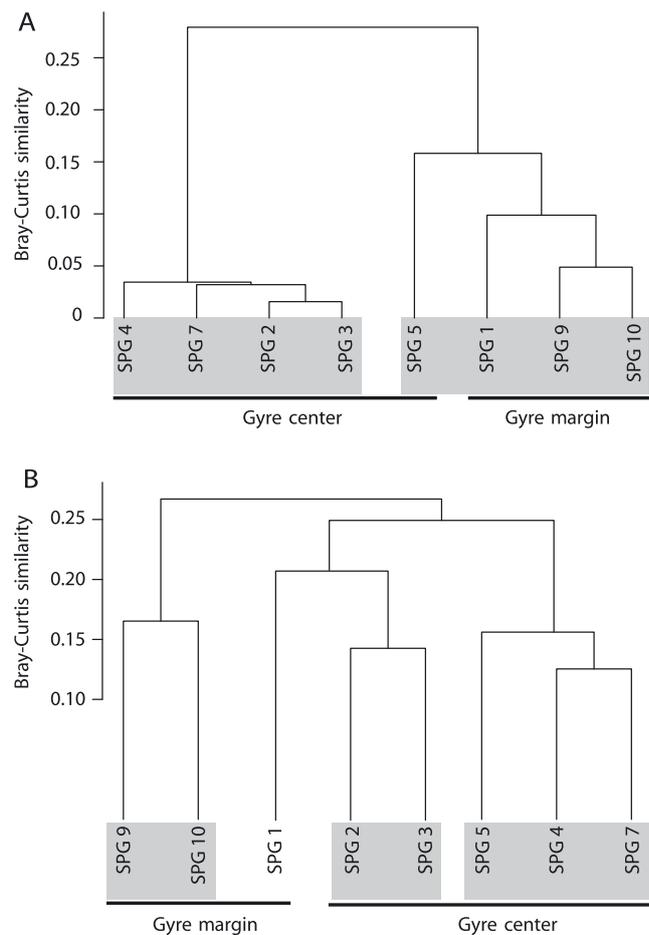


Fig. 5. Dendrograms generated from an agglomerative ‘group average’ cluster analysis of (A) archaeal and (B) bacterial community Bray-Curtis similarity (0 = 100% similar), based on species-level taxonomic assignments at the deep chlorophyll maximum of each station. Statistically identical samples ($p < 0.05$) were identified by a SIMPROF test and are highlighted by the grey boxes

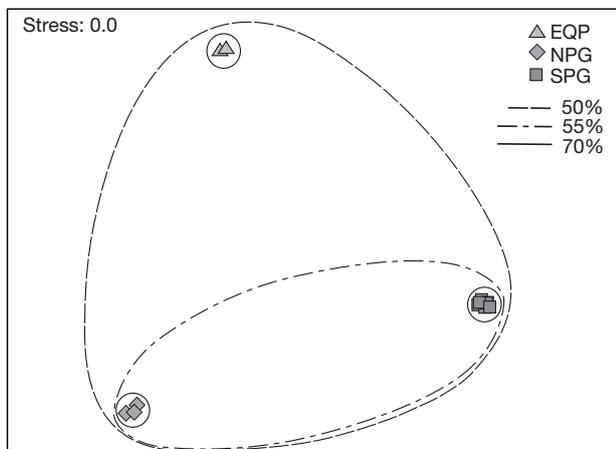


Fig. 6. Non-metric multidimensional scaling (stress: 0.0) performed on species-level taxonomic assignments and the corresponding data matrix subjected to an analysis of similarity (permutations = 999, distance = 'bray'). Samples taken from each environment clustered accordingly and remained distinct from each other with a p-value of 0.001 and $R = 1$. Dashed and solid lines denote Bray-Curtis sample similarity for the South Pacific Gyre (SPG), Equatorial Pacific (EQP), and North Pacific Gyre (NPG)

from those of the EQP and NPG despite striking similarities in their proportion and presence of identical dominant pyrotags (Fig. 6). A SIMPER analysis, which can be used to identify species that contribute most to the dissimilarity between groups, demonstrated the abundance of SAR11 to be the discriminating factor between the gyres (30% of dissimilarity) and *Flavobacteriaceae* to be influential in distinguishing the equatorial upwelling sites (15% of dissimilarity).

This distinct SPG bacterial assemblage is broadly similar throughout the central SPG despite the great distances between sites, grouping at >70% similarity at the 3% OTU level. The samples from all 3 regions (SPG, NPG, and EQP) group together at the 50% similarity level, indicating that much of the DCM bacterial community is relatively consistent regardless of region. However, at higher levels of similarity, the SPG, NPG, and EQP samples group independently from each other, indicating that these different regions host distinct DCM bacterial communities.

Environmental and geographic drivers of community structure

To evaluate whether community composition correlated with environmental variables in the SPG DCM, we performed a Spearman rank correlation

test (RELATE) that compared a matrix of each measured environmental variable with a matrix of calculated Bray-Curtis dissimilarities scores. Bacterial community composition, within the SPG, strongly correlates with sea-surface chlorophyll content ($\rho = 0.627$, significance = 0.1%), temperature ($\rho = 0.516$, significance = 0.3%), salinity ($\rho = 0.407$, significance = 3.1%), dissolved nitrate ($\rho = 0.368$, significance = 3.1%), and dissolved phosphate ($\rho = 0.511$, significance = 0.5%). Archaeal community composition also correlated with several environmental variables, including dissolved nitrite ($\rho = 0.457$, significance = 1.6%) and dissolved nitrate ($\rho = 0.374$, significance = 2%; Halm et al. 2012). There is no significant Spearman rank correlation between community compositions and either water depth or *in situ* chl *a*.

To expand on these results, we complemented these individual correlation tests with a CCA (Fig. 4). This methodology allowed us to quantify the percent variance in community composition explained by each environmental factor. The CCA revealed (1) the best predictor of both bacterial and archaeal community composition to be sea-surface chlorophyll content ($p = 0.020$, $p = 0.014$, respectively) explaining 33.1 and 36.14% of the total species (97%) variation, respectively, and (2) water depth ($p = 0.032$, $p = 0.004$), which explains 25.4 and 35.1%, respectively, across the SPG.

Because sea-surface chlorophyll content is the best individual predictor of bacterial community composition within the DCM of the SPG (Spearman, significance = 0.1%; CCA, $p = 0.020$), we used an environmental fitness analysis (EnvFit) to determine whether the relationship persists on an ocean-wide scale. To accomplish this goal, we compared bacterial community composition from the SPG, EQP, and NPG sites to annually averaged SeaWiFS sea-surface chlorophyll data. The resulting correlation is significant ($r^2 = 0.71$, $p < 0.003$).

Community composition and geographic distance

Recent investigations have challenged the once-held belief that geography does not constrain the dispersal of microorganisms (Papke et al. 2003, Whitaker et al. 2003, Schauer et al. 2010). In order to examine the influence of geographic distance on community composition regionally (within the SPG) and globally (including the EQP and NPG), we converted latitude and longitude into a pairwise distance matrix and ran a Spearman rank correlation test (permutations = 999). Both tests, regional and global,

indicate that geographic distance significantly correlates to bacterial community composition in both the SPG ($\rho = 0.529$, significance 0.1%) as well as between all stations analyzed ($\rho = 0.755$, significance 0.1%). Although these correlations to geographic distance are significant, their meaning remains unclear because the most distant sampling locations also differ in significant environmental properties.

DISCUSSION

Taxonomic richness in the DCM

Our discovery of high OTU richness for DCM bacterial assemblages in the hyper-oligotrophic SPG supports similar observations of microbial richness in other nutrient-limited marine environments. For example, high taxonomic richness of bacteria occurs at the DCM in both the Sargasso Sea and the NW Mediterranean Sea (Treusch et al. 2009, Pommier et al. 2010). Rarefaction analysis corroborates this claim by showing total bacterial taxa to be under-sampled at every SPG site. In contrast, archaeal richness is much lower than bacterial richness in our SPG samples, as in other previously studied marine environments (Aller & Kemp 2008, Winter et al. 2008).

Within the SPG, bacterial taxonomic richness was lowest at the gyre margin (Stns SPG1, SPG9, and SPG10) and highest in the gyre center (Stns SPG5 and SPG7), where the DCM extends to greater depths than anywhere else on the planet (Fig. S1). We know from satellite sea-surface chlorophyll data, a proxy for primary productivity, that this region has extremely low primary productivity. As a result, this region has a remarkably low flux of organic material to the seafloor (D'Hondt et al. 2009). This decline in energy availability from gyre margin to gyre center may therefore play a role in driving competition and diversification at these sites (Hibbing et al. 2010). It is also possible that DCM depth is responsible for the increased diversity at the gyre center, as it has been previously speculated that diversity may increase with depth due to the increasingly stable environment relative to the dynamic surface ocean (Ghioglione et al. 2012). The geographic variation in richness that we observe in the SPG is unlikely to be an artifact of short-term variation, such as phytoplankton blooms; no satellite evidence of variation in sea-surface productivity occurs at any of our stations in the week prior to the sampling date of each station. We observed no geographic trend in archaeal richness within the gyre.

In a study of the Sargasso Sea, Treusch et al. (2009) reported a shift from lower bacterial richness in surface water to higher richness at the DCM. Their study showed that the Sargasso Sea DCM contains a unique community, underscoring the importance of examining richness and community composition in the DCM (Treusch et al. 2009). Higher richness at the DCM may also explain why richness in our SPG DCM samples is high relative to richness estimated from 2 surface ocean depths (25 and 100 m) at NPG Stn ALOHA (Hunt et al. 2013). To our knowledge, the richness we observed at the SPG DCM is significantly higher than in previous studies of the DCM in other regions (Treusch et al. 2009, Pommier et al. 2010, Friedline et al. 2012). Some of these differences in richness can be attributed to methodological differences between studies; for example, T-RFLP studies, such as those of Treusch et al. (2009), generally sample only relatively abundant taxa, whereas tag-sequencing studies, such as those of Pommier et al. (2010), Friedline et al. (2012), and our study, sample both rare and abundant taxa.

Community composition in the DCM

Despite the remarkable taxonomic richness of bacterial assemblages in the SPG DCM, abundance-weighted community composition is strikingly homogeneous throughout the SPG. Most of the highly abundant organisms, including MGII *Euryarchaeota*, *Prochlorococcus*, *Pelagibacter ubique* (SAR11 clade), SAR 86, and unclassified *Flavobacteriaceae*, carry genes that encode machinery for either photosynthesis or phototrophy (Partensky et al. 1999, Sabehi et al. 2004, Campbell et al. 2008, Riedel et al. 2010), which may provide an advantage in this oligotrophic environment. Both *Prochlorococcus* and *Pelagibacter ubique* provide excellent examples of genomic streamlining in order to confer an advantage in an oligotrophic environment (Dufresne et al. 2005, Giovannoni et al. 2005, Eiler et al. 2009). Concentrations of other abundant organisms, such as SAR406 in the *Deferribacteres* cluster, have been found to be highly correlated with levels of chl *a* and have been speculated to benefit directly from phytoplankton exudates (Gordon & Giovannoni 1996).

These dominant bacterial and archaeal taxa are also present at other DCM locations such as the Sargasso Sea, Mediterranean Sea, and the Southern California Bight (Treusch et al. 2009, Ghai et al.

2010, Beman et al. 2011). Similar trends in community composition have also been observed in the Mediterranean and Sargasso Seas, such as the elevated number of tags associated with SAR11 and the overall dominance of *Alphaproteobacteria* and *Cyanobacteria* at this horizon (Treusch et al. 2009, Pommier et al. 2010). The dissimilarity between the DCM of the NPG and that of the SPG is due primarily to the abundance of SAR11 tags, which contribute to 30% of the Bray-Curtis dissimilarity between regions. The abundance of SAR11 gene copy numbers in the NPG (at Stn ALOHA) correlates with diatom abundance, primary productivity, dissolved organic carbon, and chl *a* (Eiler et al. 2009). A recent survey of photosynthetic picoeukaryotic composition (0–3 μm) in the SPG found no evidence of diatoms at either the DCM or in the surface water (Shi et al. 2009). However, diatoms were found at mesotrophic stations just outside the gyre margin. Although both the SPG and NPG are oligotrophic regions, these observations and the relatively higher concentrations of sea-surface chlorophyll at Stn ALOHA than in the SPG may explain the relatively higher proportion of SAR11 tags at Stn ALOHA than in the SPG.

Comparison of the SPG and NPG communities to the EQP communities is complicated by the use of different forward primers for the EQP samples. Differential amplification as a result of primer bias is a well-recognized problem in molecular surveys; both amplicon length and primer choice, even targeting the same region of the 16S gene, can influence estimates of microbial species richness (Engelbrektson et al. 2010). Consequently, although the SPG, NPG, and EQP samples were all analyzed with the same reverse primers and we trimmed the EQP sequences to the same size and hyper-variable region as the SPG and NPG sequences, we cannot fully rule out the possibility that differences between the EQP communities and the SPG and NPG communities resulted from primer bias.

With this caveat in mind, major differences between the EQP communities and the gyre communities are consistent with environmental differences between these oceanographic regions. In particular, the EQP sites were differentiated by their abundance of *Flavobacteriaceae* tags. This group was responsible for 15% of the Bray-Curtis dissimilarity between regions. This distinction is consistent with the high productivity of Stns EQP1 and EQP8 relative to sites in the NPG and SPG, as this family is associated with high-productivity regimes, such as phytoplankton blooms (Pinhassi et al. 2004).

Gyre margin versus gyre center

Although the dominant taxa are the same at all of our SPG sites, their relative proportions vary from western gyre margin to gyre center to southern gyre margin. This result is consistent with environmental characteristics of the gyre controlling the composition of both bacterial and archaeal assemblages in the SPG. The rotation of the gyre shapes environmental variables including DCM depth, by elevating the thermocline and nutricline at the gyre margins and depressing them at the gyre center (Pennington et al. 2006). Lack of upwelling in the gyre center renders the sea surface nutrient-deprived, reducing sea-surface primary production. Results from our CCA appear to reflect this phenomenon, as both bacterial and archaeal communities vary with sea-surface chlorophyll and depth within the gyre (Fig. 4).

Archaeal assemblage composition demonstrates a clear shift in the abundance of MGII and MGIII *Euryarchaeota* in transects from the shallow and more productive western and southern gyre margin sites toward the deeper low-productivity central gyre stations (Fig. 2). The cause of this shift is unclear; MGIII *Euryarchaeota* are generally considered to be rare (Massana et al. 2000, Galand et al. 2009). Only one other study has identified an MGIII-dominated community, which was in cold deep water of the western Arctic Ocean (Galand et al. 2009). The ecological role of this often rare but ubiquitous archeon is presently enigmatic.

Community similarity and habitat partitioning

Hewson et al. (2006) discovered that composition of bacterial communities is remarkably consistent in oligotrophic surface waters over a horizontal spatial scale of a few kilometers. They described this phenomenon as a 'patch size of homogeneity,' which they observed to break down over 15 to 50 km scales. This 'patch size' was greatly expanded by a study targeting multiple water masses along a North Atlantic transect from 60°N to 5°S, which demonstrated that samples taken thousands of kilometers apart within a single water mass resembled each other more closely in bacterial community composition than they resembled samples vertically separated by hundreds of meters in water depth (Agogue et al. 2011).

Our study expands on the results of Agogue et al. (2011) by demonstrating that both bacterial and

archaeal tags are remarkably similar within the even larger region of the SPG. Throughout the SPG, the sampled bacterial and archaeal populations differ by only 12.5–28% and 2–16%, respectively, at the 97% similarity taxonomic level in terms of a Bray-Curtis dissimilarity score. This similarity is striking, as the sampled locations are horizontally separated by 100s to 1000s of kilometers of open ocean and differ in water depth by up to 100 m.

Agogu e et al. (2011) attributed community differences between water masses to physical barriers to dispersal. However, our results suggest that environmental preferences (illustrated by strong correlations to sea-surface properties such as chlorophyll and *in situ* temperature and nutrient availability) may be more important than barriers to dispersal in defining community composition; bacterial assemblages of the environmentally analogous NPG samples more closely resemble those from our SPG samples than either resembled the assemblages from the nutrient- and chlorophyll-rich samples of the geographically closer EQP. However, it is important to acknowledge the potential for primer bias when comparing our EQP data to the NPG and SPG data.

Despite their similarities, the SPG and NPG communities have distinct taxonomic compositions (>45% dissimilarity). Differences in the proportion of the dominant taxa as well as differences in the rare community members appear unlikely to result from differences in sampling season. The SPG gyre samples were taken during the southern hemisphere summer (late December/January 2006/2007) and seasonally corresponded with NPG gyre sample HOT 194 (August 2007). The remaining NPG samples were taken in October (2006) and February (2007) and provide a broad perspective on seasonality within the NPG. As the NPG sample collections occurred in different months and seasons, and these samples are statistically indistinguishable from each other, seasonality alone does not explain differences between the SPG and NPG samples. While relationships between seasonality and DCM communities can be complicated, recent studies have demonstrated highly repeatable seasonal patterns in bacterial community composition in marine surface waters (Fuhrman et al. 2006, Needham et al. 2013). Consequently, differences between the SPG and NPG in the proportion of dominant community members are likely to result from environmental selection (Fig. 2). In contrast, differences in presence or absence of rare taxa may at least partly reflect allopatric effects, such as currents that block cross-equatorial exchange (e.g. Agogu e et al. 2011).

CONCLUSIONS

We sampled DCM prokaryotic communities throughout Earth's largest oceanographic province, the SPG. This province constitutes 15% of the global ocean area. While differences in community composition along vertical gradients are well documented, few tag-sequencing studies have examined horizontal changes and even fewer have followed an oceanographic horizon, such as the DCM. None have previously examined the SPG. Bacterial and archaeal assemblages are remarkably homogeneous in the DCM of the SPG, as well as across the Pacific. The most abundant taxa are the same at all sites examined, suggesting that these organisms have adapted to dominate this biologically significant horizon ocean-wide. Significant differences in assemblage composition appear to directly result from environmental heterogeneity. Within the SPG, sea-surface chlorophyll content and water depth are both highly correlated to bacterial community composition in the DCM. Of these environmental properties, sea-surface chlorophyll content appears to be the most significant predictor of DCM community structure within the SPG, as well as across the Pacific Ocean, despite large distances between sites and pronounced differences in the depth (up to 200 m), temperature (9.1 C), and other properties of the DCM from one station to another.

Acknowledgements. We thank Captain Tom Desjardins and the crew and the scientific party of RV 'Roger Revelle' Expedition KNOX02RR for making our SPG sampling possible. We thank Captain Kent Sheasley and the crew and scientific party of RV 'Knorr' Expedition 195-3 for making our EQP sampling possible. We thank Robert Pockalny for creating the station location map and Andrew Voorhis for trimming the EQP sequencing data. We also thank John B. Kirkpatrick and Justin Abreu for their reviews of this work. This study was funded by the Ocean Drilling and Biological Oceanography Programs of the US National Science Foundation (Grants OCE-0527167 and OCE- 0752336).

LITERATURE CITED

- Agogu e H, Lamy D, Neal PR, Sogin ML, Herndl GJ (2011) Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Mol Ecol* 20:258–274
- Aller JY, Kemp PF (2008) Are Archaea inherently less diverse than Bacteria in the same environments? *FEMS Microbiol Ecol* 65:74–87
- Baas-Becking LGM (1934) *Geobiologie of inleiding tot de milieukund*. WP Van Stockum and Zoom, The Hague
- Behrenfeld MJ, Falkowski PG (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol Oceanogr* 42:1–20

- Beman JM, Steele JA, Fuhrman JA (2011) Co-occurrence patterns for abundant marine archaeal and bacterial lineages in the deep chlorophyll maximum of coastal California. *ISME J* 5:1077–1085
- Boyle T, Smillie G, Anderson J, Beeson D (1990) A sensitivity analysis of nine diversity and seven similarity indices. *Res J Water Pollut Control Fed* 62:749–762
- Campbell BJ, Waidner LA, Cottrell MT, Kirchman DL (2008) Abundant proteorhodopsin genes in the North Atlantic Ocean. *Environ Microbiol* 10:99–109
- Clarke KR, Gorley RN (2006) PRIMER v6: user manual/tutorial. PRIMER-E, Plymouth
- D'Hondt S, Spivack A, Pockalny R, Ferdlema T and others (2009) Subseafloor sedimentary life in the South Pacific Gyre. *Proc Natl Acad Sci USA* 106:11651–11656
- Dufour P, Charpy L, Bonnet S, Garcia N (1999) Phytoplankton nutrient control in the oligotrophic South Pacific subtropical gyre (Tuamotu Archipelago). *Mar Ecol Prog Ser* 179:285–290
- Dufresne A, Garczarek L, Partensky F (2005) Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol* 6:R14
- Eiler A, Hayakawa DH, Church MJ, Karl DM, Rappé MS (2009) Dynamics of the SAR11 bacterioplankton lineage in relation to environmental conditions in the oligotrophic North Pacific subtropical gyre. *Environ Microbiol* 11:2291–2300
- Engelbrektson A, Kunin V, Wrighton KC, Zvenigorodsky N, Chen F, Ochman H, Hugenholtz P (2010) Experimental factors affecting PCR-based estimates of microbial species richness and evenness. *ISME J* 4:642–647
- Fenchel T, Finlay BJ (2004) The ubiquity of small species: patterns of local and global diversity. *BioScience* 54:777–784
- Friedline CJ, Franklin RB, McCallister SL, Rivera MC (2012) Bacterial assemblages of the eastern Atlantic Ocean reveal both vertical and latitudinal biogeographic signatures. *Biogeosciences* 9:2177–2193
- Fuhrman JA, Hewson I, Schwabach MS, Steele JA, Brown MV, Naeem S (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Natl Acad Sci USA* 103:13104–13109
- Galand PE, Casamayor EO, Kirchman DL, Potvin M, Lovejoy C (2009) Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *ISME J* 3:860–869
- Ghai R, Martin-Cuadrado AB, Molto AG, Heredia IG and others (2010) Metagenome of the Mediterranean deep chlorophyll maximum studied by direct and fosmid library 454 pyrosequencing. *ISME J* 4:1154–1166
- Ghiglione JF, Galand PE, Pommier T, Pedrós-Alió C and others (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc Natl Acad Sci USA* 109:17633–17638
- Giovannoni SJ, Tripp H, Givan S, Podar M and others (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–1245
- Gordon DA, Giovannoni SJ (1996) Detection of stratified microbial populations related to *Chlorobium* and *Fibrobacter* species in the Atlantic and Pacific oceans. *Appl Environ Microbiol* 62:1171–1177
- Gregg WW, Casey NW, McClain CR (2005) Recent trends in global ocean chlorophyll. *Geophys Res Lett* 32:L03606
- Halm H, Lam P, Ferdelman TG, Lavik G and others (2012) Heterotrophic organisms dominate nitrogen fixation in the South Pacific Gyre. *ISME J* 6:1238–1249
- Hewson I, Capone DG, Steele JA, Fuhrman JA (2006) Influence of Amazon and Orinoco offshore surface water plumes on oligotrophic bacterioplankton diversity in the west tropical Atlantic. *Aquat Microb Ecol* 43:11–22
- Hibbing ME, Fuqua C, Parsek MR, Peterson SB (2010) Bacterial competition: surviving and thriving in the microbial jungle. *Nat Rev Microbiol* 8:15–25
- Huber JA, Mark Welch DB, Morrison HG, Huse SM, Neal PR, Butterfield DA, Sogin ML (2007) Microbial population structures in the deep marine biosphere. *Science* 318:97–100
- Huisman J, Pham Thi NN, Karl DM, Sommeijer B (2006) Reduced mixing generates oscillations and chaos in the oceanic deep chlorophyll maximum. *Nature* 439:322–325
- Hunt DE, Lin Y, Church DB, Karl DM, Tringe SG, Izzo LK, Johnson ZI (2013) Relationship between abundance and specific activity of bacterioplankton in open ocean surface waters. *Appl Environ Microbiol* 79:177–184
- Huse S, Huber J, Morrison H, Sogin M, Welch D (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* 8:R143
- Huse SM, Dethlefsen L, Huber JA, Mark Welch D, Relman DA, Sogin ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genet* 4:e1000255
- Huse SM, Mark Welch D, Morrison H, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. *Environ Microbiol* 12:1889–1898
- Karl DM, Lukas R (1996) The Hawaii Ocean Time-series (HOT) program: background, rationale and field implementation. *Deep-Sea Res II* 43:129–156
- Martiny JBH, Eisen JA, Penn K, Allison SD, Horner-Devine MC (2011) Drivers of bacterial β -diversity depend on spatial scale. *Proc Natl Acad Sci USA* 108:7850–7854
- Massana R, DeLong EF, Pedrós-Alió C (2000) A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl Environ Microbiol* 66:1777–1787
- Morel A, Gentili B, Claustre H, Babin M, Bricaud A, Ras J, Tieche F (2007) Optical properties of the 'clearest' natural waters. *Limnol Oceanogr* 52:217–229
- Needham DM, Chow CET, Cram JA, Sachdeva R, Parada A, Fuhrman JA (2013) Short-term observations of marine bacterial and viral communities: patterns, connections and resilience. *ISME J* 7:1274–1285
- Oksanen J, Blanchet GF, Kindt R, Legendre P and others (2011) vegan: community ecology package. R package version 2.2-1. <http://CRAN.R-project.org/package=vegan>
- Papke RT, Ramsing NB, Bateson MM, Ward DM (2003) Geographical isolation in hot spring cyanobacteria. *Environ Microbiol* 5:650–659
- Partensky F, Hess WR, Vaultot D (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* 63:106–127
- Pennington TJ, Mahoney KL, Kuwahara VS, Kolber DD, Calienes R, Chavez FP (2006) Primary production in the eastern tropical Pacific: a review. *Prog Oceanogr* 69:285–317
- Pinhassi J, Sala MM, Havskum H, Peters F and others (2004) Changes in bacterioplankton composition under different phytoplankton regimens. *Appl Environ Microbiol* 70:6753–6766
- Polovina J, Howell E, Abecassis M (2008) Ocean's least pro-

- ductive waters are expanding. *Geophys Res Lett* 35:1–5, doi:10.1029/2007GL031745
- Pommier T, Neal PR, Gasol JM, Coll M, Acinas SG, Pedrós-Alió C (2010) Spatial patterns of bacterial richness and evenness in the NW Mediterranean Sea explored by pyrosequencing of the 16S rRNA. *Aquat Microb Ecol* 61: 221–233
- Riedel T, Tomasch J, Buchholz I, Jacobs J and others (2010) Constitutive expression of the proteorhodopsin gene by a *Flavobacterium* strain representative of the proteorhodopsin-producing microbial community in the North Sea. *Appl Environ Microbiol* 76:3187–3197
- Sabehi G, Béjà O, Suzuki MT, Preston CM, DeLong EF (2004) Different SAR86 subgroups harbour divergent proteorhodopsins. *Environ Microbiol* 6:903–910
- Schauer R, Bienhold C, Ramette A, Harder J (2010) Bacterial diversity and biogeography in deep-sea surface sediments of the South Atlantic Ocean. *ISME J* 4:159–170
- Shi XL, Marie D, Jardillier L, Scanlan DJ, Vaulot D (2009) Groups without cultured representatives dominate eukaryotic picophytoplankton in the oligotrophic South East Pacific Ocean. *PLoS ONE* 4:e7657
- Sogin M, Morrison H, Welch D, Huse S, Neal P, Arrieta J, Herndl GJ (2006) Microbial diversity in the deep-sea and the underexplored 'rare biosphere'. *Proc Natl Acad Sci USA* 103:12115–12120
- Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA, Giovannoni SJ (2009) Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J* 3:1148–1163
- Weston K, Fernand L, Mills DK, Delahunty R, Brown J (2005) Primary production in the deep chlorophyll maximum of the central North Sea. *J Plankton Res* 27:909–922
- Whitaker RJ, Grogan DW, Taylor JW (2003) Geographic barriers isolate endemic populations of hyperthermophilic Archaea. *Science* 301:976–978
- Winter C, Moeseneder M, Herndl GJ, Weinbauer M (2008) Relationship of geographic distance, depth, temperature, and viruses with prokaryotic communities in the eastern tropical Atlantic Ocean. *Microb Ecol* 56:383–389

Editorial responsibility: Gerhard Herndl, Vienna, Austria

*Submitted: June 3, 2014; Accepted: February 18, 2015
Proofs received from author(s): April 6, 2015*