OPEN

The ISME Journal (2016) 10, 979–989

www.nature.com/ismej

ORIGINAL ARTICLE

Bacterial diversity and community composition from seasurface to subseafloor

Emily A Walsh[1,2], John B Kirkpatrick[3], Scott D Rutherford[4], David C Smith[3], Mitchell Sogin[5] and Steven D'Hondt[3]

[1]*Department of Microbiology, The Forsyth Institute, Cambridge, MA, USA;* [2]*Harvard School of Dental Medicine, Boston, MA, USA;* [3]*Graduate School of Oceanography, University of Rhode Island, Narragansett Bay Campus, Narragansett, RI, USA;* [4]*Department of Environmental Sciences, Roger Williams University, Bristol, RI, USA and* [5]*Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, Woods Hole, MA, USA*

We investigated compositional relationships between bacterial communities in the water column and those in deep-sea sediment at three environmentally distinct Pacific sites (two in the Equatorial Pacific and one in the North Pacific Gyre). Through pyrosequencing of the v4–v6 hypervariable regions of the 16S ribosomal RNA gene, we characterized 450 104 pyrotags representing 29 814 operational taxonomic units (OTUs, 97% similarity). Hierarchical clustering and non-metric multi-dimensional scaling partition the samples into four broad groups, regardless of geographic location: a photic-zone community, a subphotic community, a shallow sedimentary community and a subseafloor sedimentary community (⩾1.5 meters below seafloor). Abundance-weighted community compositions of water-column samples exhibit a similar trend with depth at all sites, with successive epipelagic, mesopelagic, bathypelagic and abyssopelagic communities. Taxonomic richness is generally highest in the water-column O_2 minimum zone and lowest in the subseafloor sediment. OTUs represented by abundant tags in the subseafloor sediment are often present but represented by few tags in the water column, and represented by moderately abundant tags in the shallow sediment. In contrast, OTUs represented by abundant tags in the water are generally absent from the subseafloor sediment. These results are consistent with (i) dispersal of marine sedimentary bacteria via the ocean, and (ii) selection of the subseafloor sedimentary community from within the community present in shallow sediment.
The ISME Journal (2016) **10,** 979–989; doi:10.1038/ismej.2015.175; published online 2 October 2015

Introduction

Previous studies indicate that microbial community composition varies from one marine environment to another (Delong *et al.*, 2006; Zinger *et al.*, 2011; Hamdan *et al.*, 2013), but can be relatively consistent in similar marine environments separated by long distances (Inagaki *et al.*, 2006; Agogué *et al.*, 2011; Walsh *et al.*, 2015). For example, adjacent marine water masses with different environmental properties have distinct microbial communities (Agogué *et al.*, 2011; Hamdan *et al.*, 2013). Similarly, subseafloor sedimentary environments with different properties separated by a few tens of kilometers also have distinct communities (Inagaki *et al.*, 2006; Hewson *et al.*, 2007; Schauer *et al.*, 2010). Despite these differences over relatively short geographic distances, microbial community composition in individual deep-seawater masses can be relatively constant for thousands of kilometers (Agogué *et al.*, 2011). And broadly, similar microbial communities inhabit similar subseafloor sedimentary environments separated by thousands of kilometers (Inagaki *et al.*, 2006). These observations are consistent with the old adage, 'Everything is everywhere but the environment selects' (Baas-Becking, 1934), in which microorganisms are considered to be ubiquitously dispersed because of their small size, large numbers and low extinction rates (Martiny *et al.*, 2006). Despite these intriguing biogeographic patterns, few previous studies have examined (i) the vertical distribution of bacterial taxa throughout the entire open-ocean water column, or (ii) the relationship of microbial communities in marine sediment to those in the overlying water.

Distributions of abundant bacterial taxa in surface marine sediment (to 1.1 meters below seafloor (mbsf)) suggest that the seafloor is colonized via the overlying water column (Hamdan *et al.*, 2013). However, the relationship of microbial communities in deep subseafloor sediment (⩾1.5 mbsf) to those in

Correspondence: EA Walsh, The Forsyth Institute, 245 First Street, Cambridge, MA 02145, USA.
E-mail: ewalsh@forsyth.org
Received 9 March 2015; revised 10 August 2015; accepted 13 August 2015; published online 2 October 2015

surface marine sediment and the overlying ocean has not been previously examined.

Vertically sampling the water column and sediment in distinct oceanographic regions provides a means to evaluate the influence of local oceanographic properties such as nutrient availability, sedimentation rate and distance from land (Fuhrman *et al.*, 2006; Kallmeyer *et al.*, 2012) on microbial diversity throughout both water column and sediment. High productivity in surface water can lead to the formation of oxygen-deficient zones in deeper water, where unique communities may form in response to the deoxygenation (Beman and Carolan, 2013). In high-productivity regimes, such as the eastern equatorial Pacific upwelling region (EQP-1) and the moderately productive central equatorial Pacific upwelling region (EQP-8), a relatively high flux of organic debris to the seafloor sustains an active anaerobic subseafloor sedimentary community (D'Hondt *et al.*, 2004). In contrast, low-productivity regions, such as the North Pacific gyre (EQP-11), are characterized by a continuously oxic water column, extremely low rates of sedimentation and organic matter deposition, and a relatively low-activity aerobic subseafloor community (D'Hondt *et al.*, 2009; Røy *et al.*, 2012; D'Hondt *et al.*, 2015).

(i) To document vertical distributions of bacterial diversity and community composition, (ii) to investigate how these communities vary from one oceanographic region to another, and (iii) to test the extent to which bacterial communities of deep subseafloor sediment may originate in the water column, we used 454 pyrosequencing technology and the v4–v6 hypervariable region of the bacterial 16S ribosomal RNA (rRNA) gene to examine bacterial community composition (presence–absence and relative abundance) in the water column, near-seafloor sediment (0–10 centimeters below seafloor (cmbsf)), and subseafloor sediment at three environmentally distinct Pacific sites: the very high-productivity eastern equatorial upwelling region (EQP-1), the moderately high-productivity open-ocean central equatorial upwelling region (EQP-8) and the very low-productivity northern gyre (EQP-11) (Figure 1).

Materials and methods

Sampling methodology

We collected the sediment and seawater samples during a 43-day transect of the *R/V Knorr* through the eastern equatorial Pacific and the North Pacific gyre (*Knorr* Expedition 195-3). We collected the water samples using a 24-Niskin bottle CTD rosette (Sea-Bird SBE-911 CTD system plus package). Once the CTD/Niskin system was on deck, we clamped the Niskin bottles and applied compressed air to the vent plugs to enhance filtration. We filtered 5 l from each bottle onto 0.2 μM Supor-200 47 mm filters at a rate of approximately 200 ml min^{-1}, and then stored the filters at −80 °C. We targeted a combination of standard depths and oceanographic horizons (seasurface, O_2 minimum, chlorophyll *a* maximum, thermocline and deep water) and analyzed 7–14 sampling depths at each station from 3 to 5500 m water depth depending on location. We measured water depths using the center depth of the SeaBeam 2112 multibeam. We estimated the average sedimentation rate for each site by dividing the crustal age (Müller *et al.*, 2008) by the sediment thickness in the NDGC global thickness grid (Divins, 2003). We used mean annual average chlorophyll *a* data (Behrenfeld and Falkowski, 1997; Gregg *et al.*, 2005), collected from September 1994–December 2004, as a proxy for seasurface productivity.

At each station, we collected sediment using a multi-corer (0–30 cmbsf]), gravity corer (up to 4 mbsf) and long-piston coring device (up to 35 mbsf). We subsampled the sediment using sterile 60 cc cut-off syringes and immediately froze the subsamples at − 80 °C for later DNA extraction. For each site in this study, we analyzed one sediment sample taken with the multi-corer at the sediment-water interface

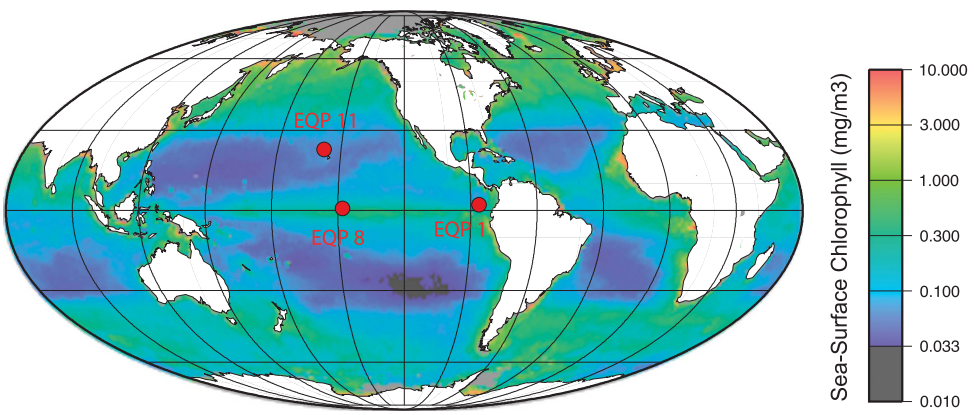

Figure 1 Locations of sampling stations. Superimposed chlorophyll *a* content is annual average data (mean from September 1994–December 2004) from the SeaWiFS satellite uploaded into GeoMapApp (Behrenfeld and Falkowski, 1997; Gregg *et al.*, 2005).

(0–10 cmbsf) and two to five sediment samples taken at greater depths (between 0.25 and 34 mbsf). In total, this study includes 39 tag-pyrosequencing samples: 28 from the water column, 3 from the sediment-water interface and 8 from subseafloor sediment.

Site descriptions

All three sites are located in the deep open ocean, with fully pelagic sediment.

EQP-1 (1° 48.21' N, 86° 11.29' W) was the easternmost site in our equatorial Pacific transect. This site is located in a relatively flat region of 1.6-Ma basement (Müller *et al.*, 2008), with water depth of 2856 m (Figure 1). It is in the eastern equatorial upwelling regime. Consequently, it contains a highly productive ecosystem and is characterized by a high mean sedimentation rate (~75 m Ma^{-1}). Its sediment is predominantly foraminifera-rich nannofossil ooze. Total sediment thickness is estimated to be 120 m (Divins, 2003). We began sampling with the CTD/Niskin cast on 14 January 2009 (bottles on deck 19:33 GMT). The site has a pronounced O_2 minimum zone. The chlorophyll maximum was strongly developed at our time of sampling (at 34 m) and located above the depth of the pycnocline. The O_2 minimum zone was within the thermocline and contained O_2 concentrations as low as 6 µmol kg^{-1} (Figure 2). In the sediment, dissolved oxygen was limited to the upper 7 mm (Røy *et al.*, 2012).

EQP-8 (0° 0.36' N, 147° 47.50' W) was the westernmost site of our equatorial transect (Figure 1). It is located on a flat abyssal plain, with basement age of 80 Ma (Müller *et al.*, 2008) and water depth of 4336 m. Total sediment thickness is estimated to be 380 m (Divins, 2003). This site is in the Central Pacific equatorial upwelling region. Consequently, it has high surface productivity for an open-ocean location. Its sediment is calcareous nannofossil ooze. The mean sedimentation rate at this site (4.8 m Ma^{-1}) is substantially lower than at EQP-1. We conducted water-column sampling here on 5 February 2009 (CTD/Niskin system on deck 4:54 GMT). At this site, the chlorophyll maximum was slightly deeper in the water column than at EQP-1 at the time of sampling (41 m), but in contrast to EQP-1, was above the depth of the pycnocline. The O_2 minimum at this site was within the thermocline, with a concentration of 29 µmol kg^{-1} (Figure 2). In the sediment, dissolved oxygen penetrated ~10 cm below seafloor (Røy *et al.*, 2012).

EQP-11 (30° 21.32' N, 157° 52.23' W) is in the North Pacific gyre. It is located on a very slight topographic high with basement age of 88.7 Ma (Müller *et al.*, 2008). With water depth of 5813 m, this is our deepest water site. Total sediment thickness is estimated to be 100 m (Divins, 2003). We began water-column sampling on the 19th of February (CTD on deck 06:12 GMT). The deep chlorophyll maximum (DCM) was in relatively deep water (127 mbsl) and not strongly developed. The O_2 minimum at this site (18 µmol kg^{-1}) was also deep in the water column (900 mbsl). As at the other sites, this O_2 minimum was located within the thermocline. Mean sedimentation rate is extremely low at EQP-11 (~1.1 m Ma^{-1}). The sediment is pelagic clay and dissolved oxygen penetrates past the greatest depth cored (28.06 mbsf; Røy *et al.*, 2012).

DNA extraction, pyrosequencing and statistical analyses

We applied a high-resolution v4–v6 tag-pyrosequencing approach of the bacterial 16S rRNA gene to a total of 39 water and sediment samples taken

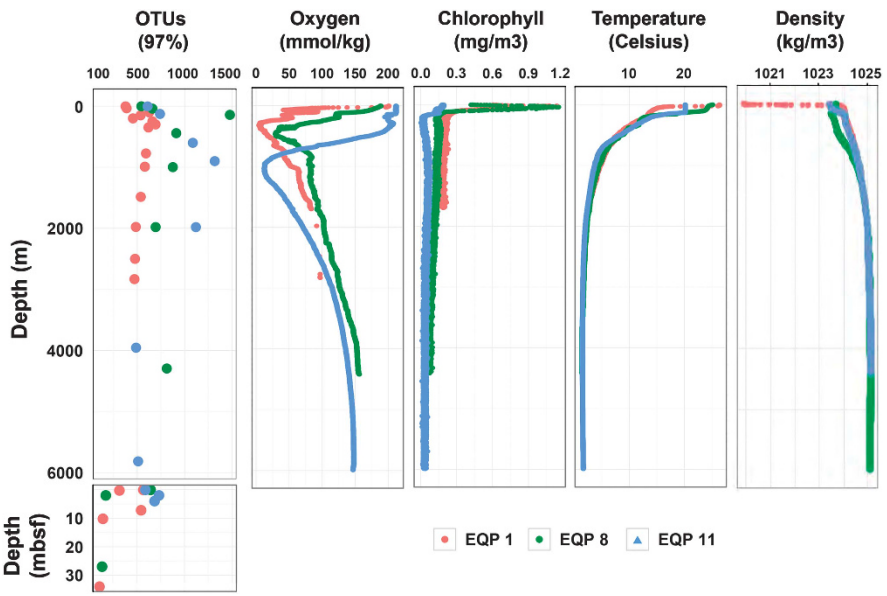

Figure 2 Scatterplot of sequencing and CTD data including (from left to right): OTUs (97%), oxygen, chlorophyll *a*, temperature and density.

Table 1 Sampling information and diversity estimates for water and sediment collected from our three sites

Sample ID	*Sample type*	*Water or sediment depth (m)*	*Oxygen (mmol kg^{-1})*	*Chlorophyll (mg m^{-3})*	*AOU (μmol kg^{-1})*	*OTUs (97%)*	*Chao-1 (97%)*	*PD*	*Evenness (97%)*
EQP-1 (3 m)	Seawater	3	201	0.61	7	376	593	51	0.69
EQP-1 (34 m)	Seawater	34	102	1.08	136	387	778	47	0.76
EQP-1 (101 m)	Seawater	101	93	0.35	159	625	1254	64	0.82
EQP-1 (150 m)	Seawater	150	45	0.24	211	535	884	60	0.83
EQP-1 (200 m)	Seawater	200	55	0.24	203	454	679	61	0.83
EQP-1 (249 m)	Seawater	249	21	0.20	241	660	1393	67	0.81
EQP-1 (299 m)	Seawater	299	6	0.22	264	693	1748	77	0.80
EQP-1 (348 m)	Seawater	348	6	0.22	268	618	1308	67	0.78
EQP-1 (774 m)	Seawater	774	43	0.20	260	592	1092	66	0.79
EQP-1 (991 m)	Seawater	991	52	0.19	258	580	1078	61	0.78
EQP-1 (1485 m)	Seawater	1485	73	0.19	249	537	1010	65	0.78
EQP-1 (1978 m)	Seawater	1978	92	NA	237	485	736	62	0.80
EQP-1 (2500 m)	Seawater	2500	96	NA	236	476	857	61	0.78
EQP-1 (2832 m)	Seawater	2832	97	NA	237	470	685	61	0.81
EQP-1-SEDA	Sediment	0.05	0	NA	NA	569	677	68	0.89
EQP-1-SEDB	Sediment	0.25	0	NA	NA	319	392	44	0.82
EQP-1-SEDC	Sediment	7.2	0	NA	NA	546	736	72	0.85
EQP-1-SEDD	Sediment	10.2	0	NA	NA	147	201	14	0.72
EQP-1-SEDE	Sediment	34	0	NA	NA	111	168	14	0.63
EQP-8 (2 m)	Seawater	2	188	0.43	18	543	1786	61	0.73
EQP-8 (41 m)	Seawater	41	191	1.07	28	660	3017	72	0.74
EQP-8 (142 m)	Seawater	142	124	0.18	105	1479	7467	143	0.89
EQP-8 (442 m)	Seawater	442	30	0.16	249	912	3370	100	0.82
EQP-8 (998 m)	Seawater	998	84	0.14	227	876	3188	91	0.81
EQP-8 (1980 m)	Seawater	1980	102	0.05	227	694	2209	83	0.79
EQP-8 (4296 m)	Seawater	4296	154	0.07	182	812	2858	99	0.78
EQP-8-SEDA	Sediment	0.05	34	NA	NA	646	927	80	0.84
EQP-8-SEDB	Sediment	2	0	NA	NA	177	229	28	0.58
EQP-8-SEDC	Sediment	27	0	NA	NA	138	173	19	0.57
EQP-11 (3 m)	Seawater	3	212	0.18	13	609	1627	69	0.71
EQP-11 (127 m)	Seawater	127	212	0.14	13	740	2378	82	0.74
EQP-11 (600 m)	Seawater	600	125	0.05	174	1087	3532	118	0.84
EQP-11 (900 m)	Seawater	900	19	0.07	297	1319	5539	138	0.87
EQP-11 (1981 m)	Seawater	1981	70	0.05	261	1120	4470	122	0.85
EQP-11 (3955 m)	Seawater	3955	136	0.03	201	487	1132	67	0.74
EQP-11 (5816 m)	Seawater	5816	148	0.03	186	509	1242	72	0.77
EQP-11-SEDA	Sediment	0.05	101	NA	NA	592	1216	74	0.80
EQP-11-SEDB	Sediment	2	73	NA	NA	733	2978	71	0.59
EQP-11-SEDC	Sediment	4	60	NA	NA	687	2924	68	0.57

Abbreviations: AOU, apparent oxygen utilization; NA, not applicable; OTU, operational taxonomic unit; PD, phylogenetic diversity.
Diversity estimates are based on 16S rRNA 454 pyrosequencing data with each sample randomly subsampled to the same number of reads as the sample with the initial lowest number of reads (2880).

vertically from seasurface to subseafloor at our three sites (Figure 1, Table 1). We extracted DNA from filtered water samples using the Power Water DNA Isolation Kit (MoBio, Carlsbad, CA, USA) and from sediment samples using the Power Soil DNA Isolation Kit (MoBio). We modified the extraction protocol for the deep EQP-11 sediment samples (2 and 4 m) to overcome effects of low biomass and DNA-binding clay (Cai *et al.*, 2006; Kallmeyer *et al.*, 2012). These modifications were (i) pooling of 10x the sediment (10 vs 1 g) and (ii) use of extended PCR cycle numbers (40 × vs 30 ×). In addition, we used a phenol–chloroform DNA extraction for water samples from EQP-1. We used fusion primers 518 f and 1064r, targeting the v4–v6 hypervariable regions of the 16S rRNA gene, to construct amplicon libraries using the following PCR conditions: 94 °C for 2 min, 30 cycles of (94 °C for 30 s, 55 °C for 20 s, 72 °C for 1 min). We pooled three separate PCR reactions per sample to eliminate the potential for early cycle PCR-induced error. We ran negative controls during all PCR reactions. All sequenced reactions produced bands, whereas controls did not. We sequenced pooled amplicons using a FLX 454 pyrosequencer (Roche, Branford, CT, USA).

We discarded low-quality reads, defined as (i) reads with undefined residues and (ii) reads that did not contain the PCR primers at the beginning of each read (Huse *et al.*, 2007). We submitted our sequence data to the GenBank database under study accession no. SRP058974. We taxonomically assigned reads, based on ⩾80% sequence similarity, using the GAST system for sequence identification (Sogin *et al.*, 2006).

To ensure inter-sample comparability for our taxonomic richness estimates and statistical analyses, we utilized the QIIME software (Boulder, CO, USA) to randomly reduce the number of reads in each sample to the lowest number of reads in any individual sample (2880). We also used the QIIME program to calculate taxonomic richness (total operational taxonomic units (OTUs), Chao-1, ACE and phylogenetic diversity (PD); Caporaso *et al.*, 2010).

We used the Primer-E software package (Clarke and Gorley, 2006) for (i) calculation of Bray–Curtis similarity indices and (ii) all ordination and statistical methods including non-metric multidimensional scaling, hierarchical group-average clustering and Spearman rank correlation tests. We used this output to construct figures using the ggplot2 package (Wickham, 2009) in R computer language (R Core Team, 2014).

To test the hypothesis that OTUs in subseafloor sediment are preferentially from the rare biosphere in the water column, rather than a random sample of water-column OTUs, we developed a Matlab simulation to randomly resample the water-column OTU frequency data with replacement. We repeated this sampling 5000 times to create a distribution of the number of OTUs randomly drawn from the rare (rank abundance >100) water-column OTUs. We then used this distribution to estimate the probability of randomly picking 'n' rare OTUs from the water-column abundance distribution.

Results

Similarity among bacterial communities

Hierarchical 'group-average' clustering and non-metric multidimensional scaling analyses separate our samples into four distinct abundance-weighted community compositions, based on Bray–Curtis similarity (Figure 3a). Regardless of site, water-column samples partition into (i) photic-zone communities, which include samples from at or above the DCM, and (ii) aphotic communities (below the DCM). The GAST taxonomy demonstrated that Cyanobacteria, Flavobacteria and Alphaproteobacteria dominate the photic-zone samples, whereas the aphotic-zone samples are heavily dominated by Alphaproteobacteria, Gammaproteobacteria, Deltaproteobacteria and Deferribacteres (Figure 4).

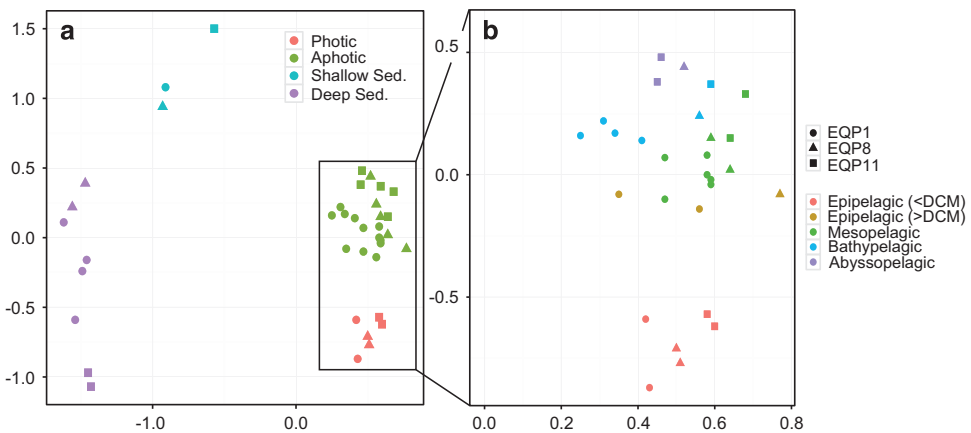

Figure 3 (**a**) Non-metric multidimensional scaling (nMDS) plot of the water column and sediment community compositions. Axes do not represent any measured parameter, but define a 2-D space that allows the best spatial representation of sample similarity, based on Bray–Curtis similarity indices. (**b**) Enlargement of the nMDS plot for the water-column communities with a superimposed color gradient representing depth from seasurface (red) to the abyssopelagic water of EQP-8 and EQP-11 (blue).

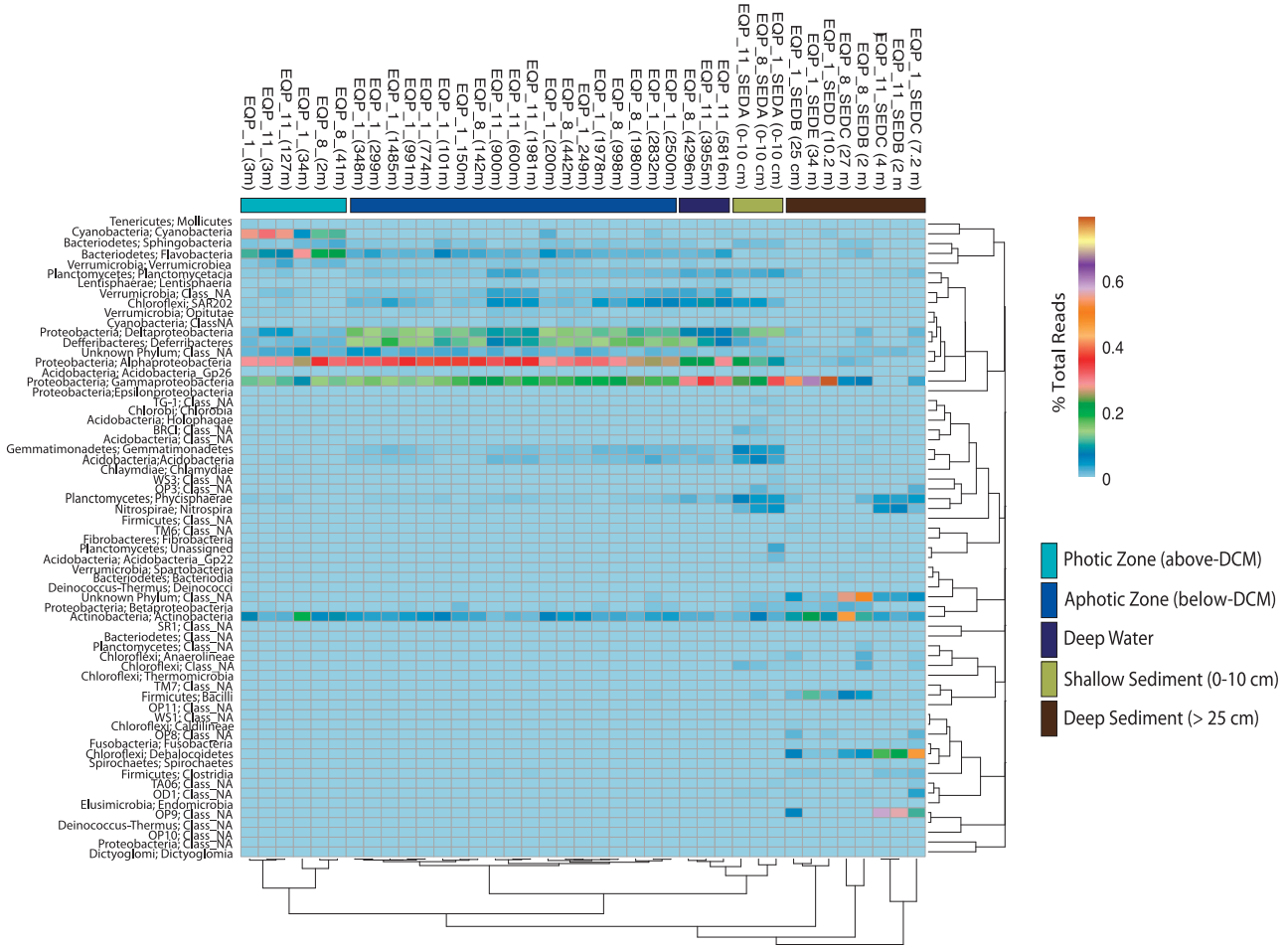

Figure 4 Heatmap clustering visualization of water column and sediment bacterial communities (x axis) and class-level GAST taxonomy (y axis). The heatmap is color coded by the percentage contributed by a given taxon to any given sample. Hierarchal clustering is based on the Morista–Horn similarity index.

Regardless of site, aphotic-zone community composition exhibits a consistent relationship to water depth, with a clear compositional gradation from epipelagic communities below the DCM to abyssopelagic communities at depths between 3500 and 6000 mbsl (Figure 3b). The photic-zone samples cluster by site, whereas the aphotic-zone samples vary less from site to site. Water-column samples decline in similarity with increasing water depth regardless of site location (Figure 3b).

We also observed two distinct groups in the sediment: a near seafloor (0–10 cmbsf) community and a subseafloor (0.25–34 mbsf) community. The common members of the shallow sedimentary community overlap with those of the deep-water community. The shallow sedimentary community is dominated by Alphaproteobacteria, Gammaproteobacteria and Deltaproteobacteria, as well as *Nitrospira* and Planctomycetes (Figure 4). The deep sedimentary community primarily comprises members of the obsidian pool (OP) candidate phylum, including OP9, OP8 and OP3, as well as Chloroflexi (Dehalococcoidetes and Anaerolineae), Actinobacteria and another phylum of unknown lineage (Figure 4).

The rank abundance and distribution of shared OTUs

To examine the extent to which OTUs are shared between the previously described communities (photic, aphotic, shallow sediment and deep sediment) (Figure 3a), we generated rank abundance histograms of the 100 most abundant OTUs (97% similarity) in each community. Figures 5a and c highlight the OTUs shared between the water column (photic and aphotic combined) and shallow sediment communities, between shallow and deep ($\geqslant$ 1.5 mbsf) sediment communities, and between the deep sedimentary and water-column communities. Although our clustering results binned an EQP-1 sample from 0.25 mbsf with the deep sedimentary community (Figure 4), we excluded this sample from this shared-OTU analysis because it was taken close to the seafloor and we wish to ensure that only OTUs truly selected for subseafloor conditions are compared.

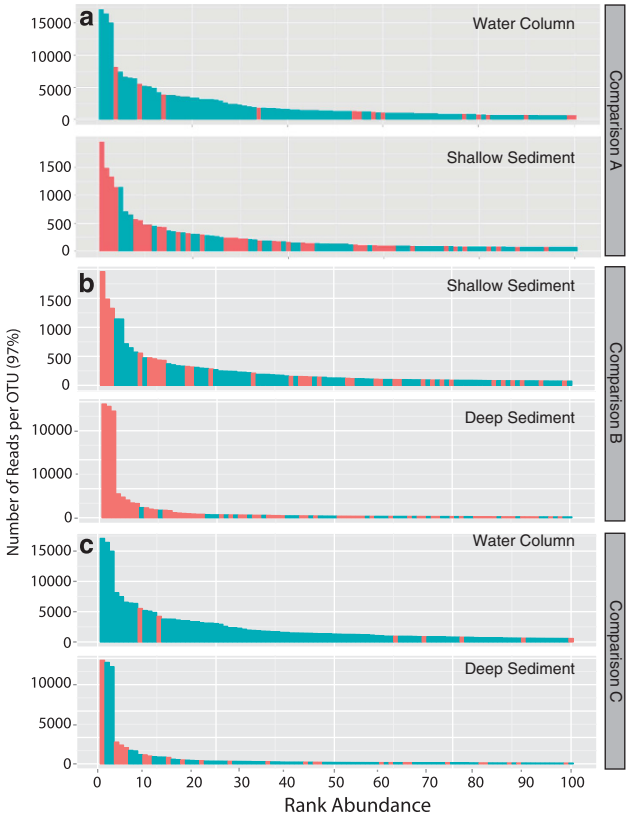

Figure 5 Rank abundance histograms of community overlap between (**a**) water column and shallow sediment (0–10 cmbsf), (**b**) shallow sediment and deep sediment (≥1.5 mbsl) and (**c**) water column and deep sediment (≥1.5 mbsl). OTUs highlighted in pink are shared between compared environments, whereas OTUs highlighted in blue are unique to each environment.

Table 2 Shared OTUs in the 100 most abundant OTUs found in the pooled data from each environment (water column, shallow sediment and deep sediment)

Top 100 OTUs	*Number of OTUs/% reads that overlap with:*		
Base data set	*Water column*	*Shallow sediment*	*Deep sediment*
Water column	– / –	14/11.9%	7/5.8%
Shallow sediment	41/56.3%	– / –	33/40.6%
Deep sediment	24/37.0%	62/87.8%	– / –

Abbreviation: OTU, operational taxonomic unit.
For the environment shown in each row, each pairwise comparison shows the number of shared OTUs and the percent of total reads in the top 100 OTUs comprised by those shared OTUs.

Many of the abundant OTUs in our shallow and deep sedimentary communities are present in the water at trace levels (<0.025% of total reads) (Table 2, Figures 5a and c). To quantify the statistical significance of so many shared OTUs being rare in the water column, we randomly resampled with replacement the water-column OTU frequency data until 24 unique OTUs were selected. These 24 randomly selected OTUs represent the 24 OTUs in the 100 most abundant deep-sediment OTUs that are

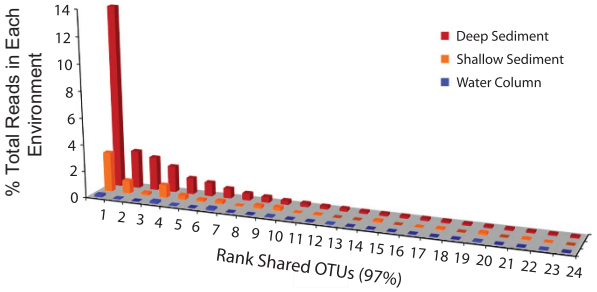

Figure 6 Rank abundance histograms for the top OTUs shared (x axis) between the oceanic and sedimentary environments. The y axis indicates the percentage of the total number of reads per OTU within our subseafloor sediment samples (≥1.5 mbsf, red), shallow sediment samples (orange) and water-column samples (blue).

shared with the water column (pink bars in lower panel of Figure 5c). Of these 24 shared OTUs, 22 are rare in the water column. Given the water-column abundance distribution and 5000 repetitions of our random resampling procedure, the probability of randomly selecting 22 of 24 OTUs from the pool of OTUs rarer than the top 100 in the water column is vanishingly small ($P<0.0001$).

These shared OTUs belong to Dehalococcoidetes, OP9, OP8 and other taxa. Furthermore, the OTUs that are abundant in the deep (subseafloor) sediment are also often abundant in the shallow sediment (Figure 5b). In contrast, most of the OTUs abundant in the water column are not present in our subseafloor sediment tags. The most abundant water-column OTUs that appear in our deep sedimentary tags include members of the Actinobacteria, Planctomycetes and Firmicutes (Figure 4). In short, the individual OTUs that are abundant in the deep sediment are also often abundant in the shallow sediment and often present but very rare (<0.025% of the tag population) in the overlying water column (Figure 6). The exceedingly small probability of getting so many rare water-column OTUs (22 of 24) by randomly sampling the water column ($P<0.0001$) indicates that water-column OTUs surviving in subseafloor sediment are preferentially from the rare community in the water column.

Taxonomic richness, evenness and PD

Vertical trends in bacterial diversity (# of observed OTUs, Chao-1, ACE and PD) and evenness (H') are similar at all three sites, with all measurements of richness and evenness peaking within the mesopelagic zone (Table 1). Overall, diversity is highest in the water column and lowest in the deep subseafloor sediment (Figure 2a). Within the water column, the vertical profile of bacterial diversity mirrors the O_2 profile and matches the profile of apparent oxygen utilization at each site (Figures 2a and b, Table 1). Diversity in the water column is generally lowest in the surface ocean and highest at the O_2 minima (Figures 2a and b). In the sediment, richness is

highest near the seafloor (Figure 2a). Bacterial diversity in the water column varies from site to site; for example, richness is highest in the oxygen minimum zone of North Pacific gyre site EQP-11 and much lower in the oxygen minimum zone of eastern equatorial upwelling Site EQP-1 (Figure 2a).

Discussion

Vertical trends in bacterial diversity

Water-column bacterial communities at all sites exhibit a clear depth profile, with samples declining in similarity from the seasurface to the abyssopelagic waters, in accordance with previously described trends (for example, DeLong *et al.*, 2006; Brown *et al.*, 2009; Treusch *et al.*, 2009). Although changes in community composition with depth have previously been well documented, studies of taxonomic richness with depth have yielded mixed results. Some studies have reported declines in taxonomic richness and/or PD with increasing ocean depth (Brown *et al.*, 2009; Agogué *et al.*, 2011; Bryant *et al.*, 2012), whereas others have observed higher richness deeper in the water column (for example, Pommier *et al.*, 2010, Kembel *et al.*, 2011; Ghiglione *et al.*, 2012). By sampling the water column more comprehensively than previous studies, we find that bacterial richness (taxonomic richness, PD and Chao-1) and evenness in the water column is lowest at the surface and highest within the O_2 minima of the mesopelagic zone at each of our sites. These results are consistent from site to site regardless of the diversity metric applied. Consequently, they provide clear evidence of vertical changes in bacterial richness across a range of environments. Some previous studies also reported elevated bacterial richness within the mesopelagic zone (Kembel *et al.*, 2011; Jing *et al.*, 2013; Ladau *et al.*, 2013), but did not explicitly associate this result with the O_2 minimum zone. The relationship we observe between oxygen and taxonomic richness suggests that respiration may have a significant role in shaping bacterial richness in the open ocean.

Only a few studies have examined bacterial diversity in both the water column and marine sediment (Feng *et al.*, 2009; Zinger *et al.*, 2011; Hamdan *et al.*, 2013). None of these studies included sediment deeper than 1.1 mbsf. Most previous studies of seafloor bacterial communities agree that those communities are highly diverse (Ravenschlag *et al.*, 1999; Madrid *et al.*, 2001; Luna *et al.*, 2004; Hewson *et al.*, 2007; Schauer *et al.*, 2010). Some studies have suggested that seafloor communities are taxonomically richer than communities in the water column (Lozupone and Knight, 2007; Feng *et al.*, 2009; Zinger *et al.*, 2011). In our study, taxonomic richness was generally higher in the water and lower in the sediment. This difference from previous studies may be partly due to the absence of subseafloor (>1 mbsf) sedimentary communities in those studies (>1 mbsf communities at our sites tend to be consistently less diverse than seafloor communities). It may also be due to the different oceanographic context of our studies relative to some previous studies (for example, Feng *et al.*, 2009). It is possible that the open-ocean seafloor sediment is less taxonomically rich than seafloor sediment of continental margins and estuaries. This difference may also partly result from different sampling approaches. Our study and Feng *et al.* (2009) sample both seawater communities and sediment communities at each location, while Luzopone and Knight (2007) and Zinger *et al.* (2011) compared seawater communities from diverse locations with sediment communities from very different locations.

Within the sediment at all of our sites, taxonomic richness is highest at the seafloor and declines with increasing sediment depth (Figure 2). There is extensive taxonomic overlap between the communities of the deep sediment and those of the shallow sediment (Figure 5). These results suggest that communities in deep marine sediment are derived from a subset of the community that lived in the sediment when it was near the seafloor.

Taxonomic overlap between the ocean and sediment

Zinger *et al.* (2011) examined the bacterial beta-diversity of seafloor and seawater ecosystems and demonstrated that pelagic communities and epibenthic communities differed greatly at all taxonomic levels. Our study supports this observation and extends it to subseafloor sediment, as the abundance-weighted bacterial community compositions of the sediment are clearly distinct from those in the overlying water column (Figure 3).

To examine the extent of taxonomic overlap between the ocean communities and the sediment communities, we separately pooled the DNA sequences from the water samples of all three sites and the DNA sequences from the sediment samples of all three sites. We pooled the data for these comparisons, rather than doing the comparisons site by site, because sinking particles may be advected large distances, and in different directions at different water depths, before settling to the seafloor. Consequently, the material that settles to the seafloor rarely originates in the immediately overlying water. Although our three sites are separated by thousands of kilometers, spanning over 70 degrees of longitude and 30 degrees of latitude, we think it is justifiable and necessary to pool the samples because (i) intense sequencing at a single site has demonstrated that the marine bacterial rare biosphere is highly cosmopolitan (Gibbons *et al.*, 2013) and (ii) pelagic communities from the same oceanographic horizon (the DCM) are highly similar for thousands of kilometers (Walsh *et al.*, 2015).

When we consider the presence–absence of individual taxa, we observe overlap between environments. The most prominent example of this overlap

is the presence in seawater and shallow sediment of many OTUs that are abundant in the deep-sediment samples (Figures 5 and 6). The overlap is very lopsided considering relative abundance; OTUs represented by abundant sequences in the sediment tag populations are often present but represented by rare sequences in the seawater tag population, whereas OTUs with abundant tags in the seawater population are rarely present in the sediment populations. This result is consistent with microbes being carried through the ocean and deposited in the sediment. The absence of known phototrophs (for example, *Prochlorococcus*, *Synechococcus*) in our subseafloor tag populations indicates that taxonomic overlap between the seawater and subseafloor sedimentary communities does not simply result from preservation of detritus. Instead, the subseafloor conditions appear to select for preferential survival of taxa that were represented by relatively few sequences in the seawater populations. The relatively low abundance of these sequences in the seawater populations is not surprising, as subseafloor sedimentary conditions are very different from water-column conditions; for example, anoxic conditions and low energy availability are much more ubiquitous in subseafloor sediment than in the overlying ocean (Jørgensen and D'Hondt, 2006). It is possible that microbes associated with the sedimentary environment are transported long distances through the ocean as components of the rare biosphere, from which they may or may not recolonize the seafloor. Although resuspension and downslope sediment transport may be important sedimentary depositional processes in some marine environments, they are not significant processes at our sites; all three sites are located on flat abyssal regions or topographic highs and the cores from all three sites show no evidence of turbidites, indicating little to no redeposition of sediment. Consequently, advective transport through the water column provides the simplest explanation of microbial input to this sediment, consistent with discovery of thermophiles in marine sediment far from hydrothermal vents (Lee *et al.*, 2005; Hubert *et al.*, 2009), and with Hamdan *et al.*'s (2013) identification of hydrography as a forcing factor for OTU composition in near-seafloor marine sediment.

Microbial transport through the water column may explain the incidence of large-scale biogeographic patterns in both near-seafloor sediment (Hamdan *et al.*, 2013) and subseafloor sediment. A previous study (Inagaki *et al.*, 2006) observed that similar taxa (90% similarity, for example, Dehalococcoidetes, OP9 and Desulfobacteriales) are present in environmentally similar subseafloor sedimentary environments separated by large geographic distances. Our results show this similarity at a much finer taxonomic resolution (97% similarity) and indicate that these clades may dominate in a broad range of subseafloor sedimentary environments. Transport through the water column is far easier than through a sedimentary matrix where both advection and biologically accessible electron donors are limited.

Conclusions

Our study provides the first vertical profiles of microbial diversity from seasurface to subseafloor sediment (down to 34 mbsf). Bacterial richness is generally higher in the water column than in the sediment. It is highest in the O_2-minimum zone of the water column and lowest in the subseafloor sediment. Despite the differences in abundance-weighted community compositions in the water column, the shallow sediment and the subseafloor sediment, the most abundant taxa in the subseafloor sedimentary communities are also often abundant in the shallow sedimentary communities and present but rare in the water column. This result suggests that deep subseafloor communities (i) constitute a subset of the diverse taxa that were present in the sediment at the time of burial and (ii) are ultimately seeded via the water column.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

We thank Captain Kent Sheasley, and the crew and scientific party of *R/V Knorr* Expedition 195-3 for our samples. We also thank Robert Pockalny for mapping and helping to select the Expedition 195-3 sampling sites. This study was funded by the Biological Oceanography Program of the US National Science Foundation (grant OCE--0752336) and by the NSF-funded Center for Dark Energy Biosphere Investigations (grant NSF-OCE-0939564). This is C-DEBI publication 269.

References

Agogué H, Lamy D, Neal PR, Sogin ML, Herdnel GJ. (2011). Water mass-specificity of bacterial communities in the North Atlantic revealed by massively parallel sequencing. *Mol Ecol* **20**: 258–274.

Baas-Becking LGM. (1934). *Geobiologie of Inleiding tot de Milieukund*. WP Van Stockum and Zoom (in Dutch): The Hague, The Netherlands.

Behrenfeld MJ, Falkowski PG. (1997). Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol Oceanogr* **42**: 1–20.

Beman JM, Carolan MT. (2013). Deoxygenation alters bacterial diversity and community composition in the ocean's largest oxygen minimum zone. *Nat Commun* **4**: 2705.

Brown MV, Philip GK, Bunge JA, Smith MC, Bissett A, Lauro FM *et al.* (2009). Microbial community structure in the North Pacific Ocean. *ISME J* **3**: 1374–1386.

Bryant JA, Stewart FJ, Eppley JM, DeLong EF. (2012). Microbial community phylogenetic and trait diversity

declines with depth in a marine oxygen minimum zone. *Ecology* **93**: 1659–1673.
Cai P, Huang Q, Zhang X, Chen H. (2006). Adsorption of DNA on clay minerals and various colloidal particles from an Alfisol. *Soil Biol Biochem* **38**: 471–476.
Caporaso GJ, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK *et al.* (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**: 335–336.
Clarke KR, Gorley RN. (2006). PRIMER v6: User Manual/ Tutorial.: PRIMER-E, Plymouth.
D'Hondt S, Jørgensen BB, Miller DJ, Batzke A, Blake R, Cragg BA *et al.* (2004). Distributions of microbial activities in deep subseafloor sediments. *Science* **306**: 2216–2221.
D'Hondt S, Inagaki F, Zarikian CA, Abrams LJ, Dubois N, Engelhardt T *et al.* (2015). Presence of oxygen and aerobic communities from sea floor to basement in deep-sea sediments. *Nature Geosci* **8**: 299–304.
D'Hondt S, Spivack A, Pockalny R, Ferdleman T, Fisher J, Kallmeyer J *et al.* (2009). Subseafloor sedimentary life in the South Pacific Gyre. *Proc Natl Acad Sci USA* **106**: 11651–11656.
DeLong EF, Preston CM, Mincer T, Rich V, Hallam SJ, Frigaard NU *et al.* (2006). Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496–503.
Divins DL. (2003). Total Sediment Thickness of the World's Oceans & Marginal Seas, vol. Boulder, CO.
Feng B-W, Li X-R, Wang J-H, Hu Z-Y, Meng H, Xiang L-Y *et al.* (2009). Bacterial diversity of water and sediment in the Changjiang estuary and coastal area of the East China Sea. *FEMS Microbiol Ecol* **70**: 236–248.
Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Nat Acad Sci USA* **103**: 13104–13109.
Ghiglione J-F, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K *et al.* (2012). Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc Nat Acad Sci USA* **109**: 17633–17638.
Gibbons SM, Caporosa JG, Pirrung M, Field D, Knight R, Gilbert JA. (2013). Evidence for a persistent microbial seed bank throughout the global ocean. *Proc Nat Acad Sci USA* **110**: 4651–4655.
Gregg WW, Casey NW, McClain CR. (2005). Recent trends in global ocean chlorophyll. *Geophys Res Lett* **32**: L03606.
Hamdan LJ, Coffin RB, Sikaroodi M, Greinert J, Treude T, Gillevet PM. (2013). Ocean currents shape the microbiome of Arctic marine sediments. *ISME J* **7**: 685–696.
Hewson I, Jacobson/Meyers ME, Fuhrman JA. (2007). Diversity and biogeography of bacterial assemblages in surface sediments across the San Pedro Basin, Southern California Borderlands. *Environ Microbiol* **9**: 923–933.
Hubert C, Loy A, Nickel M, Arnosti C, Baranyi C, Brüchert V *et al.* (2009). A constant flux of diverse thermophilic bacteria into the Cold Arctic Seabed. *Science* **325**: 1541–1544.
Huse SM, Huber JA, Morrison HG, Sogin ML, Welch DM. (2007). Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biol* **8**: R143.
Inagaki F, Nunoura T, Nakagawa S, Teske A, Lever M, Lauer A *et al.* (2006). Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proc Nat Acad Sci USA* **103**: 2815–2820.
Jing H, Xia X, Suzuki K, Liu H. (2013). Vertical profiles of bacteria in the tropical and subarctic oceans revealed by pyrosequencing. *PLoS One* **8**: e79423.
Jørgensen BB, D'Hondt S. (2006). A starving majority deep beneath the seafloor. *Science* **314**: 932–934.
Kallmeyer J, Pockalny R, Adhikari RR, Smith DC, D'Hondt S. (2012). Global distribution of microbial abundance and biomass in subseafloor sediment. *Proc Nat Acad Sci USA* **109**: 16213–16216.
Kembel SW, Eisen JA, Pollard KS, Green JL. (2011). The phylogenetic diversity of metagenomes. *PLoS One* **6**: e23214.
Ladau J, Sharpton TJ, Finucane MM, Jospin G, Kembel SW, O'Dwyer J *et al.* (2013). Global marine bacterial diversity peaks at high latitudes in winter. *ISME J* **7**: 1669–1677.
Lee Y-J, Wagner ID, Brice ME, Kevbrin VV, Mills GL, Romanek CS, Wiegel J. (2005). Thermosediminibacter oceani gen. nov., sp. nov. and Thermosediminibacter litoriperuensis sp. nov., new anaerobic thermophilic bacteria isolated from Peru Margin. *Extremophiles* **9**: 375–383.
Lozupone CA, Knight R. (2007). Global patterns in bacterial diversity. *Proc Nat Acad Sci USA* **104**: 11436–11440.
Luna GM, Dell'Anno A, Giuliano L, Danovaro R. (2004). Bacterial diversity in deep Mediterranean sediments: relationship with the active bacterial fraction and substrate availability. *Environ Microbiol* **6**: 745–753.
Madrid VM, Aller JY, Aller RC, Chistoserdov AY. (2001). High prokaryote diversity and analysis of community structure in mobile mud deposits off French Guiana: identification of two new bacterial candidate divisions. *FEMS Microbiol Ecol* **37**: 197–209.
Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL *et al.* (2006). Microbial biogeography: putting microorganisms on the map. *Nat Rev Micro* **4**: 102–112.
Müller RD, Sdrolias M, Gaina C, Roest WR. (2008). Age, spreading rates, and spreading asymmetry of the world's ocean crust. *Geochem Geophys Geosyst* **9**: Q04006.
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.
R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, AustriaURL http://www.R-project.org/.
Ravenschlag K, Sahm K, Pernthaler J, Amann R. (1999). High bacterial diversity in permanently cold marine sediments. *Appl Environ Microbiol* **65**: 3982–3989.
Røy H, Kallmeyer J, Adhikari RR, Pockalny R, Jørgensen BB, D'Hondt S. (2012). Aerobic microbial respiration in 86-million year old deep-sea red clay. *Science* **336**: 922–925.
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.
Sogin M, Morrison H, Welch D, Huse S, Neal P, Arrieta J *et al.* (2006). Microbial diversity in the deep-sea and

the unexplored 'rare Biosphere'. *Proc Nat Acad Sci USA* **103**: 12115–12120.

Treusch AH, Vergin KL, Finlay LA, Donatz MG, Burton RM, Carlson CA *et al.* (2009). Seasonality and vertical structure of microbial communities in an ocean gyre. *ISME J* **3**: 1148–1163.

Walsh EA, Smith DC, Sogin ML, D'Hondt SL. (2015). Bacterial and archaeal biogeography of the deep chlorophyll maximum in the South Pacific Gyre. *Aquatic Microbial Ecology* **75**: 1–13. doi:10.3354/ame01746.

Wickham H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer: New York.

Zinger L, Amaral-Zettler LA, Furhman JA, Horner-Devine MC, Huse SM, Welch DBM *et al.* (2011). Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. *PLoS One* **6**: e24570.