Microbial diversity of oligotrophic marine sediments

Alan Durbin

A thesis submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in the Department of Marine Sciences

Chapel Hill
2009

Abstract

Alan Durbin
Microbial diversity of oligotrophic marine sediments
(Under the direction of Andreas Teske)

Oligotrophic marine sediments cover nearly 90% of the ocean's floor, yet in contrast to the continental margins, their biogeochemistry and microbiology has been little explored. Oligotrophic sediments exhibit dramatic nutrient limitation due to the slow rain rate of photosynthetic carbon, and consequently display a greatly expanded zone of higher-energy electron accepting activities relative to margin sediments. The working hypothesis that the most-energetic available electron acceptor structures subsurface microbial communities predicts that different microbial lineages occur in oligotrophic vs. organic-rich margin sediments. This study examines this hypothesis by considering the archaeal diversity, as revealed in clone libraries, for an ultraoligotrophic, fully oxic site, as well as the archaeal and bacterial diversity of an oligotrophic, oxic/suboxic site. Finally, I compare the archaeal diversity of available oligotrophic and representative continental margin sites, and find evidence for oligotrophic sediments as hosting a distinct, deeply divergent assemblage of Archaea, possibly adapted to exploit the higher-energy electron acceptors present therein.

Acknowledgements

I would like to acknowledge the members of the crew and shipboard scientific party of cruise Knox02RR, and in particular I thank Jens Kallmeyer, Heather Schrum, Art Spivack, Jan Fischer, and Tim Ferdelman for providing unpublished data from this cruise. The help and advice in the laboratory and in writing afforded me by graduate students in Marine Sciences has been invaluable. I would like to thank my advisor Andreas Teske and committee members Marc Alperin and Barbara MacGregor for their guidance and support during the research and writing that went in to this thesis. Finally I thank my friends and family for their essential support and encouragement, and in particular my sister Sarah Durbin for helping editing and proofreading this manuscript.

Table of Contents

List of Tables

List of Figures

List of Abbreviations and Symbols

α……MG-I subclade Alpha

β……MG-I subclade Beta

γ……MG-I subclade Gamma

δ……MG-I subclade Delta

ι……MG-I subclade Iota

ε……MG-I subclade Epsilon

η……MG-I subclade Eta

κ……MG-I subclade Kappa

θ……MG-I subclade Theta

υ……MG-I subclade Upsilon

ζ……MG-I subclade Zeta

AAG……Ancient Archaeal Group

BSA……Bovine Serum Albumin

DHVE-3……Deep Hydrothermal Vent Euryarchaeota 3

DHVE-4……Deep Hydrothermal Vent Euryarchaeota 4

DHVE-5……Deep Hydrothermal Vent Euryarchaeota 5

DHVE-6……Deep Hydrothermal Vent Euryarchaeota 6

DHVE-8……Deep Hydrothermal Vent Euryarchaeota 8

DHVEG-II……Deep-sea Hydrothermal Vent Euryarchaeotal Group II

DSAG……Deep Sea Archaeal Group

DSEG-1……Deep Sea Euryarchaeotal Group 1

DSEG-2……Deep Sea Euryarchaeotal Group 2

DSEG-3……Deep Sea Euryarchaeotal Group 3

DSEG-4……Deep Sea Euryarchaeotal Group 4

HWCG-I……Hot-Water Crenarchaeotal Group I

HWCG-IIa……Hot-Water Crenarchaeotal Group IIa

HWCG-IIb……..…….....……Hot-Water Crenarchaeotal Group IIb

HWCG-III…….…….……Hot-Water Crenarchaeotal Group III

LMP…….…….……..…..Low Melting Point agarose

LPA…….…….……...…….Linear Polyacrylamide

MBG-A.…….……...……Marine Benthic Group A

MBG-B……..……..Marine Benthic Group B

MBG-D……..Marine Benthic Group D

MBG-E…….……..……..Marine Benthic Group E

MCG.…….……...……Miscellaneous Crenarchaeotal Group

MG-I.....Marine Group I Crenarchaeota

MG-II.....Marine Group II Euryarchaeota

MG-III…….……...…….Marine Group III Euryarchaeota

MG-IV…….……..….Marine Group IV Euryarchaeota

MG-V…….……...…..Marine Group V Euryarchaeota

ML…….……..……Maximum Likelihood

ODP……..Ocean Drilling Program

OTU…….……Operational Taxonomic Unit

SAGMEG.…….…….South African Gold Mine Euryarchaeotal Group

SPG10…….…….……..…….South Pacific Gyre Site 10

SPG11…….…….……..……South Pacific Gyre Site 11

SPG12…….……..South Pacific Gyre Site 12

ThetaYC.....Yue and Clayton's (2005) similarity index

TMEG…….……..……..Terrestrial Miscellaneous Euryarchaeotal Group

Microbial Diversity of Oligotrophic Marine Sediments

Introduction.

Marine sediments constitute one of the largest biomes, by area and volume, on planet Earth, and have been shown to host microbial communities that may account for $1/10^{th}$ to $1/3^{rd}$ of global microbial biomass (Parkes et al. 1994, Whitman et al. 1998). Organic carbon hundreds of thousands to millions of years old sustains active microbial communities (e.g.,Wellsbury et al. 2002, D'Hondt et al. 2004, Parkes et al.2005, Biddle et al. 2006) with extraordinarily slow community turnover rates of centuries or more (Biddle et al. 2006). Studies of deep marine sediments have so far focused on continental margin sediments, where the bulk of carbon burial (Dunne et al. 2007) and large-scale production and consumption of methane (D'Hondt et al. 2002, 2004) occur. Here sulfate, a relatively low-energy electron acceptor, is consumed slowly over a scale of meters to hundreds of meters, compared to centimeters to meters in the estuarine sediments, due to lower organic substrate input rates. Beneath the sulfate reduction zone only the least energetically favorable electron acceptors persist, such as CO_2 or oxidized portions of organic molecules. Thus, as a result of electron donor limitation, availability of only low-energy electron acceptors, or both, the continental margin deep subsurface exhibits strong energy limitation.

Although the continental margins have been the principal focus of subsurface microbiological and geochemical studies to date, abyssal sediments >2000m water depth cover a much larger extent of the ocean floor (~89%; Dunne et al. 2007). Despite accounting for 80% of global surface primary production, the abyssal ocean accounts for only 13.5% of deposition and 2% of burial. While 13.5% is a nontrivial amount in terms of balancing geochemical budgets, 2% is nearly negligible; this signals that while a significant proportion of global sedimentary remineralization occurs in oligotrophic sediments, only a very tiny fraction escapes microbially mediated oxidation. Conversely, this means that microbial communities in oligotrophic sediments are under extreme electron donor limitation, causing extreme energy limitation and contributing to the lowest cell densities yet observed (D'Hondt et al. 2009).

This energy limitation occurs despite the advantages offered by high-energy oxidants: the slow flux of carbon substrates means high-energy electron acceptors are available over a much wider sediment depth range in the oligotrophic seafloor compared to the continental shelf. Although in broad view, continental margin sediments can be regarded as anoxic, often there are spatially compressed, millimeter or centimeter-scale oxic/suboxic zones at the sediment-water interface (where suboxic is defined by the absence of oxygen and lack of sulfide, indicating minimal net sulfate reduction). At least in the case of oxygen, the expanse of these strata principally determines burial efficiency in organic-rich sediments (Hartnett et al. 1998). The greatly expanded oxic or suboxic zones of oligotrophic sediments may be more easily-accessible analogues of the compressed oxic and suboxic zones in organic-rich sediments. Alternatively, the combination of low electron donor availability and high sediment oxidation state may pose unique challenges for microbial life, shaping a fundamentally different habitat in oligotrophic sediments. Lower organic carbon concentrations may come in the form of the least preferable, most recalcitrant molecules (Wakeham et al. 1997, Hedges et al. 2001, Lee et al. 2004), and may be more inaccessible due to adsorption to minerals (lower organic carbon production means a lower carbon/mineral ratio; Dunne et al. 2007). Oligotrophic suboxic or oxic sediments may be fundamentally different from oxic/suboxic strata atop eutrophic sediments due to the interactive effects of different variables, such as the combination of high sediment oxidation state and low electron donor availability, leading to higher biosynthesis costs as a proportion of cellular energy budget (McCollom and Amend 2005). Finally, different suboxic redox processes are possible in neutral pH, non-sulfidic oligotrophic sediments than in suboxic acidic and/or sulfidic margin sediments (Schippers and Jørgensen 2001, 2002, Severmann et al. 2006).

The dominant uncultured archaeal lineages identified in the marine subsurface to date (e.g., Teske and Sørensen 2008) are principally known from environments with relatively low-energy electron acceptors, such as sulfur, sulfate, and CO_2. It is not clear whether the same groups that have been identified as dominant subsurface lineages also persist in the more extensive habitat of oligotrophic marine sediments, which are geochemically distinct. Since microbes that specialize in using the most energetic available electron acceptor are hypothesized to outcompete other subsurface microbes, the most-energetic available electron acceptor may play a key role in structuring microbial communities (e.g. Inagaki et al. 2006). This

expectation is examined in Chapter 1 of this work, a comparison of the archaeal diversity (as revealed in 16S rRNA clone libraries) of an oxic, ultraoligotrophic sediment column at three near surface sediment depths and the overlying water sample from a multicore; this study examines potential gradients in microbial diversity between the sediment and water column, and within the sediment itself, when the same electron acceptor (oxygen) is shared throughout. Chapter 2 looks at the archaeal and bacterial 16S rRNA gene diversity of an oligotrophic shallow sediment column with oxic and suboxic strata, in conjunction with porewater chemical data, the first study to do so at high resolution in oligotrophic sediments. Finally, Chapter 3 examines the oligotrophic seafloor as a potential distinct microbial habitat, compared to the better-characterized margin environments, and explores the association of sediment trophic state and the appearance of different deeply branching, uncultured archaeal lineages.

Chapter 1: Sediment-Associated Microdiversity within the Marine Group I Crenarchaeota

Introduction

Most microbiological studies of deep subsurface sediments have focused on organic-rich, reduced and anoxic continental margin and slope sediments; these are geochemically and microbiologically distinct from organic-poor, oxidized sediments of the open ocean (e.g., D'Hondt et al. 2004). The novel, deeply branching archaeal lineages found in 16S rRNA clone libraries from anoxic marine subsurface sediments (e.g., Vetriani et al. 1999, Inagaki et al. 2003, Parkes et al. 2005, Sørensen and Teske 2006, Biddle et al. 2006, Inagaki et al. 2006) differ from those in the oxic water column (e.g., DeLong et al., 1992; Fuhrman et al. 1992; DeLong et al. 1994; Fuhrman and Davis 1997; Massana et al. 2000; Bano et al. 2004). Thus, the deep marine sedimentary subsurface is regarded as a distinct biome with a specialized, predominantly heterotrophic microbial community (Biddle et al. 2006, Lipp et al. 2008) and with a distinct genetic repertoire (Biddle et al. 2008). However, this emerging picture of the deep marine sedimentary community is strongly biased towards datasets from organic-rich sediments. Oligotrophic marine subsurface sediments are massively underrepresented in current subsurface studies (Sørensen et al. 2004; Teske and Sørensen 2008; Fry et al. 2008) and are therefore of great interest for microbial community analyses.

The depth distribution of electron acceptors in the subsurface is hypothesized to control the depth profile and phylogenetic composition of microbial communities (Teske and Sørensen, 2008), since microorganisms that specialize in the highest-energy electron acceptors available have the energetic resources to outmultiply and outcompete others. In organic-rich sediments with high rates of organic carbon input, microbial activity depletes high-energy electron acceptors such as oxygen and nitrate within a few millimiters or centimeters of the sediment-water interface; in oligotrophic ocean basins, the oxic zone extends to a scale of tens of centimeters or a meter (Wenzhöfer and Glud 2002). Ultra-oligotrophic sediments underlying the centers of subtropical gyres may be oxic throughout the entire sediment column, on a scale of tens of meters (Gieskes and Boulègue 1986). Thus, if electron acceptor is the dominant factor

in structuring marine sedimentary microbial communities, oxidized sediments could host the same microbial assemblage as the overlying oxic water column. To test the working hypothesis of overlapping microbial community structure in oxic water column and sediments, this study aims at a detailed comparison of aerobic, oligotrophic subsurface communities and aerobic water column communities from the South Pacific gyre.

The principal target microorganisms of this investigation are pelagic Crenarchaeota, the major archaeal assemblage in the marine water column, based on quantitative studies using abundance of crenarchaeotal lipids (Damsté et al. 2002) and FISH probes specific to Crenarchaeota (Karner et al. 2001, Church et al. 2003, Teira et al. 2006, Varela et al. 2008). The pelagic Crenarchaeota consist predominantly of the Marine Group I (MG-I) phylum, based on 16S rRNA clone libraries (DeLong et al., 1992; Fuhrman et al. 1992; DeLong et al. 1994; Fuhrman and Davis 1997; Massana et al. 2000, Bano et al. 2004), but also include other crenarchaeotal phyla (e.g., Mincer et al. 2007, Coolen et al 2007, Agogué et al. 2008). Two cultured representatives of the MG-I crenarchaeota exist: the pure culture isolate *Nitrosopumilus maritimus* (Könneke et al. 2005), and the candidate species *Cenarchaeum symbiosum* from a natural enrichment growing within the tissue of a marine sponge, *Axinella* sp. (Preston et al. 1996).

Here, we focused on different phylogenetic groups within the MG-I archaea, and identified groups that are specifically recovered from sediments, as opposed to water column MG-I archaea. Previous work has delineated several clades within the MG-I phylogeny, based on 16S rRNA gene sequences and bootstrap support for mutually exclusive clusters (Massana et al. 2000, Takai et al. 2004). Massana et al. (2000) adopted a Greek alphabetical nomenclature for the MG-I subphyla, which was followed by Takai et al. 2004 and expanded by Sørensen et al. (2004) and this study. Of these, clade MG-I α (alpha) is the best-represented in public databases and contains most of the water-column sequences, including *Nitrosopumilus maritimus*. The sponge-associated candidate species *Cenarchaeum symbiosum* is rather divergent from other MG-I 16S rRNA clades and appears in a separate clade β (beta) (Massana et al. 2000). The MG-I γ (gamma) group includes clones from ambient seawater at Indian and Pacific Ocean hydrothermal vents (Takai et al. 2004) and from the North Pacific mesopelagic water column (Mincer et al. 2007). The MG-I δ (delta) group is a smaller cluster that appears to group within the MG-I γ. Other clades encountered in oligotrophic sediments include the ε (epsilon), ζ (zeta), and η (eta) groups (Sørensen et al.

2004).

Results.

Geochemical and site description data.

In this study, we examined the phylogenetic structure of MG-I crenarchaeotal communities from two oligotrophic South Pacific gyre sites, one from within the gyre abyssal plain at 5076 m depth, at 41°51'S and 153°06'W (SPG11), and another from just outside the gyre proper at 5306 m depth and 45°58'S and 163°11'W (SPG12), in conjunction with available geochemical data. Both sediments represent the oligotrophic endmembers of microbial community studies. The SPG12 samples are from a gravity core. The SPG11 sediment samples were obtained from a multicorer push core; SPG11 supernatant water was taken from the same pushcore. Intact microstructures of very fine-grained clay sediment at the sediment-water interface, and undisturbed oxygen profiles determined *ex situ* on fresh multicores collected simultaneously with the cores sampled in this study, indicated an intact sediment/water interface (Fischer et al. 2009). Cores affected by sediment resuspension were easily recognized due to the extremely slow settling rate of fine clay sediments, and were not sampled.

Sedimentation rates and the resulting carbon flux to the sediment are approximate indicators of sediment trophic states (D'Hondt et al. 2002). All available geochemical data indicate that sites SPG 11 and 12 are strongly oligotrophic, SPG11 more so than SPG12. The sedimentation rates for SPG Sites 11 and 12 are 0.9m/My and 1.8m/My, respectively; they are 4-10 and 2-5 times lower than the rates of the slowest-accumulating site studied on ODP Leg 201, Peru Basin Site 1231 (D'Hondt et al. 2003; D'Hondt et al., 2009; Figure 1). The sediment DIC concentrations, indicators of microbial net metabolism, are lower at the SPG sites than at the oligotrophic Peru Basin Site 1231 (D'Hondt et al. 2003). For SPG11, alkalinity (96% of which is DIC at *in situ* pH) does not change from the overlying water values (Figure 1A). Total organic carbon (TOC), as determined by the subtraction of total inorganic carbon (TIC) from total carbon (TC) as a weight-percent of sediment, was 0.594% in the surface 1cm at SPG11, decreased to 0.49% at 2 cm depth, and then more gradually to 0.45% by 9 cm depth (D'Hondt et al. 2009). Values for SPG12 were not available at time of writing.

Most of the drawdown of dissolved oxygen at SPG11 occurs over the upper 10 cm of sediment.

c[O2] declines from 176.59 µM at 5 cmbsf to 160.95 µM at 25-30 cm depth, a 15.64 µM decrease (Figure 1A). In-situ dissolved oxygen profiles of bottom water at different SPG sites, including nearby SPG Site 10, revealed a bottom-water oxygen concentration of ~220 µM across the SPG.

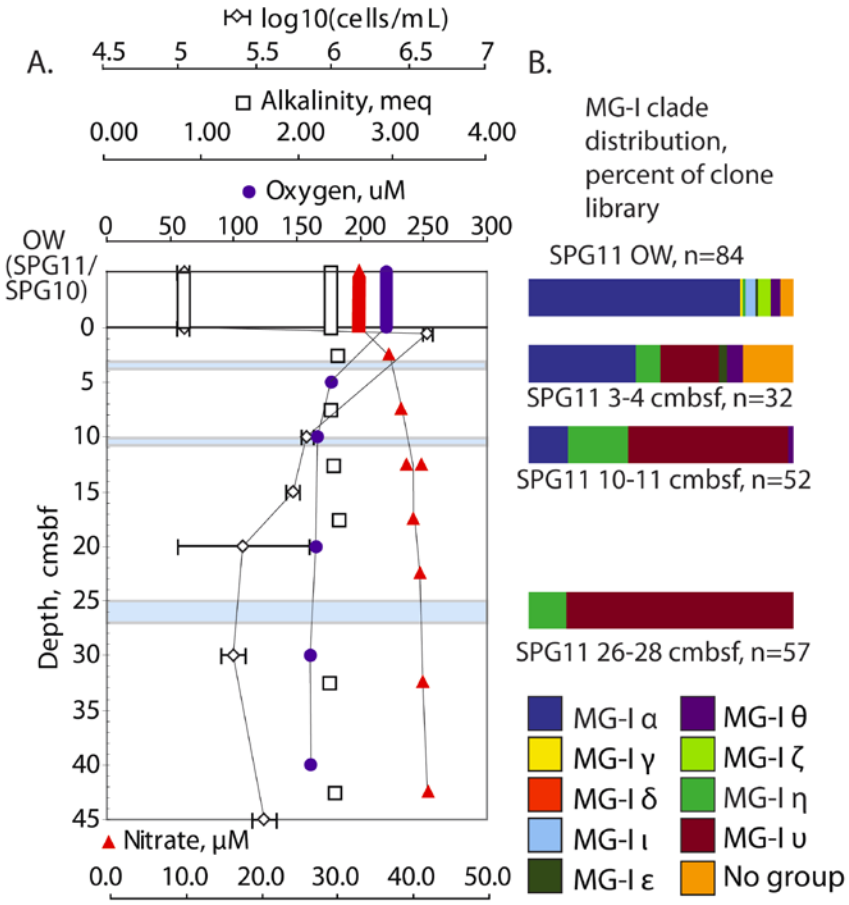


Figure 1. Depth trends in porewater constituents, cell densities and MG-I subclade abundance at SPG11. (A) Nitrate, alkalinity, and oxygen data (Fischer et al. 2009, D'Hondt et al. 2009), as well as cell counts (D'Hondt et al. 2009). Values from the overlying water are the topmost values shown, at 0 cmbsf. (B) Abundance of clones belonging to different well-supported MG-I clades, with sampling interval indicated by light blue bars.

Thus, the overall drawdown of oxygen over the upper ~30 cm depth at SPG11 is ~60 µM, after which little change is observed (Figure 1A). Oxygen appears not to be depleted throughout the ~70m sediment column (Fischer et al. 2009). The SPG11 nitrate concentration profile revealed a sharp increase in concentration over the upper ~30 cm of sediment, from 33.3 µM in the overlying water to 43.2 µM at 25-30 cmbsf (Figure 1A). The deepest interval sampled for nitrate from the multicore, 30-35 cmbsf, showed a slight decline to 41.7 µM, which is consistent with the deep sediment measurements made with the trigger core (D'Hondt et al. 2009).

At Site SPG12, the nitrate, oxygen, and alkalinity profiles indicate slightly more reducing conditions than at SPG 11 (Figure 2A). Oxygen declines from 114.76 µM at 0.05 cmbsf to ~2 µM at 179 cmbsf, while nitrate declines from 43.8 µM at 10-15 cmbsf to zero at 253-258 cmbsf, suggesting a successive depletion or near-depletion of oxygen followed by depletion of nitrate. No multicore or overlying water sample was taken for Site SPG12. The drawdown profile of oxygen appears exponential over the upper 100-150 cmbsf, while the buildup of DIC over the same interval appears linear until it reaches nearly its maximum value at around 100-150 cmbsf (Figure 2A). Taken together, rates of sedimentation, oxygen and nitrate depletion depths, and maximum DIC values indicate that SPG12 is significantly more oligotrophic than the most nutrient-limited marine sediments examined during ODP Leg 201 (D'Hondt et al. 2004).

Cell counts at site SPG11 revealed a peak in cell densities (4.27×10^6 cells/mL) at the sediment water interface, followed by a 20-fold decline to 2.29×10^5 cells/mL at 30 cmbsf (Figure 1A). A slightly higher cell density of 3.63×10^5 was detected at the next deepest interface, 45 cmbsf, and afterwards cell concentrations decline more slowly with depth, by less than an order of magnitude from 3.46×10^4 at 75 cmbsf to 7.41×10^3 cells/mL. Site SPG12 exhibited higher cell abundances than Site SPG11 (Figure 2A). The highest recorded densities at Site SPG12 occur not at the very surface but at 70-119 cmbsf; a gravity corer could have resulted in the loss of the more-active sediment-water interface. The subsurface peak in cell densities at ~1m depth may be associated with a redox interface between oxygen and reduced nitrogen species or metals, possibly supporting lithotrophic metabolisms, and coincides with the depth at which the

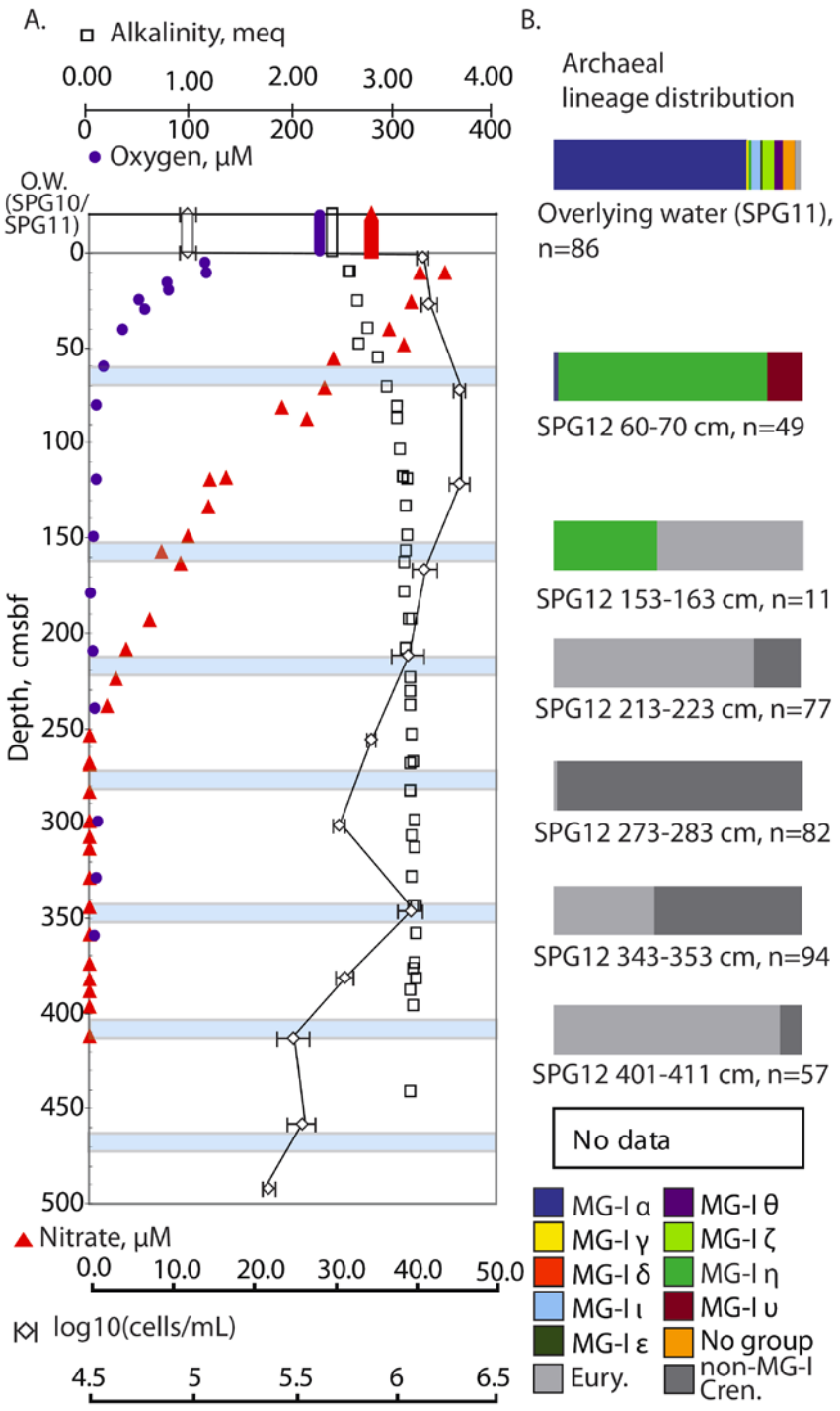


Figure 2. SPG12 depth trends in porewater constituents, cell densities and MG-I subclade abundance. (A) SPG12 nitrate, alkalinity, and oxygen data, as well as cell counts (Fischer et al. 2009, D'Hondt et al. 2009). 0 cmbsf values for cell counts and oxygen represent overlying water values from the SPG11 multicorer samples. (B) Relative abundance of MG-I subclades with depth at SPG12 with the SPG11 overlying water Archaeal clone library included for comparison. Clone libraries from deeper depths at SPG12 consisted solely of non-MG-I Crenarchaeota and Euryarchaeota, discussed elsewhere (Durbin, 2009). Light blue horizontal bars indicate sampling interval.

alkalinity maximum begins. Heterotrophic metabolism likely also plays an important role both through the dissimilatory oxidation of organic matter as well as by providing the reducing equivalents responsible for generating reduced metals or by production of ammonia or organic N. This is in contrast with Site SPG11, where geochemical profiles suggest dominance of aerobic respiration, and where electron donors are insufficient to spark redox cycling of NO_X or metals (Figure 1A, 2A).

Molecular survey results and phylogenetic analysis.

For Site SPG11, 226 nearly full length archaeal 16S rRNA gene clones derived from 3 different sediment horizons and from overlying water were sequenced and analyzed. For Site SPG12, 49 nearly full length archaeal 16S rRNA gene clones were obtained from the 60-70 cmbsf horizon and 11 from 153-163 cmbsf. All sediment-derived sequences and all but 2 water column-derived sequences collected at Site SPG11 grouped with the Marine Group I (MG-I) Crenarchaeota (DeLong 1992). One of the remaining two clones grouped with Marine Group II (DeLong 1992), the other with Marine Group III (Fuhrman and Davis 1997). Archaeal clone libraries from the 60-70 cmbsf horizon at SPG12 (Figure 2B) yielded solely MG-I sequences. Although few clones were recovered from the 153-163 cmbsf interval, MG-I archaea were also recovered in significant proportions at this horizon (4/11 sequences, ~36%), in a putatively suboxic zone well into the nitrate drawdown interval. Remaining sequences from this depth belonged to deeply branching Euryarchaeotal lineages (Durbin, 2009).

A total of 29 unique phylotypes at the 3% OTU level (Figure 3; Table 1) and 76 unique phylotypes at the 1% OTU level (Appendix A, Appendix B) were identified. MG-I subgroups were defined based on the major secondary or tertiary clusters visible in the tree with interior branch (IB; Sitnikova et al. 1995) support greater than 95%, with the exception of subgroups θ (theta) and ζ (zeta), which had IB support >90%. Non-chimeric representatives of previously delineated subgroups (e.g., Massana et al. 2000, Takai et al. 2004, Sørensen, et al 2004) were used to assign labels to subgroups recovered in the current trees. The identified clades were found to be stable, and had the same high interior branch support in both trees. Major subgroups for the SPG dataset overall were α, η (eta) and υ (upsilon), together accounting for 253/277 MG1 clones (~90%). A between-group distance matrix revealed maximum divergences for most subgroup pairs of 7-9% (Table 2). Divergence within most subgroups was shallow, with maxima ranging

Table 1. Representative phylotypes for MG-I, based on 3% OTU defintion.

Clone name	Accession number	Best BLAST % ID	Best NCBI match accession #	Best NCBI match clone name	MG-I subgroup	Overlying water	SPG11 3-4 cmbsf	SPG11 10-11 cmsbf	SPG11 26-28 cmbsf	SPG12 60-70 cmbsf	SPG12 153-163 cmbsf	Total
SPG11_H2O_A59	FJ487503	94%	AJ567637	MBAA50	η	1	2	11	8	39	4	65
SPG11_10_11_A23	FJ487550	94%	AJ567664	MBMPA55	η			2				2
SPG11_3_4_A29	FJ487535	94%	AJ870928	PS2ARC22	α		1					1
SPG11_10_11_A8	FJ487551	94%	AJ567663	MBMPA54	υ		2	25	42	5		74
SPG12_60_70_A21	FJ487517	95%	AJ567644	MBMPA7	υ		2	5	5	2		14
SPG11_3_4_A18	FJ487529	95%	AJ870325	MBWPA51	υ		2	2				4
SPG11_H2O_A91	FJ487481	95%	EU048592	MD2902-A3	ε	1	1					2
SPG12_60_70_A22	FJ487518	94%	AJ567649	PS2ARC24	η					2		2
SPG11_3_4_A3	FJ487536	95%	EF069355	PS2bARC23	υ		1	1		1		3
SPG11_3_4_A40	FJ487541	94%	EF069384	PS8ARC45	υ		1					1
SPG11_H2O_A82	FJ752631	94%	AJ567622	MBAA16	-	1						1
SPG11_H2O_A21	FJ487473	95%	FJ150794	116exp4	α	61	6	7				74
SPG11_3_4_A4	FJ487540	95%	EF069382	SSM263-NA06	α	1	2					3
SPG11_H2O_A62	FJ487493	97%	EU650239	050exp4	γ	2						2
SPG11_H2O_A109	FJ487483	95%	FJ150804	085exp4	α	1	2	1				4
SPG11_3_4_A42	FJ487544	93%	AJ567656	MBWPA32	η		1					1
SPG11_H2O_A1	FJ487499	96%	FJ150804	085exp4	α	1	1					2
SPG11_H2O_A110	FJ487472	95%	EF069378	PS8ARC36	α	3						3
SPG11_3_4_A12	FJ487525	93%	EF069369	PS7ARC24	-		1					1
SPG11_3_4_A10	FJ487524	94%	EU048600	MD2902-A13	-		2					2
SPG11_H2O_A10	FJ487476	95%	AB177110	ODP1230A22.16	θ	2	2					4
SPG11_H2O_A45	FJ487485	95%	EF069362	PS2ARC16	ζ	4						4
SPG11_3_4_A34	FJ487538	94%	AB177110	ODP1230A22.16	θ		1	1				2
SPG11_3_4_A9	FJ487547	94%	FJ175618	A5-82	θ		1					1
SPG11_H2O_A68	FJ487496	95%	AJ870321	MBWPA44	ι	3						3
SPG11_H2O_A88	FJ487507	93%	AB426388	WHA21-7	-	1						1
SPG11_H2O_A87	FJ487487	92%	AY627454	Urania-1A-30	-	1						1

Table 1. Representative phylotypes for MG-I, based on 3% OTU defintion. Relevant information on best BLAST match is included, along with MG-I subgroup designation, and phylotype abundance for each sample from which MG-I was recovered. Dash indicates if no subclade designation could be made.

Table 2. Average between-clade, within-clade, and maximum within-clade divergence levels for major MG-I subgroups recovered in this study.

	Mean between-clade distance					Mean within-clade	Max. within-clade
α: *Nitrosopumilus*						0.01938	0.0489
α						0.02887	0.0617
υ	0.0858					0.02654	0.0343
η	0.0839	0.0555				0.02364	0.0262
γ	0.0891	0.0827	0.0769				
ζ-ε-θ	0.0946	0.0788	0.0797	0.0872		0.03641	0.0445
ι	0.1021	0.0655	0.0785	0.0848	0.0609		

Table 2. Mean between-clade, within-clade, and maximum within-clade divergences for MG-I subgroups. Average between-clade, within-clade, and maximum within-clade divergence levels for major MG-I subgroups recovered in this study. Within-clade distances were not calculated for γ and ι clades, since very few representatives of these were recovered. All distances were estimated using a composite maximum likelihood (ML) substitution model with gamma-corrected site rates. Maximum within-clade distances approximately reflect the phylogenetic depth of clades. α:*Nitrosopumilus* shows the intragroup diversity within the crown-group MG-I α clade containing *Nitrosopumilus* and most of the SPG α diversity. Between-clade distances are not included for α:*Nitrosopumilus* as a subset of MG-I α.

from 2.6-6.2% maximum values, and mean within-group distances of 2.4-3.6%. The ζ-ε-θ cluster, with by interior branch (IB) support of 92%, is of approximately the same phylogenetic depth as the other major subgroups. The MG-I α clade containing the cultured MG-I representative, *Nitrosopumilus maritimus*, was supported by an IB value of 99%.

The phylogenetic differentiation of MG-I archaea between the water column and sediment, and among sediment layers, was clear at a finer taxonomic resolution. At SPG11, the dominant MG-I subcluster in the overlying water was MG-I α, comprising 67/84 MG-I clones (~80%). Fewer clones were found for MG-I γ (1 clone), ι (3), η (1) ζ (4), ε (1), and θ (3); 4 sequences were not affiliated with any of these groups (Figure 1B, Figure 3). At 3-4 cm sediment depth, MG-I α clones contributed 13 of 32 (~41%) clones, along with 3 MG-I η (~9%), 7 υ(~22%), 1 ε, 2 θ, and 6 ungrouped clones (Figure 1B). Notably, MG-I υ and η, both uncommon groups in the overlying water sample, were proportionally more abundant at 3-4 cm depth, representing ~22% and ~9% of the archaeal 16S rRNA clone library, respectively. At 10-11 cm sediment depth, MG-I υ increased to 61% of all clones. At 26-28 cm, the deepest depth analyzed for SPG 11, the

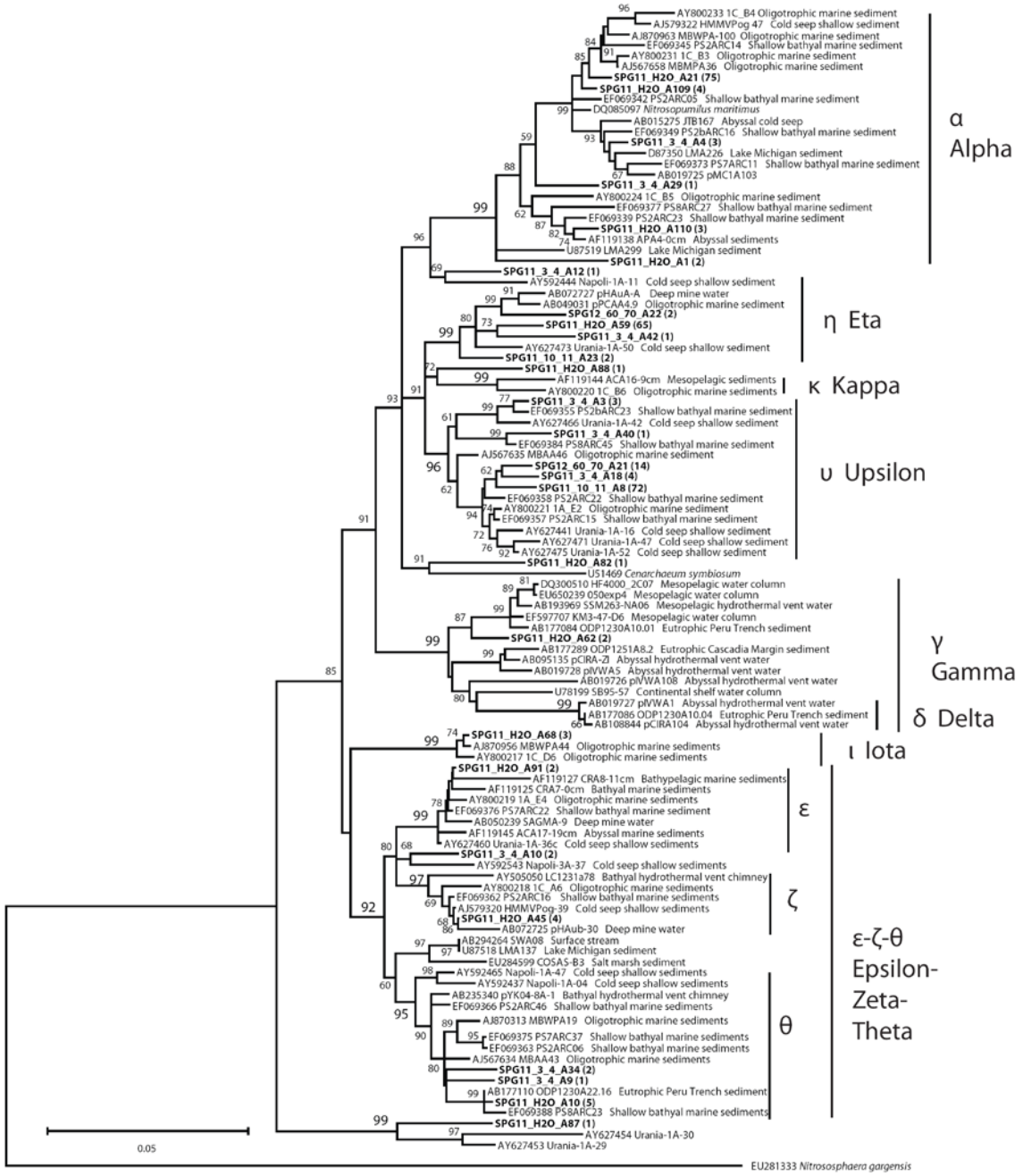


Figure 3. MG-I 16S rRNA gene phylogeny for representative phylotypes from SPG11 and SPG12, using 3% OTU definition. Distances were estimated using a neighbor-joining algorithm with gama-corrected site rates in MEGA4 (Tamura et al. 2007), using interior-branch node support, on an 800-bp alignment.

entire archaeal 16S rRNA clone library consisted of MG-I υ (49/57, ~86%) and η (8/57, ~14%). Site SPG12 lacked a fine-scale depth resolution, and therefore the depth trend of MG-I archaea is harder to infer at this site. However, as in the deepest SPG11 sediment layers, the MG-I clones in the SPG12 sediments

consisted almost exclusively of MG-I υ and η. At the shallowest depth analyzed at SPG12, 60-70cmbsf, MG-I υ and η were predominant, with η the most abundant group at 41/49 clones (~84%), along with 7 υ clones (~14%) and 1 α clone (Figure 2B). The deepest sample at SPG12 to yield any MG-I clones, 153-163cmbsf depth, resulted in a small number of MG-I η clones, four out of 11 archaeal clones (Figure 2B). At deeper sediment depths, MG-I archaea could not be detected, and the archaeal community of SPG 12 shifted towards entirely different archaeal phyla (Durbin 2009).

Diversity analysis.

Based on a whole-phylogeny-based analysis (Schloss and Handelsman 2006), each archaeal assemblage at SPG11 was revealed to have a structure significantly different from any other ($P < 0.00001$). However, further analysis revealed some statistical similarities between successive communities, in line with expectations from the observed depthwise changes in MGI clade relative abundances (Figure 1B, Figure 2B). Comparison of the shared Chao1 richness estimates of the overlying water and 3-4 cmbsf assemblages with the estimates for the two assemblages by themselves revealed a large overlap in community memberships at both the 1% and 3% OTU levels (Table 3A, B). Chao1 comparisons between 3-4 cmbsf and 10-11 cmbsf revealed a small degree of shared diversity at the 1% level, but the Chao1 estimate for shared 3% OTUs between these communities was greater than the estimate for 10-11 cmbsf 3% OTU alone, suggesting that these comparisons may be unreliable at extremely low diversity levels and small sample sizes.

The Jaccard index, which considers the total abundance of overlaps in phylotype membership (Yue and Clayton 2001), suggested approximately 5-25% of sequences in the SPG11 overlying water (OW) and 3-4 cmbsf sample sets belonged to shared OTUs (Table 3B). The ThetaYC estimate, which considers the "overlap" in both membership and library-specific relative abundance of OTUs (Yue and Clayton 2005), for any comparison involving the overlying water (OW) was highest (~50%) between the OW and 3-4 cmbsf samples; comparisons of the OW with deeper horizons quickly declined to 0. Both Jaccard and ThetaYC values were highest for comparisons between depths that were closest to each other

Table 3(A). Diversity indices for SPG11/SPG12 MG-I, 1% OTU definition.

	Sample indices			Comparative indices			
	# MG-I clones	Chao1 species diversity (95% C.I.)	Simpson diversity index	OW	3-4 cmbsf	10-11 cmbsf	26-28 cmbsf
OW	83	66.0 (46-125))	0.2865				
3-4 cmbsf	32	45.0 (28-101.9)	0.1512	26.5 0.145 (+/- 0.104) 0.0457 (+/- 0.046)			
10-11 cmbsf	52	30.3 (22-67)	0.2836	2.5 0.0509 (+/- 0.0504) 0.0443 (+/- 0.0266)	3.5 0.135 (+/- 0.078) 0.0671 (+/- 0.067)		
26-28 cmbsf	57	3.0 (2)	0.7544	0 0 0	1.5 0.0259 (+/- 0.0288) 0.0073 (+/- 0.0236)	6 0.519 (+/- 0.196) 0.102 (+/- 0.102)	
60-70 cmbsf (SPG12)	49	11.60 (11.1-18.1)	0.6547	0 0 0	1 0.0403 (+/- 0.0406) 0.0344 (+/- 0.0292)	9.5 0.1923 (+/- 0.101) 0.0826 (+/- 0.0393)	2.5 0.1088 (+/- 0.129) 0.0572 (+/- 0.0282)

Table 3(B). Diversity indices for SPG11/SPG12 MG-I, 3% OTU definition.

	Sample indices			Comparative indices			
	# MG-I clones	Chao1 species diversity (95% C.I.)	Simpson diversity index	OW	3-4 cmbsf	10-11 cmbsf	26-28 cmbsf
OW	83	15.5 (13.4-30.0)	0.0476				
3-4 cmbsf	32	22.5 (14.0-66.2)	0.0645	7.5			
				0.357 (+/- 0.207)			
				0.470 (+/- 0.115)			
10-11 cmbsf	52	5.0 (5.0)	0.0897	5	7.5		
				0.183 (+/- 0.267)	0.624 (+/- 0.206)		
				0.102 (+/- 0.045)	0.318 (+/- 0.120)		
26-28 cmbsf	57	2.0 (2.0)	0.6165	0	5	3	
				0	0.188 (+/- 0.127)	0.769 (+/- 0.208)	
				0	0.153 (+/- 0.079)	0.760 (+/- 0.114)	
60-70 cmbsf (SPG12)	49	3.0 (3.0)	0.1598	0	8.5	3	3
				0	0.219 (+/-0.137)	0.616 (+/- 0.219)	0.755 (+/-0.208)
				0	0.0749 (+/-0.0448)	0.373 (+/- 0.109)	0.195 (0.0650)

Table 3. Diversity indices for MG-I clone libraries from SPG 11 and SPG 12 based on 1% and 3% OTU definitions. (A) ~1% furthest-neighbor operational taxonomic unit (OTU)-based diversity analyses, and (B) ~3% furthest-neighbor OTU-based diversity analyses, calculated using DOTUR (Schloss and Handelsman 2005) and SONS (Schloss and Handelsman 2006). For comparative indices, first row is the shared Chao1 estimate (Chao 1984), the second is abundance-based Jaccard (Yue et al. 2002), and the third is ThetaYC similarity (Yue and Clayton 2005).

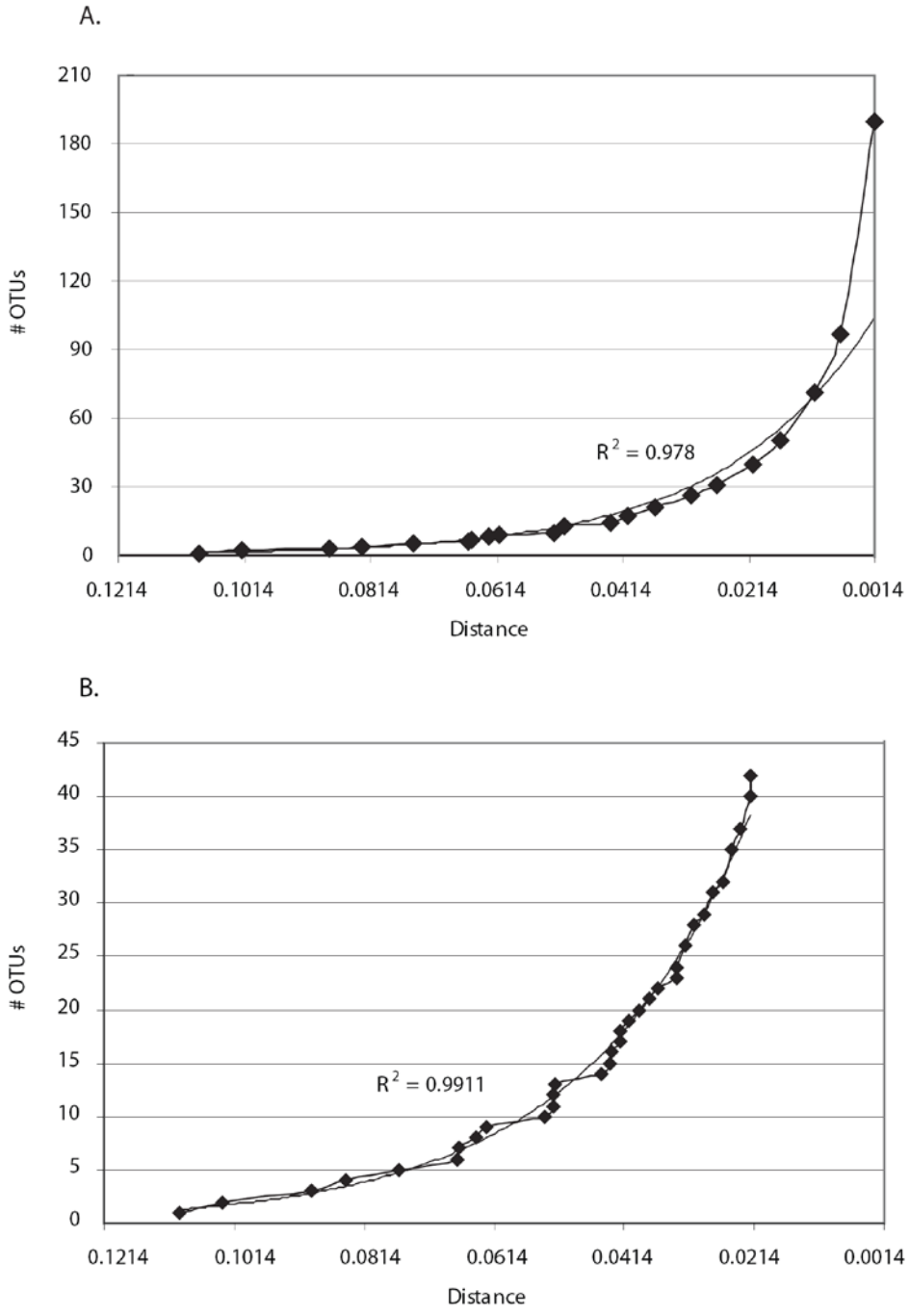


Figure 4. Lineage-through-time plots of cumulative abundance of OTUs vs. distance. (A) Lineage-through-time plot of number of OTUs vs distance, in 0.5% increments. (B) Lineage-through-time OTU/distance plot, excluding OTUs of less than 2% divergence. An exponential distribution trendline is included in both. (B) shows putative discontinuous bursts in cladogenesis rates, possibly associated with the divergence of the major MG-I subgroups recovered in this study, particularly MG-I α. A value of 0.14% was chosen as zero value, since this corresponds to the per-sequence variation introducted by Taq polymerase error (see Methods) and so is not natural microdiversity. All distances were calculated using the ML composite model with gamma rate correction in MEGA4.

(i.e., values increase up and to the right in Table 3); an exception to this pattern was the comparison of 60-70 cmbsf with 26-28 cmbsf; in this case, although these two samples were the closest depthwise, their comparison did not exhibit the highest ThetaYC value for any 60-70cmbsf, consistent with these two samples deriving from different cores. In some cases ThetaYC error values overlapped, likely due to low sample size. This pattern of higher similarity for adjacent sediment depths found much stronger statistical support in the Jaccard values. Together, the Jaccard and ThetaYC values for the archaeal 16S rRNA clone library data in the SPG11 sediment column are consistent with an archaeal assemblage that transitions in structure from one depth to the next: at least one shared 16S rRNA phylotype is abundant in one of two adjacent depths, while overall community structure of adjacent horizons is distinct. These observations are consistent with the trend observed from the percent-abundance plots (Figure 1, Figure 2), from an MG-I α -dominated overlying water sample to an MG-I η and υ dominated sediment community at 3-4 cmbsf and below.

Simpson's indexes sharply decrease both with shallower depth and when OTUs are defined at 1% rather than 3% divergence levels (Table 3), indicating a sharp increase in overall species richness and evenness both spatially (towards the water column) and at finer phylogenetic scales. The Simpson's indices for all MG-I combined (not shown) were 0.003741 for a 0% (unique) OTU, 0.04112 at 1% OTU and 0.1632 at 3% OTU.

A lineage-through-time plot of number of OTUs vs OTU definition distance level revealed that ~58% of MG-I diversity, or branches, are less than 1% divergent from their closest relatives, with 79 OTUs detected at a 1% OTU, compared to 190 unique OTUs, after adjustment for Taq error (see Materials and Methods; Acinas et al. 2004), and ~49% of diversity (branches) occurred at less than 0.5% divergence from closest relatives (Figure 4A). This high contribution of finest-scale diversity to total diversity was also noted in an extensive sequence effort on coastal bacterioplankton (Acinas et al. 2004). In the current study, an exponential trendline was found to describe the OTU abundance distribution with an r^2 value of 0.978, although clearly the data depart from this distribution at certain distance levels, particularly at the finest-scale divergences (Figure 4A). A consideration of only higher-level OTUs raised the possibility of discontinuous intervals of rapid cladogenesis between 3 and 6% divergence levels (Figure 4B), possibly associated with the branching of the major subgroups of MG-I represented in this dataset, particularly α,

which also has maximum internal divergence levels of ~5-6% (Table 2).

Discussion

A central question of this study was to determine if the oxic, oligotrophic sediments underlying the majority of the ocean surface represent a distinct microbial habitat, as revealed by a transition in community structure from the water column. The same archaeal phylum, the Marine Group I Crenarchaeota, was found to dominate in clone libraries from both the multicore overlying water and the sediment at SPG11, illustrating the major role the shared terminal electron acceptor, oxygen, has in shaping available microbial niches. However, finer-scale analysis revealed that distinct MG-I clades occur in the sediment and overlying water. A distinct transition from a putative pelagic to a sedimentary-associated MG-I assemblage was observed at SPG11, despite sharing a dominant terminal electron accepting activity with the water column. SPG12 shared the same dominant sediment-associated lineages as SPG11 sediment, although in a different relative ratio. These results suggest that different controls and constraints impact microbial community structure in the oxic water column and in oxic, ultraoligotrophic sediments.

At SPG 11, MG-I subphylum α accounts for ~80% of archaeal clones recovered from the overlying water, becomes less abundant at 3-4 cm and 10-11 cm depth, and is not found at 26-28 cm depth. This is reflected in the decreasing community-structural similarity (ThetaYC) between the overlying water and sediment as deeper horizons are considered (Table 3). Conversely, groups found sparingly (0-1%) in the overlying water, MG-I υ and η, became steadily more abundant with depth, comprising nearly all clones at 10-11 cm depth and all at 26-28. The high Jaccard similarities but low ThetaYC similarities between successive depths, combined with low Jaccard and Theta similarities comparing 3-4 cmbsf with 26-28 cmbsf, indicate that diversity and abundance overlaps are greatest between successive depths. This supports the characterization of the observed diversity pattern as a transition from a water-column (or shallowest sediment) MG-I assemblage to a sediment-associated assemblage in deeper sediments. The diversity maximum observed at 3-4 cm depth, where MG-I α, υ and η clones co-occur with some of the minor clusters found in the overlying water, e.g. MG-I ε and θ, can be explained as an example of an ecotone, wherein the boundary between two biomes hosts greater biodiversity due to the co-occurrence of species native to both habitats (Odum and Barrett 2005).

Interestingly, a lineage-through-time plot of number of OTUs vs sequence divergence revealed that more than 50% of 16S rRNA gene sequence diversity (i.e., number of branches) occurred at <1% divergence (Figure 4A). The same pattern was observed in studies of coastal marine bacterioplankton using extensive rRNA clone library sequencing (Acinas et al. 2004, Thompson et al. 2005). Here, diversity at less than 1% 16S rRNA divergence doubled the cumulative number of branches in a whole-community phylogenetic tree, compared to diversity seen at more than 1% divergence. This microdiversity below the 1% sequence divergence level is interpreted as functionally neutral variation, as judged by genome size variation (genome size variation in prokaryotes is principally driven by functional and not junk DNA content); competitive and selective forces are too weak to favor a single microhabitat-adapted organism in the bulk water column (~1 mL or greater). In contrast, the reduced number of branches at levels at and above 1% divergence are indicative of periodic selective sweeps reducing diversity at higher levels (Thompson et al. 2005). Since the MG-I subgroups are posited here to be adapted to sedimentary (or water column) conditions, one might expect a similar jump in diversity across the divergence thresholds associated with the MG-I subgroups, with putative selective sweeps associated with the foundation of these habitat-associated clades (founder effects), followed by relaxed selection which allowed greater diversity to accumulate and persist. The MG-I dataset shows some "jumps" of increasing OTU numbers at moderate (4-6.5 %) divergence levels when excluding finest-scale (<2%) diversity (Figure 4B). The major habitat-associated MG-I subgroups identified in this study, namely MG-I α, υ and η, have intergroup divergences of approximately the same divergence level (Table 2). These intergroup divergences could result from selective sweeps or adaptation to certain environments. Subsequent intragroup evolution and accumulation of 16S rRNA microdiversity may represent variation that is more or less functionally neutral within the context of sediment or water-column environments.

At SPG11, the decline in species-level Chao1 diversity with depth, particularly at the 3% OTU level (Table 3B), may indicate that deeper sediments are populated by remnants of the near-surface community. The cell counts are consistent with a surface-associated peak in microbial diversity and activity that undergoes a sharp dropoff over the upper 30 cm, wherein a much reduced community persists deeper into the sediment and declines slowly with depth. Notably, many of the cell counts made for SPG11 are below previous detection limits of 1.8×10^5 (Cragg and Kemp, 1995) or the 4.7×10^4 cells/mL limit

estimated for ODP Leg 201 (D'Hondt et al., 2003), and are lower than any other published number for the equivalent sediment horizon in marine sediments (D'Hondt et al. 2003, 2009). Accurate and precise cell counts for such low-biomass sediments were made possible by the recent development of a novel technique for separating cells from sediments prior to counting (Kallmeyer et al. 2008).

This apparent decline in species diversity also may result from diminishing DNA concentration with depth, leading to stochastic PCR bias where only a few phylotypes being amplified during initial PCR cycles, and subsequently dominate the reaction. Yet given a conservative 25% recovery of DNA from ~10 mL of sediment containing at least 100,000-300,000 cells per mL of sediment, and using 10-15% of this in each PCR reaction, at least 10,000 MG-I 16S rRNA gene molecules should be present in each PCR, assuming the 2/5 ratio of Crenarchaeota to Bacteria found in the deep North Pacific water column (Karner et al. 2001). Low DNA concentrations would result in jackpot amplification of only a few sequences being amplified early on and dominating all subsequent rounds of amplification. PCR inhibition could impact the amplified sequence diversity similarly, since inhibition result in a few templates randomly amplifying past some point at which the reaction is no longer inhibited, resulting in a selective sampling of 16S genes. In this study, three PCR reactions were combined during cloning to minimize jackpot amplification bias. Quantitative studies and powerful amplification techniques (e.g., φ29 polymerase; Biddle et al. 2008), are needed to confirm that a depth-related decline in sequence diversity results from low survivorship and not low DNA yield.

Metabolic activity of MG-I crenarchaeota in ultra-oligotrophic sediments.

Cultivation and genomic analysis suggest that the pelagic Crenarchaeota, and the MG-I in particular, are either obligate (*Nitrosopumilus maritimus*; Könneke et al. 2005) or facultative (*Cenarchaeum symbionsum*; Preston et al. 1996, Hallam et al. 2006) nitrifiers. MG-I Crenarchaeota have been shown to be the dominant nitrifiers in the North Sea and northern North Atlantic (Wuchter et al. 2006). MG-I also appear to be the dominant supplier of nitrite to the anammox reaction in the Black Sea, via a tight coupling between anaerobic ammonium oxidation (anammox) and microaerophilic or transiently oxic nitrification (Lam et al. 2007); through this linkage, pelagic Crenarchaeaota may be responsible for a significant portion of fixed nitrogen loss in oxygen minimum zones (Lam et al. 2007, Kuypers et al. 2005). Estimates based on calculations of per-cell autotrophic production (Herndl et al. 2005) suggest that nitrification by meso-

and bathypelagic Crenarchaeaota may regenerate all of the export flux of reduced N below the photic zone (Ingalls et al. 2006). However, comparative quantifications of 16S rRNA genes and ammonia monoxygenase genes from the North Atlantic suggests that nitrification is the dominant metabolism for MG-I populations only in waters north of 30 °N and in the subphotic zone (Agogué et al. 2008).

The upper 35 cm of sediment at SPG11 displays a clear vertical mirroring of oxygen and nitrate profiles, compatible with nitrification as a net metabolic activity that proceeds to at least ~30 cmbsf. This finding is consistent with the dominance of the putatively nitrifying MG-I crenarchaeota. Interestingly, the sharp decline in clade MG-I α from the water column to 3-4 cmbsf mirrors the sharp drawdown in oxygen over the same depth interval, and may also track the decline of substrates consumed by water column/sediment-water interface MG-I archaea. Thus, the shift to the MG-I η and υ groups at SPG11 may represent a shift in abundance of electron donors within the sediment, either reduced carbon or nitrogen, compared to the water column. Other explanations may include changes in the lability or accessibility of carbon substrates, availability or presence of specific reduced nitrogen species, size constraints due to pore size, and differences in viral predation pressure (e.g., Corinaldesi et al. 2007).

The MG-I sequences at SPG 12 originate from a sediment with different geochemistry; oxygen is found in lower concentrations (10-20 µM) at 60-70 cmbsf, compared to near-seawater concentrations throughout the sediment column of SPG11. Oxygen is nearly absent from SPG12 153-163 cmbsf layer (<10 uM), where nitrate is being drawn down. Despite these differences in oxygen availability, the MG-I η- and υ-dominated deep SPG12 and SPG11 sediment clone libraries are more similar to each other than either is to that of SPG11 overlying water. This supports the interpretation of the MG-I η and υ clades as consistently sediment-associated lineages in SPG11 and SPG12 sediments, excluded from the SPG11 water column.

One explanation for differing MG-I assemblage composition is a transition from a predominantly nitrifying, autotrophic metabolism to one based on heterotrophic substrates. Agogué et al. (2008) found a roughly one-to-one ratio between 16S rRNA genes and *amoA*, the marker gene for archaeal nitrification, only at the base of the euphotic zone and in subtropical-to-subpolar mesopelagic waters; noting that ammonia is in extremely low concentrations elsewhere, they propose that MG-I and other Crenarchaeota in bathy- to abyssopelagic waters and in non-subeuphotic tropical seas rely on a metabolism other than

ammonia oxidation, presumably heterotrophy. Further, Agogué et al. (2008) found that the proposed metabolic provinces were associated with different phylogroups of MG-I, principally α in the nitrifying regions, and γ elsewhere (results not shown). Thus, the community structure transition observed between the overlying water sample and the SPG11 and SPG12 sediment samples, and between SPG11 and SPG12 sediments, may be due to changing availability of key substrates, and perhaps the energetic feasibility of autotrophic nitrification versus heterotrophy.

The MG-I archaea are found in many different oxic marine environments, and have a strong potential to contaminate sediment cores and generally confound habitat-specific phylogenetic signatures of any marine biome, if the phylogenetic fine structure and habitat-associated diversification are not considered. Finding a specific association between a habitat within the marine realm, namely between oxic/suboxic marine sediments and the MG-I η to MG-I υ clades within the Marine Group I phylum, may therefore be important to future investigations of archaeal diversity in the oceans and sediments. The specific biogeochemical controls influencing the differentiation of sediment and water column lineages remain to be determined; they may possibly reflect a transition from autotrophic nitrification to heterotrophic nitrification or heterotrophy, as observed in Agogué et al. (2008). Here, studies of additional sites, perhaps also including fine-scale measurements of organic-N availability and rates of enzyme activity, as well as overlying water column samples taken independently of coring, will be required to further characterize the oxic marine sediment niche and to discriminate between explanatory hypotheses of the observed differences in community structure. Oxic gyre sediments thus represent a microcosm where most-energetic available electron acceptor is shared with the overlying water column, removing a key variable and allowing a window into the constraints structuring subsurface microbial communities.

Methods.

Geochemical and sedimentation data.

The sedimentation rate, estimated as a function of the age of the underlying crust and the thickness of the sediment column, was calculated to be ~1.8 m/Myr for SPG12, and ~0.9 m/Myr for SPG11. Cell

abundances and porewater chemistry measurements, including alkalinity, oxygen, and DIN, were determined as described elsewhere (Fischer et al. 2009; D'Hondt et al. 2009).

Sample collection.

Samples were obtained during the Knox02-RR cruise in January 2007 from two sites, one on the edge of the South Pacific gyre center, South Pacific Gyre Site 11 (SPG11) at 41 51.1281 S, 153 06.3849 W, water depth 5076m, and the other, SPG12, in the gyre margin at 45 7.855 S, 163 11.051 W at a water depth of 5306m. All SPG11 samples used in this study were taken from a multicoring device that sampled the upper 20-40 cm of sediment, as well as the overlying water. As soon as the multicorer came on deck, multicores containing sediment and overlying water were removed to a 4° C chamber, where all multicore subsampling took place. The near-bottom water column was sampled by dipping two 50-mL Falcon tubes into the overlying water of the multicore, which were subsequently frozen at -80° C. SPG11 multicore sediments were sampled by three autoclaved 60mL syringes with the tip cones cut off. To sample sediments, the core barrel was pulled down so that it was nearly flush with the sediment surface. Syringes were then pushed into the sediment, taking care to avoid sampling within ~1cm of the core barrel. Pulling the core barrel down further to just above the bottom of the cutoff syringe, the syringes were pulled laterally out of the core. The sediment was then extruded from the syringe and rounds of sediment were sliced off at specific intervals using a flame-sterilized scraper. Sampling intervals were 1 cm for 0-20 cm depth, 2 cm for 20-30 cm depth, and 5 cm for 30-40 cm depth. Samples from the same interval were pooled in a sterile Whirl-Pak bag and then frozen at -80° C. The sampling horizon was reestablished after each syringe sampling by using a flame-sterilized scraper to scrape the sediment flush with the syringe depressions.

SPG12 sediments used in this study were collected via a gravity core, of which total recovery was 401 cm. Cores were split into approximately 1m-long sections, which were then further sectioned into ~30 cm rounds. Cores were sampled by pushing autoclaved cut-off 60 mL syringes into the core cross section, avoiding an approximately 1-2 cm margin inside the core barrel, and then extruding the sample 1 cm at a time while slicing off subsamples using a flame-sterilized spatula. Subsamples were stored in sterile Whirl-Pak bags or sterile 50-mL Falcon tubes (BD Biosciences) and frozen at -80 until analysis.

Nucleic acid extraction, amplification and sequencing.

Genomic DNA for SPG12 samples was extracted from samples collected from 60-70 centimeters below the sediment surface (cmbsf) and 153-163 cmbsf, using a standard phenol-chloroform extraction protocol (Zhou et al. 1996) with some modifications. For SPG11, sediment horizons at 3-4 cmbsf, 10-11cmbsf, and 26-28 cmbsf were extracted for DNA, as well as overlying water (OW) from the multicore. Blank extractions with only buffer were also carried out under the same conditions as the appropriate sample.

DNA extraction of overlying water was undertaken using ~50 mL of frozen bottom water. For the DNA extraction of SPG11 overlying water, approximately 50 mL of -80 C frozen water was melted and filtered first through a 0.2 -µm Anodisc filter (Millipore). The filtrate was pipetted out of the receiving flask, and refiltered through a 0.02-µm Anodisc filter. Both filters were then frozen at -80 °C until DNA extraction. All tools and glassware contacting either the filter or filtrate were baked at 160 °C overnight, and previous to use filters were UVC-irradiated for approximately 30 min at a distance of 30 cm from a 15-watt UVC source.

To extract genomic DNA, filters were first added to a 60-mL Teflon tube (Oak Ridge) that had been baked at 160 °C overnight, and crushed using a baked spatula. Next were added 5 mL of phenol, 5 mL of extraction buffer (100 mM phosphate, 250 mM sodium acetate, 50 mM EDTA, 50 mM NaCl, pH 8), 5g of a 60/40 wt% mix of 0.1 µm and 0.2 µm low-binding zirconium silicate beads (OPS Diagnostics), 10 µg of poly[dIdC] (Barton et al. 2006) and 20% UVC-irradiated SDS to a final concentration of 0.9%. This mixture was then vortexed at highest speed for 2 minutes, frequently turning the tube on its side to mix beads thoroughly. The remainder of the extraction followed a standard phenol-chloroform protocol (Zhou et al. 1996), except that organic phases were extracted twice. The crude DNA extract was then precipitated by adding, sequentially, 2.5 ul/mL of UVC-irradiated GenElute linear polyacrylamide coprecipitant (LPA) (Sigma) to maximize precipitation of dilute nucleic acids, 0.5 volume of 7.5 M ammonium acetate (NH_4CH_3COO), and 1 volume of isopropanol. The precipitation was carried out at -20 °C overnight.

Beads used for all SPG12 extractions were a 60/40 mix of 0.1 mm and 0.2 mm low-protein-binding zirconium silicate beads (Shi 2005) (OPS Diagnostics). For the fine clay and low biomass

sediments of SPG11, a larger amount of sediment (12 g) was used for 10-11 cmbsf and 26-28 cmbsf, and a relatively higher bead-to-sediment mass ratio of ~0.3-0.4 was used. Additionally, beads used for SPG11 3-4 cmbsf sample were 0.1 mm zirconium silicate beads (Biospec). All other SPG11 extractions were performed with the same 60/40 mix of low-protein-binding zirconium silicate beads.

Sediment extraction for both SPG12 and SPG11 proceeded similarly with a few differences, based on what resulted in successful DNA recovery. For SPG12, 7-8 g of sediment was slurried in 7.5 mL of extraction buffer (100 mM phosphate, 250 mM sodium acetate, 50 mM EDTA, 50 mM NaCl, pH 8), and then added to pre-baked bead-beating canisters containing zirconium beads with a sediment-to-beads mass ratio of ~0.6-0.7. Approximately 30 µg of a carrier DNA, polydeoxyinosine-polydeoxicytosine (poly[dIdC]) (Barton et al. 2006), was then added to act as a blocking agent and to soak up nucleases to protect sample DNA. No carrier DNA was added to the SPG12 70 cmbsf and SPG11 3-4 cmbsf samples, as higher cell densities nearer the sediment surface reduced the need for a carrier DNA. Next, 15 mL of pH 8 buffered phenol and 875 µL of 20% sodium dodecyl sulfate were added, and the samples were subjected to bead-beating in a Braun MSK Cell Homogenizer for 20 s at high speed. The remainder of the extraction protocol followed a standard phenol-chloroform extraction and cleanup protocol, except that the sediment pellet was reextracted with a second volume extraction buffer with no SDS added, and this second aqueous phase was processed through the same organic phases as the first, thus re-extracting the organic phases and reducing DNA loss. The crude DNA extract was precipitated by addition of, in order, 2.5 ul/mL of UVC-irradiated GenElute linear polyacrylamide coprecipitant (LPA) (Sigma) to maximize precipitation of dilute nucleic acids, 0.5 volume of 7.5 M ammonium acetate, and 1 volume of isopropanol. The precipitation was carried out at -20 °C overnight. Following resuspension in PCR-grade water, the crude DNA extract was then processed through a MoBio PowerSoil kit as per manufacturer's instructions, with the omission of Solution C1, used in the cell-disruption step.

The primers A8f/A1492r (Teske et al. 2002) were used to amplify an approximately 1500-base pair (bp) fragment of the 16S gene of Archaea, for both sample and blank extractions. Sample PCRs were carried out in triplicate and then combined to maximize yield and to minimize "jackpot" amplification bias (e.g., Cha and Thilly 1993), wherein by chance a few sequences are amplified early on, and due to low DNA concentrations, these initial amplicons dominate subsequent rounds of amplification. Each 25-µL

PCR reaction contained 1 µL (SPG12 60-70 cmbsf), 2 µL (3-4 cmbsf), 10 µL (10-11cmbsf) or 15 µL (26-28cmbsf) of DNA template, 2.0 µL of dNTP solution (2.5 mM each dNTP), 0.250 µL (1.25 units) Takara Bio Inc. SpeedSTAR hot-start *Taq* enzyme, 2.5 uL of Takara Fast Buffer I (30mM Mg^{2+}), 1 µL of 100 mg/mL bovine serum albumin (BSA), and 1.5 uL each of 10 mM forward and reverse primers. The conditions for PCR were as follows: denaturation at 94 °C, followed by 35 cycles, each consisting of 5 s denaturation at 94 °C, 15 s at the annealing temperature, and 20 s of elongation at 72 °C, followed by one 10-minute elongation cycle at 72 °C. . Triplicate PCR reactions were then combined, target bands extracted and purified from a 2.0% low-melting-point (LMP) agarose (Promega) gel using SYBR Gold and the Invitrogen S.N.A.P. Miniprep kit, following the manufacturer's instructions. A gel fragment at approximately 1500 bp was also extracted and purified for the PCR and extraction negative controls, whether or not a band was present. All purified PCR products were further concentrated by precipitation with 0.5 µL of GenElute linear polyacrylamide (LPA) (Sigma). The pellets were then dried, and resuspended in 4-20 uL of PCR H_2O.

Sequencing was performed at the Josephine Bay Paul Center for Molecular Biology and Evolution at the Marine Biological Laboratory (Woods Hole, MA), using an ABI Prism 3730 sequencer and the same primers used in the initial PCR amplification of each sample. Sequence traces were trimmed, visually checked, and then assembled in Sequencher 4.7 (Genes Codes Corp.). All SPG sequences, as well as closest GenBank relatives identified via BLAST searches, were aligned in ARB (Ludwig et al. 2004) with final adjustments made by eye. Initial NJ trees were made in ARB, containing the GenBank and NCBI sequences that would be included in the final phylogenies. These datasets were then checked for chimeric sequences using the GreenGenes pipeline, which first requires alignment using the NAST aligner (DeSantis et al. 2006). Aligned sequences were then submitted to Bellerophon 3, using a window size of 200 bp, a "similarity to core set threshold" set to 99%, and the divergence-ratio set to 1.10, the last indicating that a sequence will be identified as a chimera if fragments derived from it have 'parent' sequences in the database that are more than 10% divergent from each other (Huber et al., 2004). MG-I Crenarchaeota sequences were additionally submitted to a more stringent chimera-check, with the same settings as above but a divergence ratio of 1.02, which was then repeated 2 additional times. The second 'stringent' run

resulted in a much smaller number of identified chimeras, and the third identified none, indicating the dataset was completely sanitized at that stringency level.

Operational taxonomic units (OTUs) were defined using DOTUR (Schloss and Handelsman 2005), and a single representative sequence was then selected from each OTU defined at the 1% level to be included in phylogenies. Additionally, 3% OTU representative phylotypes were selected for MG-I sequences. The best model of sequence evolution was then determined using Modeltest 3.7 (Posada and Crandall 1998) according to the Akaike information criterion (AIC), and was always a general time reversible (GTR) model with a proportion of invariant sites and a gamma distribution of site evolutionary rates. For the MG-I tree, a different approach than bootstrap resampling was necessary, since high numbers of sequences can result in serious underestimation of branch support using bootstrap methods (Rodrigo 1993, Sitnikova et al 1995, Sanderson and Wojciechowski 2000). The neighbor joining method with gamma corrected site rates can exceed the ability of the ML algorithm to calculate the correct topology and branch lengths, at least when a gamma-correction is not also used for the ML algorithm (Tateno et al. 1994), and performs equally well when considering at least ~1000 sites and low (<~5%) sequence, characteristics that nearly approximate the MG-I dataset. Therefore, a neighbor-joining tree (Saitou and Nei 1987), with a maximum-likelihood-estimated model of evolution and a gamma-corrected rate distribution, was calculated for the MG-I dataset using MEGA 4.0 (Kumar et al. 2008). Branch support estimated using a 1000-replicate interior branch test (Nei and Kumar 2000), which uses a hypothesis-testing approach to determine the probability that a particular interior branch has a length greater than 0, i.e., that the node in question should not be collapsed into a polytomy. Values greater than 95% should be considered strong support for a particular node (Nei and Kumar 2000). Using these methods, phylogenies were estimated for both the 1% and 3% MG-I OTUs, using the same sequences downloaded from GenBank as markers.

Treeclimber (Schloss and Handelsman 2004) was used to test the hypothesis that the overall structure of two communities is statistically the same. P values indicate the probability that a similar association between phylogeny and sample origin would occur by chance (Schloss and Handelsman 2006). SONS (Schloss and Handelsman 2006) was used to estimate the Chao1 (Chao 1984) species diversity of each depth, as well as cross-comparisons of community structure using abundance-based Jaccard (Smith et

al. 1996, Yue et al. 2001) and Theta (Yue and Clayton 2005) diversity indices. Both individual community Chao1 estimates and estimates of the Chao1 diversity shared between two communities (shared Chao1) were calculated. The abundance-based Jaccard index measures the intersection of the combined assemblage of phylotypes, and can be considered the probability that given a phylotype present in one assemblage, it is present in both assemblages (Schloss and Handelsman 2006). By contrast, ThetaYC corrects for the differences in relative abundance of phylotypes in each assemblage, taking into account both the incidence of shared phylotypes and the similarity of their relative abundances in each assemblage. Thus, it can be considered a percentile measure of the similarity of community structures.

Additionally, a plot of number of unique OTUs versus distance was made based on a collector's curve using all SPG11 and SPG12 MG-I sequences as input, calculated in DOTUR. A calculation of the contribution of Taq error to divergence was made according to Acinas et al. (2004), using an error rate of $2x10^{-5}$ per site per duplication (Klepac-Ceraj et al. 2004) and an average amplicon length of 1400 base pairs, yielding an average of 1.6 errors per sequence, or ~0.11%. This was used as a correction when interpreting the OTU# vs. distance collector's curve.

Chapter 2

Microbial diversity of oligotrophic abyssal sediments at the southern edge of the South Pacific Gyre

Introduction.

The organic-rich sediments of continental shelves and slopes are the chief sites of carbon burial on Earth, representing ~98% of marine carbon burial by one estimate (Dunne et al. 2007). Numerous microbiological studies of deep subsurface sediments with the Ocean Drilling Program (ODP) have focused on these locations and have recovered phylogenetically diverse, metabolically active bacterial and archaeal communities (D'Hondt et al. 2004, Parkes et al. 2005, Webster et al. 2006; Inagaki et al. 2006, Biddle et al. 2006, 2008; Lipp et al. 2008; Teske and Sørensen 2006; Teske and Sørensen 2008, Fry et al. 2008). In these organic-rich sediments, microbial activity depletes high-energy electron acceptors such as oxygen and nitrate within the surface few millimeters to centimeters; beyond this thin surface layer, subsurface sediments are entirely dominated by carbon remineralization processes with low standard Gibbs free energy, such as sulfate reduction, methanogenesis and fermentations.

In contrast, only a few microbiological studies have focused on the deep subsurface sediments of oligotrophic, open ocean sediments and abyssal basins, e.g., the Peru Basin (ODP Site 1231) and the eastern equatorial Pacific (ODP Site 1225) (Sørensen et al. 2004, Teske and Sørensen 2008). The more limited the carbon substrate and electron donor supply, the greater the depth of oxygen penetration. The upper few centimeters to meters of the sediment column contain the zones of maximal metabolic activity and cell numbers, which coincide with oxygen depletion and nitrate drawdown (D'Hondt et al., 2004, 2009). The microbial communities of these oligotrophic sediments are composed in part of species closely related to aerobic seawater Archaea (Sørensen et al. 2004, Durbin et al., 2009) and differ from the microbial communities in anoxic, organic-rich sediments (Teske and Sørensen 2008, Fry et al. 2008).

Microbial analyses of oligotrophic marine sediments require undisturbed cores of the upper sediment layers, and sufficient sampling resolution to account for microbial community changes in

response to oxygen and nitrate consumption. However, these critical upper sediment layers are generally heavily perturbed, resuspended and mixed with seawater by borehole initiation during deep subsurface drilling (Lever et al. 2006). To account for missing or low quality data from surficial sediments, deep subsurface modeling studies have focused on sediments below threshold depths (1.5 meter, D'Hondt et al. 2002).

In this study, we examined the phylogenetic structure of archaeal and bacterial communities from an abyssal, oligotrophic, oxic to suboxic sediment column (water depth 5306 m) from the margin of the ultraoligotrophic South Pacific Gyre (Site SPG12, position 45°58'S, 163°11'W), by 16S rRNA clone library sequencing, in comparison with cell counts and geochemical profiling, using undisturbed gravity core samples. The results are discussed with special attention to similar microbial communities from surficial marine sediments (7-100cm) of moderately to highly oligotrophic sites (Inagaki et al. 2001, Wang et al. 2004, Xu et al. 2005, Wang et al. 2005, Li et al. 2008).

Results.

Geochemistry. Rates of sedimentation, oxygen and nitrate depletion depths, and maximum DIC values indicate that SPG12 is significantly more oligotrophic than the most nutrient-limited marine sediments examined from ODP Leg 201. Net heterotrophy due to remineralization of organic matter to CO_2 causes an increase in DIC concentration (D'Hondt et al. 2003). The maximum alkalinity value (of which ~96% is DIC at seawater pH) for SPG12 revealed this site to be more oligotrophic than sites 1231 and 1225, at 3.23 mM vs. 3.6 and 3.98 mM, respectively, and is much lower than the maximum alkalinity values found for ODP Leg 201 Peru Margin sites, which range from ~20 mM to ~160 mM DIC (D'Hondt et al. 2003, 2009; Figure 5). The increase in alkalinity (DIC) with depth at SPG12 appears to be linear and not exponential in shape, and reaches its maximum and stable value of ~3.2 meq at about 100 cmbsf (Figure 5). As at Leg 201 oligotrophic sites, oxygen is depleted on a scale of tens of centimeters of the surface (D'Hondt et al. 2004, 2009). Oxygen decreases exponentially from a value of 114.76 µM at 5 cmbsf to an apparent detection limit of 2-10 µM below ~100-150 cmbsf from, while nitrate declines from a value of 43.8 µM at 10-15 cmbsf to 0 µM at 253-258 cmbsf, suggesting drawdown of oxygen followed by

depletion of nitrate (no multicore or overlying water sample was taken for Site 12 due to station time constraints).

A total of 373 archaeal and 249 bacterial 16S rRNA gene clones from seven different horizons at SPG12 were amplified using the bacterial and archaeal versions of primers 8f/1492r (Teske et al. 2002), then cloned, sequenced and analyzed phylogenetically.

Archaeal diversity and lineage distribution.

Phylogenetic analysis of the Crenarchaeotal and the Euryarchaeotal 16S rRNA clone libraries revealed several characteristics of these archaeal communities at SPG12. Archaeal phyla found in previously studied marine sedimentary environments, such as the Peru Margin, were either not present or were very rarely detected in the SPG12 samples. Most archaeal sequences found in the SPG12 sediment column formed unnamed, mutually exclusive phylogenetic lineages that contain few representatives in Genbank (Figure 6 and 7). The new groups include Marine Benthic Group (MBG) A1 to A3 within the crenarchaeota (Figure 7); the MBG-A2 and A3 archaea are new sister lineages of the MBG-A archaea that were originally found in North Atlantic sediments (Vetriani et al. 1999). Most of the Euryarchaeotal clones are affiliated with four lineages termed here the Deep-Sea Euryarchaeotal Groups (DSEG) 1 to 4 (Figure 7). DSEG 1 is synonymous with the DSEG lineage (without number) as described previously (Takai et al. 2001a), which is in turn synonymous with DHVE-3 (Takai and Horikoshi 1999). The only archaea in the SPG12 sediments that have cultured representatives are the Marine Group (MG) I archaea, a mostly pelagic marine phylum represented so far by aerobic and nitrifying strains (Könneke et al. 2005, Hallam et al. 2006). For each sediment horizon, total numbers of archaeal clones, phylogenetic affiliation, 1% OTU representative phylotypes, Genbank numbers and best BLAST hits are listed in Appendix C.

The archaeal 16S rRNA gene clone libraries show significant changes with sediment depth. The archaeal community above the oxygen depletion depth was found to be distinct from the communities found at deeper depths; a clear vertical community succession across the oxic-anoxic transition was evident when considering the overall occurrence of different archaeal phyla (Figure 5). The MG I archaea dominated the clone libraries in the upper sediment layer, the zone of oxygen and nitrate availability (Figure 5). The shallowest depth analyzed at SPG12, 60-70 cmbsf, corresponds to the bottom of the oxic

zone, just above the zone of gradual oxygen depletion (80-150 cmbsf) and the zone containing the highest recorded cell densities at SPG12 (~70-120 cmbsf) (Figure 5A). Here, the MG-I subgroups υ (upsilon) and η (eta) (Durbin et al. 2009) were predominant, with η the most abundant group at 41/49 clones (~84%), along with 7 υ clones (~14%) and 1 α clone. The next sampling horizon with MG-I clones, 153-163cm depth, was located in the suboxic zone, and yielded four MG-I η clones within a total sample of 11 archaeal clones (Figure 5, Figure 6). Based on sediment and bottom water analyses of other oligotrophic sediments and SPG sites, MG-1 υ and η appear to be specifically sediment-associated and distinct from the MG-I clades that typically appear in the water column, supporting their characterization as indigenous members of the sediment community (Durbin et al. 2009).

The archaeal 16S rRNA clone library from 213-223 cmbsf, just above the nitrate depletion depth, consisted of a unique archaeal assemblage dominated by novel phylogenetic lineages of uncultured euryarchaeota, termed here DSEG-4 (Deep-Sea Euryarchaeotal Group 4) and MBG-A3 (Marine Benthic Group A3). DSEG-4 archaea were only found in this depth horizon and accounted for a significant fraction of the clones (45/79). Overall, the 213-223 cmbsf horizon had the highest archaeal phylum-level diversity of clone libraries with significant sample size (Figure 5); the archaeal clone library included members of two crenarchaeotal (MBG-A2,-A3) and three euryarchaeotal (DSEG-3, DSEG-4 and MG-V) lineages.

Archaeal clone libraries from deep sediments in the suboxic and nitrate-depleted zone (273-283 cmbsf, 343-353 cmbsf) were dominated by uncultured crenarchaeota of the MBG-A1 and MBG-A2 lineages (Figure 5). The MBG-A1 contained 65 of 153 non-MGI Crenarchaeotal sequences from the SPG12 sediments, with nearly all clones (64) represented by a single 1% OTU phylotype (Appendix C). MBG-A sequences from public databases included several from oligotrophic sediment sequences. The MBG-A2 accounted for 84 sequences and 2 phylotypes of the non-MGI Crenarchaeota. All existing GenBank representatives of MBG-A2 were included in the crenarchaeotal tree (Figure 6), and were derived from oligotrophic sedimentary environments, including the eastern equatorial Pacific (Teske and Sørensen 2008) and western North Pacific (Xu et al. 2005). The MBG-A3 group contained 2 distinct subgroups (Figure 6), with one containing 4 SPG sequences represented by one 1% OTU phylotype, as well as a sequence from oligotrophic sediments at the East Pacific Rise (Li et al. 2008). Additional groups related to the MBG-A lineages come from very different environments, including hydrothermal vents, terrestrial hot

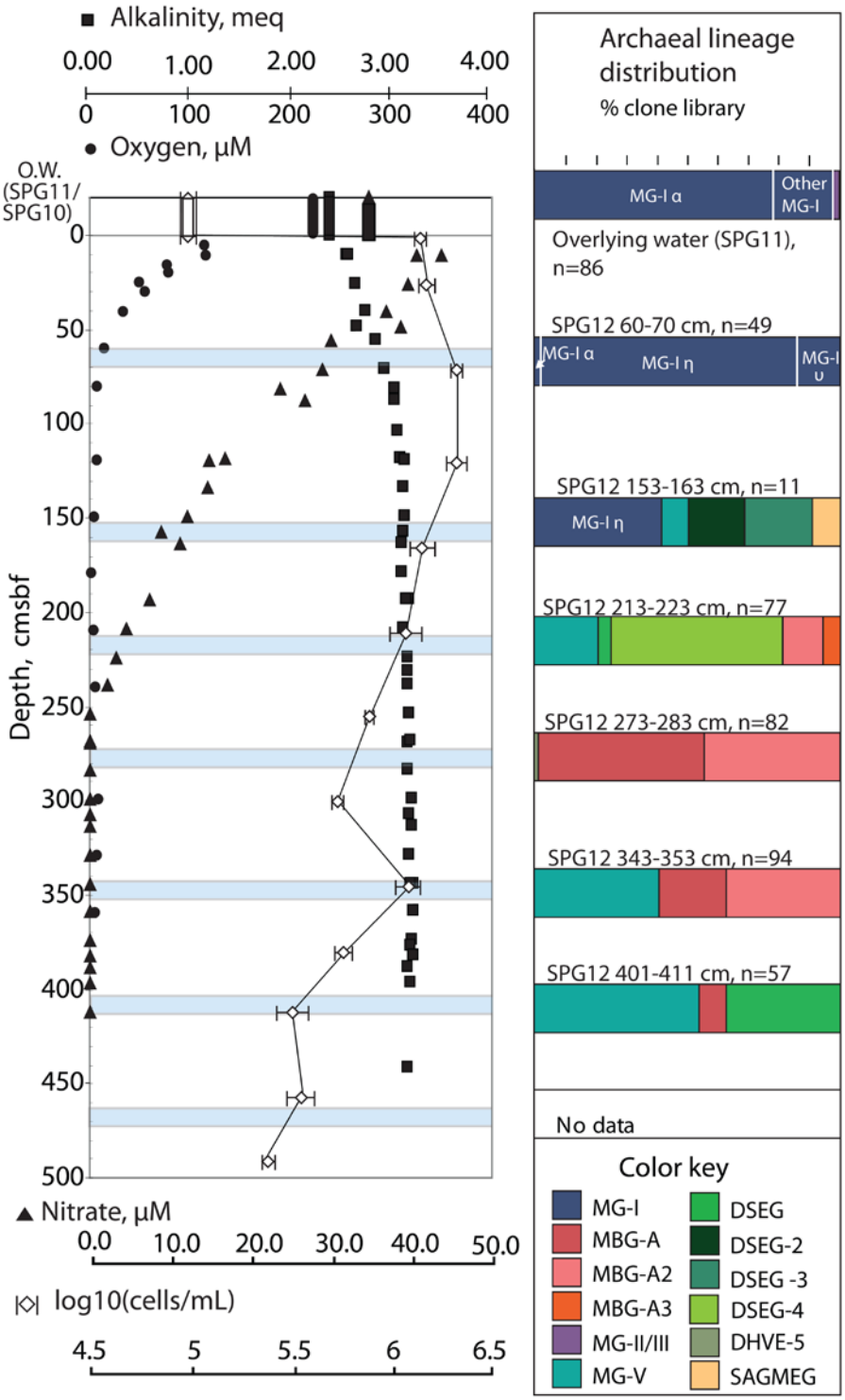


Figure 5. Depth trends in archaeal 16S rRNA clone library composition, porewater constituents, and cell counts. Marine Group I subclade abundance distributions in the upper two sediment layers and in the water column are based on Durbin et al. 2009. Porewater nitrate, alkalinity, and oxygen data, as well as cell counts (D'Hondt et al 2009) are plotted vs depth. Zero cmbsf values for cell counts and oxygen represent overlying water values from the SPG11 and SPG10 multicorer samples (D'Hondt et al. 2009), respectively, since no multicore was available for SPG12.

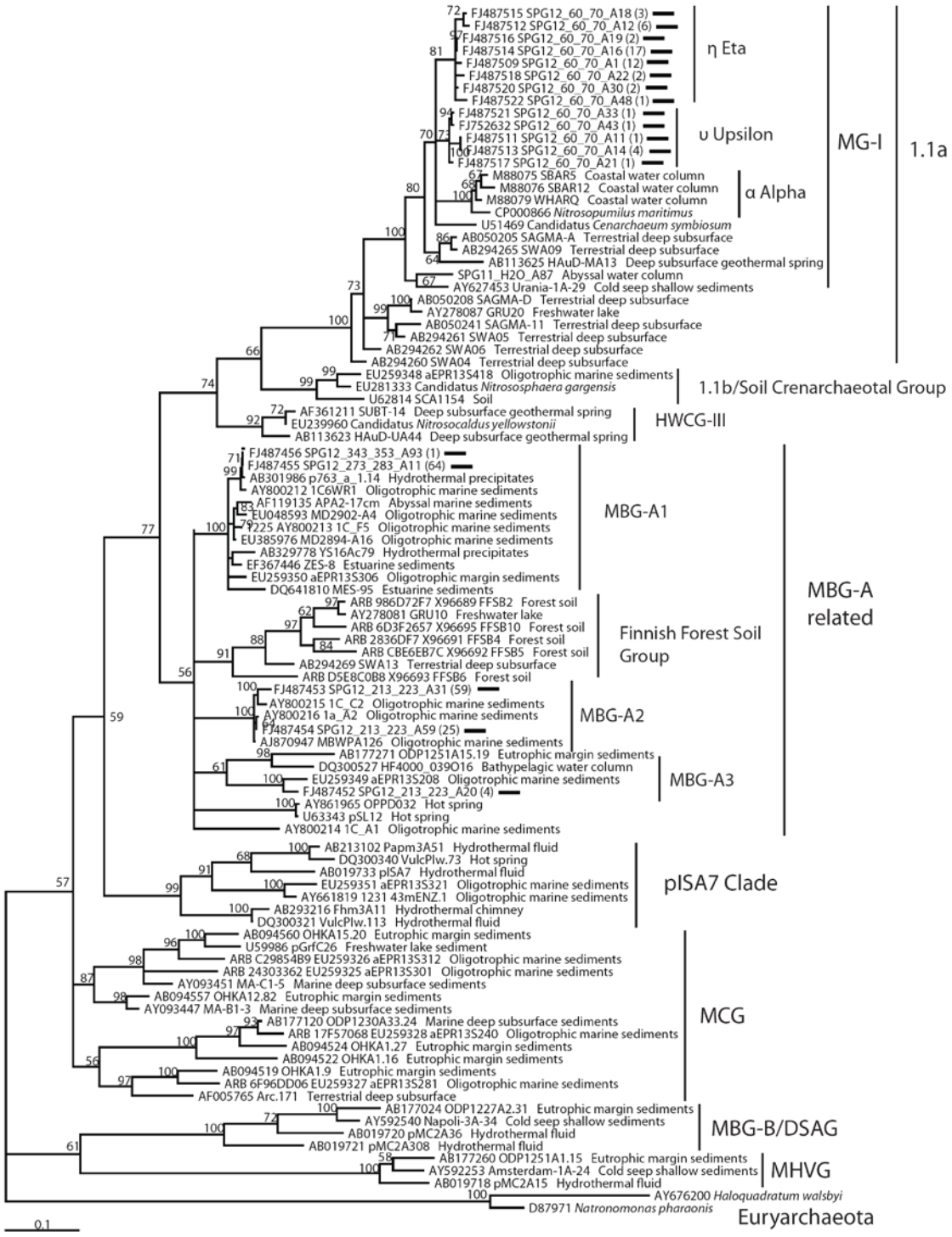


Figure 6. Maximum-Likelihood (ML)- estimated 16S rRNA gene phylogeny of Crenarchaeota, using a 701-bp alignment, and 1000 iteration ML bootstrap node support. Black horizontal bars indicate SPG12 phylotypes. The scale bar corresponds to expected number of substitutions/site.

springs (Barns et al. 1996, Spear et al. 2005), or boreal forest soil in Finland (Finnish Forest Soil Group, FFSG; Jurgens et al. 1997) (Figure 6)

.The deepest sediment sample was at 401-411 cmbsf, below the nitrate depletion depth, and yielded predominantly uncultured euryarcheota of the MG (Marine Group) V and DSEG-1 lineages (Figure 5, 6). MG-V accounted for 84/166 or 50% of all euryarchaeotal sequences recovered, with the bulk (79) represented by a single 1% OTU (Appendix C); clone library representation of this lineage increases with sediment depth (Figure 5). Only 5 sequences from GenBank grouped within the MG-V lineage, including two sequences from a tropical estuary (Singh et al., unpublished), and another from abyssal Antarctic Ocean shallow sediments (Gillan and Danis 2007) (Figure 6). MG-V is the sister lineage to the predominantly pelagic marine archaeal group MG-II (DeLong et al. 1992, Delong 1994). MG-II archaea are most abundant in the upper mixed layer of the ocean (Fuhrman et al. 1997, Church et al. 2003, Galand et al. 2009) and some are suspected to be at least facultatively phototrophic, via a bacteriorhodopsin proton pump (Delong et al. 2006, Frigaard et al. 2006). The MG-V and MG-II groups have 22% sequence divergence between each other, less than several of the mean intragroup distances between the new DSEG groups in the SPG12 sequence dataset (Table 4). Along with the principally pelagic MG-III (Furhman and Davis 1997) and sediment-associated MBG-D (Vetriani et al. 1999) and TMEG (Takai et al. 2001a), the MG-V and MG-II lineages are members of a large, but well-supported monophyletic phylum-level group, the Thermoplasmatales and their uncultured relatives (Figure 7).

The SPG12 euryarchaeotal clones included four deeply-branching lineages not abundant or not found in clone libraries from the Peru Margin or Peru Trench (Biddle et al. 2006, Inagaki et al. 2006), the DSEG/DHVE-3 clade (Takai et al. 2001a, Takai and Horikoshi 1999), here renamed the DSEG-1, as well as the novel DSEG-2, DSEG-3, and DSEG-4 (Deep-Sea Euryarchaeotal Groups), defined in the current study. Clones belonging to these lineages were obtained from shallow marine sediments, a suboxic water column, hydrothermal vent surfaces, sediments and fluids, and continental deep subsurface; there are no cultured members or enrichments (Figure 7). Corroborating the designation of these lineages as deeply branching, distinct, and reciprocally monophyletic phylogroups, the DSEG lineages identified here had higher mean differentiation from the next-closest related lineages available in the SILVA SSU REF v. 95 database than maximum or average intra-group divergences (Table 4). These phylum-level lineages also

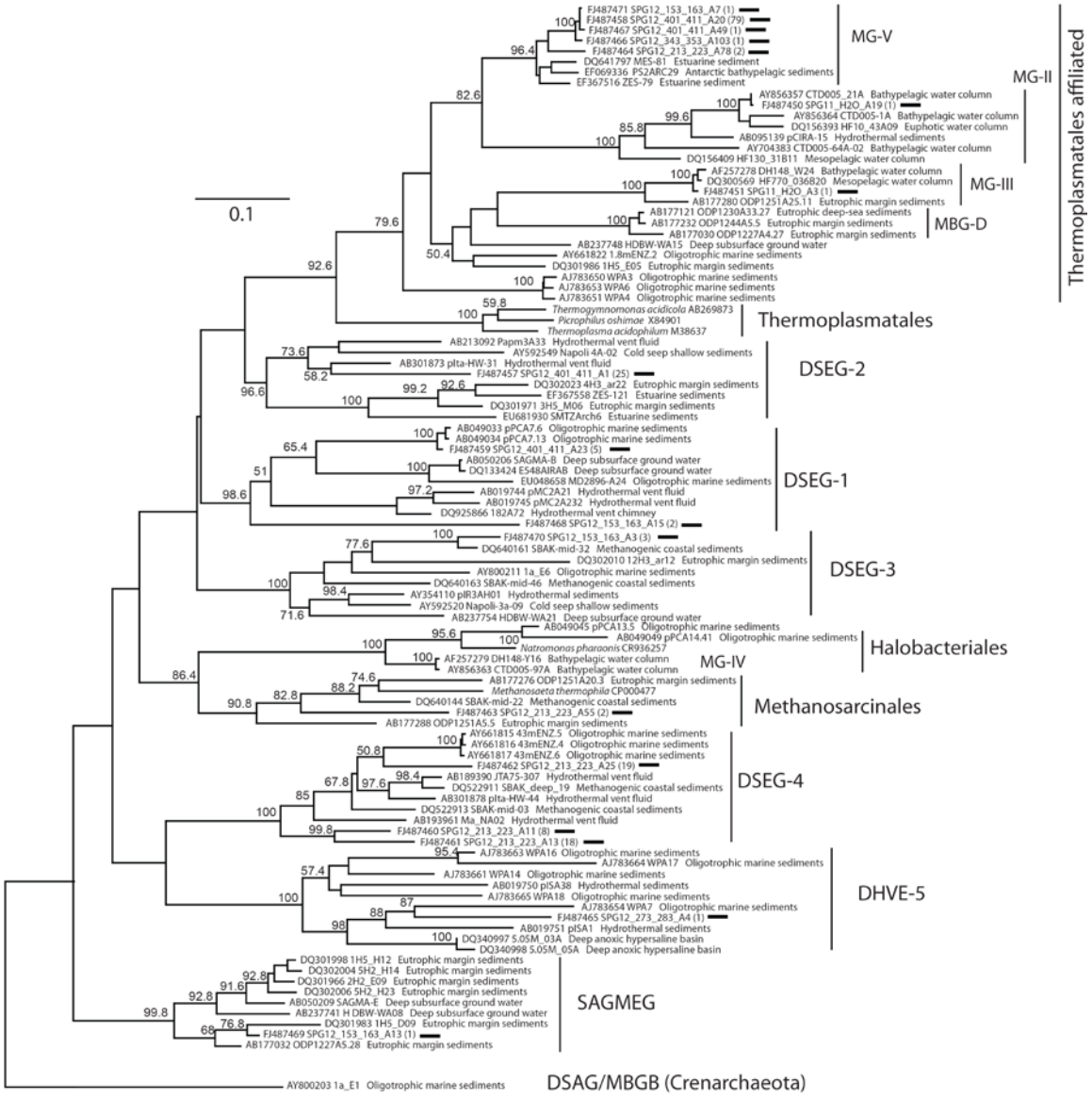


Figure 7. Maximum likelihood (ML)-estimated 16S rRNA gene phylogeny of SPG12 Euryarchaeota. A 702-base-pair (bp) alignment was used in order to accommodate shorter sequences. Node support is based on 1000 ML bootstrap iterations. Color coding indicates sequence origin, while SPG12 phylotypes are indicated by black horizontal bars. The scale bar corresponds to 0.1 expected substitutions/site.

have relatively few representatives in public databases (Table 4), with most representatives included in the phylogeny (Figure 7). For example, the second-most abundant Euryarchaeotal lineage at SPG12, DSEG-4 (Table 5), was represented by only 11 sequences in the SILVA v.95 database of archaeal 16S rRNA gene sequences >900bp (Pruesse et al. 2007), released 14 July 2008. A BLAST search (1 April 2009) of a

Table 4. Descriptive characteristics of the novel Euryarchaeotal clades identified in this study.

Lineage designation	Number of SPG12 sequences	Number of sequences greater than 83% max BLAST ID to SPG12 sequences (1 April 2009)	Number of sequences within lineage, in SILVA SSU REF v.95	Maximum divergence within lineage, for sequences in SILVA SSU REF v.95	Mean divergence within lineage, gamma-corrected distances	Mean distance between next-closest lineage not included in phylum definition	Representative sequences
DSEG-1	7	34	24	30-32%	26%	vs. DSEG-3: 37%	AB019743, EU635927, AB050206, AF355859, FJ487459
DSEG-2	25	20	18	30-32%	26%	vs. DSEG-3: 37%	EF687554, DQ640163, AY592520, FJ487457
DSEG-3	3	10	20	30-32%	19%	vs. DSEG-1: 37%	AB177269, AB301866, AB301873, FJ487470
DSEG-4	45	14	11	30-33%	32%	vs. DHVE-4: 41%	DQ522902, AY661815, DQ640157, FJ487460
MG-V	84	>50	3	5-7%	(MG-II: 11%) 6%	vs. MG-II: 26%	AF355880, EF069336, EF367516, FJ487458

Table 4. Descriptive characteristics of the novel Euryarchaeotal clades identified in this study, DSEG 1 to 4, and MG-V. Number of sequences is number of clones, not unique phylotypes. Selected representative SPG12 phylotypes were compared using BLAST (1 April 2009) and the number of sequences with maximum %ID significantly above ~80%, defined as >83%, are presented. The SILVA SSU REF v.95 database, released July 2008, contains all publicly available Archaeal 16S rRNA sequences above 900 bp and passing minimum quality standards (~10,000 sequences). Representative sequences include representatives of all most-basal lineages from within each clade, among sequences available in SILVA REF v.95 or from the current study

representative DSEG-4 SPG12 sequence returned only 14 sequences with greater than 83% maximum identity.

Bacterial diversity and lineage distribution.

The bacterial community showed depth-related changes in clone-library assemblage structure, both across the oxic/suboxic interface and with further increasing depth (Figure 8). For every sediment horizon, total numbers of bacterial clones, representative phylotypes, Genbank numbers and best BLAST hits are listed in Table 5. The clone library from the oxic 60-70 cm layer consisted of members of the Alphaproteobacteria, OP11/OD1 candidate phyla, Actinobacteria, Gemmatimonadetes, Planctomycetes, and Chloroflexi. In contrast, 60-90% of all sequences recovered from deeper sediment horizons grouped within the Planctomycetes and Chloroflexi phyla (Figure 8). Within these dominant phyla, it was determined whether previously defined subgroups were present (e.g., Elshahed et al., 2007, Inagaki et al., 2006, Morris et al., 2004, Kittelman and Friedrich, 2008, Kirkpatrick et al. 2006, Kohler et al., 2008, Harrison et al., 2009) or novel lineages had to be defined. Based on the taxonomic division of the Chloroflexi into several subphylum-level groups (Inagaki et al. 2006), the dominant Chloroflexus clade in the oxic zone at 60-70 cmbsf was identified as Chloroflexi VIb, whereas in all deeper sediment layers the subphylum Chloroflexi IV dominated (Figure 9B). A single sequence grouped within the SAR202 cluster (Morris et al. 2004), here designated the SAR202 branch 5, and several sequences grouped into an uncultured cluster within the Anaerolineae. Among the Planctomycetes, several novel, well-supported subphylum-level clusters were abundantly represented in SPG12 sediments, labeled here I, II, III and IV, as were several uncultured groups within the Planctomycetaceae (Figure 9A; Appendix D). The appearance and abundance of the Planctomycetes subgroups revealed a peak in both class-level diversity and relative abundance at 213-223 cmbsf (Figure 8), the zone of nitrate depletion. A single clone from the zone of nitrate depletion (213-223 cmbsf) grouped within the candidate genus *Scalindua,* which are anaerobic, ammonia-oxidizing marine bacteria (Figure 8).

Other uncultured, division-level bacterial clades recovered in lower numbers included AC-1 (Harris et al. 2004, Isenbarger et al. 2008), NT-B9 and NT-B2 (Reed et al. 2002), OP11 and OD1 (Harris et al. 2004), and Termite Group I (Okhuma and Kudo, 1996, Hongoh et al., 2008) (Figure 9C; Appendix D). A group here designated the Abyssal Sediment Group (ASG) may constitute a novel division-level clade of

Bacteria; it consisted of a single NCBI sequence, with all other clones of this group recovered from SPG12 at 60-70 cmbsf, 343-353 cmbsf, and 461-471 cmbsf (Figure 9C). Among phyla with cultured representatives, several clades of Deltaproteobacteria were recovered, including the *Nitrospina*, Desulfobacteriaceae, and SAR324 (Wright et al. 1997) lineages (Figure 9C).

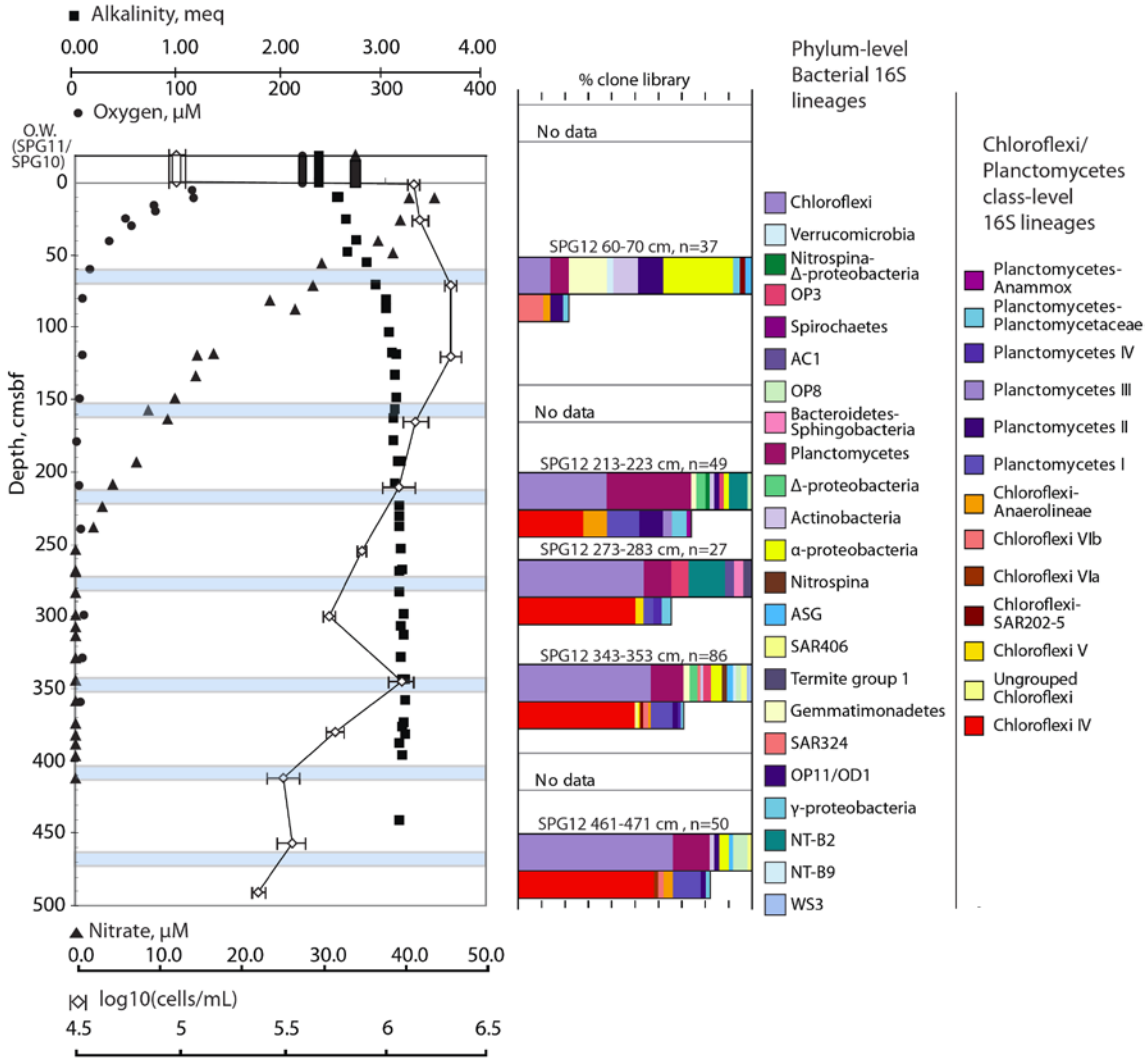
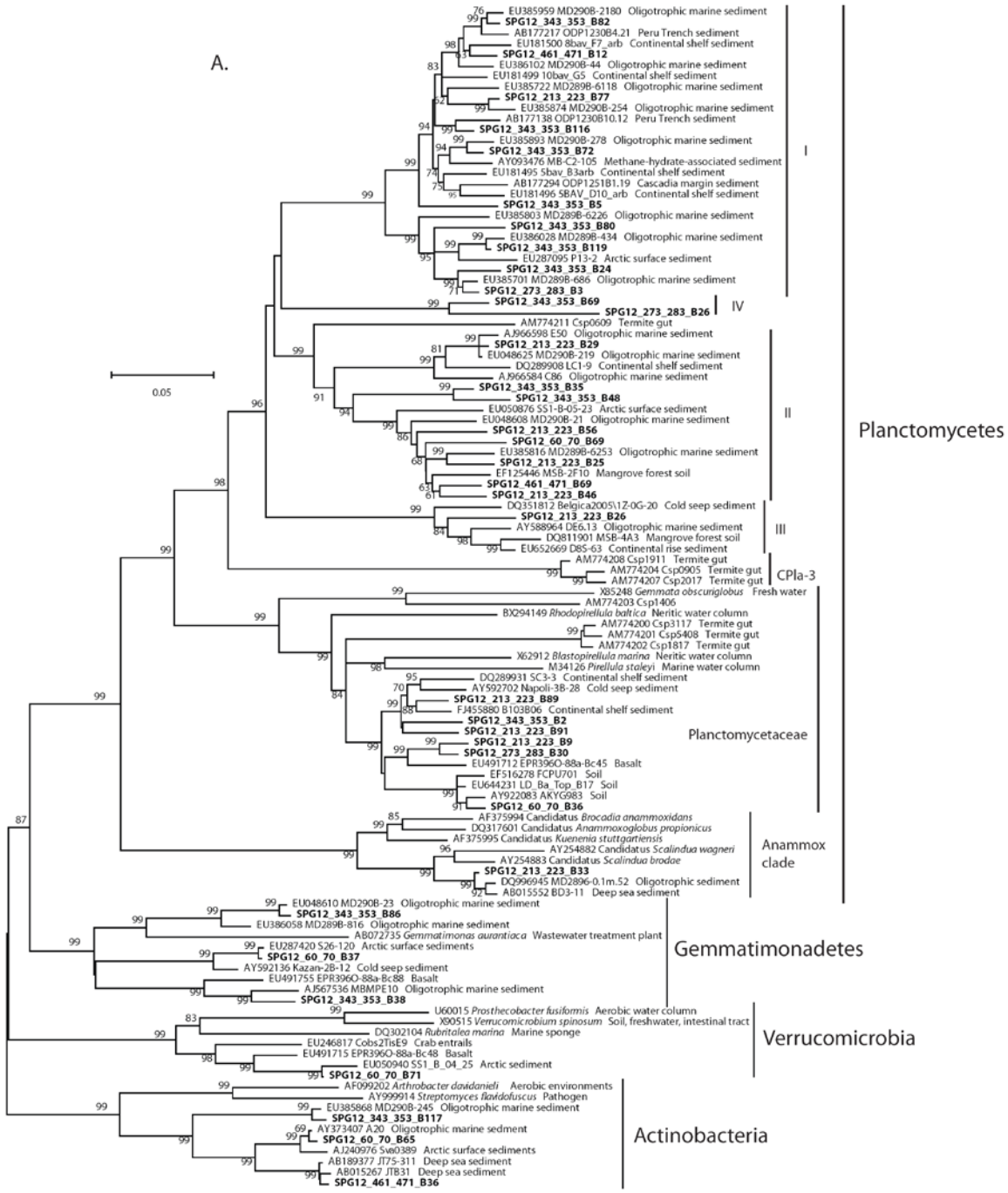


Figure 8. Porewater constituent data, cell counts and bacterial clone library composition (phyla and class/order abundances) at SPG12. Included are phylum level clades as upper bars, and class- or order-level clades for the two dominant phyla, Chloroflexi and Planctomycetes, as lower bars. SPG12 nitrate, alkalinity and oxygen data, as well as cell counts are provided at left (D'Hondt et al. 2009). 0 cmbsf values for cell counts and oxygen represent overlying water values from the SPG11 and SPG10 multicorer samples, respectively, since no multicore was available for SPG12.


A.
0.05
EU385959 MD290B-2180 Oligotrophic marine sediment
SPG12_343_353_B82
AB177217 ODP1230B4.21 Peru Trench sediment
EU181500 8bav_F7_arb Continental shelf sediment
SPG12_461_471_B12
EU386102 MD290B-44 Oligotrophic marine sediment
EU181499 10bav_G5 Continental shelf sediment
EU385722 MD289B-6118 Oligotrophic marine sediment
SPG12_213_223_B77
EU385874 MD290B-254 Oligotrophic marine sediment
AB177138 ODP1230B10.12 Peru Trench sediment
SPG12_343_353_B116
EU385893 MD290B-278 Oligotrophic marine sediment
SPG12_343_353_B72
AY093476 MB-C2-105 Methane-hydrate-associated sediment
EU181495 5bav_B3arb Continental shelf sediment
AB177294 ODP1251B1.19 Cascadia margin sediment
EU181496 5BAV_D10_arb Continental shelf sediment
SPG12_343_353_B5
EU385803 MD289B-6226 Oligotrophic marine sediment
SPG12_343_353_B80
EU386028 MD289B-434 Oligotrophic marine sediment
SPG12_343_353_B119
EU287095 P13-2 Arctic surface sediment
SPG12_343_353_B24
EU385701 MD289B-686 Oligotrophic marine sediment
SPG12_273_283_B3
SPG12_343_353_B69
SPG12_273_283_B26
IV
AM774211 Csp0609 Termite gut
AJ966598 E50 Oligotrophic marine sediment
SPG12_213_223_B29
EU048625 MD290B-219 Oligotrophic marine sediment
DQ289908 LC1-9 Continental shelf sediment
AJ966584 C86 Oligotrophic marine sediment
SPG12_343_353_B35
SPG12_343_353_B48
EU050876 SS1-B-05-23 Arctic surface sediment
EU048608 MD290B-21 Oligotrophic marine sediment
SPG12_213_223_B56
SPG12_60_70_B69
EU385816 MD289B-6253 Oligotrophic marine sediment
SPG12_213_223_B25
EF125446 MSB-2F10 Mangrove forest soil
SPG12_461_471_B69
SPG12_213_223_B46
II
DQ351812 Belgica2005\1Z-0G-20 Cold seep sediment
SPG12_213_223_B26
AY588964 DE6.13 Oligotrophic marine sediment
DQ811901 MSB-4A3 Mangrove forest soil
EU652669 D8S-63 Continental rise sediment
III
AM774208 Csp1911 Termite gut
AM774204 Csp0905 Termite gut
AM774207 Csp2017 Termite gut
CPla-3
X85248 Gemmata obscuriglobus Fresh water
AM774203 Csp1406
BX294149 Rhodopirellula baltica Neritic water column
AM774200 Csp3117 Termite gut
AM774201 Csp5408 Termite gut
AM774202 Csp1817 Termite gut
X62912 Blastopirellula marina Neritic water column
M34126 Pirellula staleyi Marine water column
DQ289931 SC3-3 Continental shelf sediment
AY592702 Napoli-3B-28 Cold seep sediment
SPG12_213_223_B89
FJ455880 B103B06 Continental shelf sediment
SPG12_343_353_B2
SPG12_213_223_B91
SPG12_213_223_B9
SPG12_273_283_B30
EU491712 EPR396O-88a-Bc45 Basalt
EF516278 FCPU701 Soil
EU644231 LD_8a_Top_B17 Soil
AY922083 AKYG983 Soil
SPG12_60_70_B36
Planctomycetaceae
AF375994 Candidatus Brocadia anammoxidans
DQ317601 Candidatus Anammoxoglobus propionicus
AF375995 Candidatus Kuenenia stuttgartiensis
AY254882 Candidatus Scalindua wagneri
AY254883 Candidatus Scalindua brodae
SPG12_213_223_B33
DQ996945 MD2896-0.1m.52 Oligotrophic sediment
AB015552 BD3-11 Deep sea sediment
Anammox clade
Planctomycetes
EU048610 MD290B-23 Oligotrophic marine sediment
SPG12_343_353_B86
EU386058 MD289B-816 Oligotrophic marine sediment
AB072735 Gemmatimonas aurantiaca Wastewater treatment plant
EU287420 S26-120 Arctic surface sediments
SPG12_60_70_B37
AY592136 Kazan-2B-12 Cold seep sediment
EU491755 EPR396O-88a-Bc88 Basalt
AJ567536 MBMPE10 Oligotrophic marine sediment
SPG12_343_353_B38
Gemmatimonadetes
U60015 Prosthecobacter fusiformis Aerobic water column
X90515 Verrucomicrobium spinosum Soil, freshwater, intestinal tract
DQ302104 Rubritalea marina Marine sponge
EU246817 Cobs2TisE9 Crab entrails
EU491715 EPR396O-88a-Bc48 Basalt
EU050940 SS1_B_04_25 Arctic sediment
SPG12_60_70_B71
Verrucomicrobia
AF099202 Arthrobacter davidanieli Aerobic environments
AY999914 Streptomyces flavidofuscus Pathogen
EU385868 MD290B-245 Oligotrophic marine sediment
SPG12_343_353_B117
AY373407 A20 Oligotrophic marine sedment
SPG12_60_70_B65
AJ240976 Sva0389 Arctic surface sediments
AB189377 JT75-311 Deep sea sediment
AB015267 JTB31 Deep sea sediment
SPG12_461_471_B36
Actinobacteria


B.
AB177209 ODP1230B33.10 Peru Trench sediments
SPG12_343_353_B101
AY869672 FS266-57B-03 Basalt crustal fluid
SPG12_343_353_B51
EU385734 MD2896-B134 Oligotrophic marine sediments
SPG12_343_353_B16
SPG12_461_471_B2
SPG12_343_353_B1
SPG12_343_353_B14
SPG12_213_223_B2
SPG12_343_353_B52
AF005748 H1.4.f Terrestrial deep subsurface
SPG12_273_283_B5
EU385714 MD2896-B106 Oligotrophic marine sediments
SPG12_461_471_B31
SPG12_343_353_B92
SPG12_461_471_B51
AY869675 FS117-3B-02 Basalt crustal fluid
SPG12_343_353_B31
EU385911 MD2902-B108 Oligotrophic marine sediments
SPG12_343_353_B104
AY869674 FS117-62B-02 Basalt crustal fluid
SPG12_343_353_B10
SPG12_343_353_B66
EU386009 MD2894-B12 Oligotrophic marine sediments
EU385930 MD2902-B139 Oligotrophic marine sediments
SPG12_213_223_B27
DQ223084 W2 Chlorinated ethene-contaminated aquifer
SPG12_343_353_B75
AJ519643 Sh765B-TzT-20 Mine tailings
SPG12_343_353_B11
AY559065 OTU-2 PCB-dechlorinating enrichment
AY584740 ER-A4-2 Estuarine sediment
SPG12_461_471_B8
SPG12_343_353_B99
SPG12_273_283_B34
EU181490 5bav_B7_ready Continental shelf sediment
SPG12_213_223_B31
EU385933 MD2902-B142 Oligotrophic marine sediments
SPG12_343_353_B41
SPG12_343_353_B110
AY704390 FS117-8B-02 Basalt crustal fluid
SPG12_343_353_B30
SPG12_213_223_B37
EU181494 8bav_H5_arb Continental shelf sediment
SPG12_343_353_B18
AY559072 OTU-9 PCB-dechlorinating enrichment
AY559073 OTU-10 PCB-dechlorinating enrichment
AF393781 Dehalobium chlorocoercia Dechlorinating culture
AY559070 OTU-7 PCB-dechlorinating enrichment
AF004928 Dehalococcoides ethenogenes Dechlorinating culture
U81649 vadinBA26
AB177132 ODP1230B1.19 Peru Trench sediment
AB177135 ODP1230B10.04 Peru Trench sediment
AY704394 FS118-62B-02 Basalt crustal fluid
EU181503 8BAVF8 Continental shelf sediment
EU385821 MD2896-B263 Oligotrophic marine sediments
SPG12_343_353_B76
SPG12_343_353_B103
EF018589 Amb_16S_858 Forest soil
EF018388 Amb_16S_1000 Forest soil
SPG12_461_471_B20
SPG12_343_353_B17
SPG12_60_70_B56
AY592140 Kazan-2B-16 Cold seep shallow sediment
AM935723 AMBB8 Soil
SPG12_343_353_B54
FJ479350 p36h18ok Prairie soil
U20797 SAR202 North Atlantic gyre seawater
AY534099 D92_22 North Atlantic gyre seawater
AJ567556 MBMPE42 North Atlantic gyre seawater
AY534094 SAR259 North Atlantic gyre seawater
AY534088 SAR272 North Atlantic gyre seawater
AY534095 SAR256 North Atlantic gyre seawater
AF186417 PAWS52f Marine sponge
AY534089 SAR242 North Atlantic gyre seawater
SPG12_343_353_B3
AY534098 SAR194 North Atlantic gyre seawater
AY534092 SAR292 North Atlantic gyre seawater
AB243673 Longilinea arvoryzae Anaerobic digestor
AB243672 Bellilinea caldifistulae Anaerobic digestor
AB046413 Anaerolinea thermophila Anaerobic digestor
EF208653 CI35cm.2.05a Sandy carbonate sediment
SPG12_213_223_B80
EU385894 MD2902-B80 Oligotrophic marine sediment
SPG12_343_353_B44
EU134954 FFCH14411 Soil
SPG12_60_70_B29
EU134962 FFCH17601 Soil
SPG12_461_471_B56
EU386099 MD2900-B26 Oligotrophic marine sediment
EU386014 MD2894-B17 Oligotrophic marine sediment
SPG12_213_223_B49
EU386021 MD2894-B24 Oligotrophic marine sediment
SPG12_343_353_B29
AM712351 MS11-3-F12 Hydrothermal vent
EU385667 MD2896-B45 Oligotrophic marine sediment
SPG12_343_353_B107
EU048672 MD2896-B13 Oligotrophic marine sediment
EF602487 Zplanct26 Euxinic hot spring
SPG12_273_283_B14
EU386007 MD2894-B10 Oligotrophic marine sediment
SPG12_213_223_B94
AB015559 BD4-9 Marine sediments
SPG12_343_353_B114
0.05
IV
Dehalococcoidetes
II
V
VIa
VIb
VI
SAR202-3
SAR202-4
SAR202-1
SAR202-5
SAR202-2
SAR202 clade
Anaerolineae
Chloroflexi
OP11
OD1
OP3

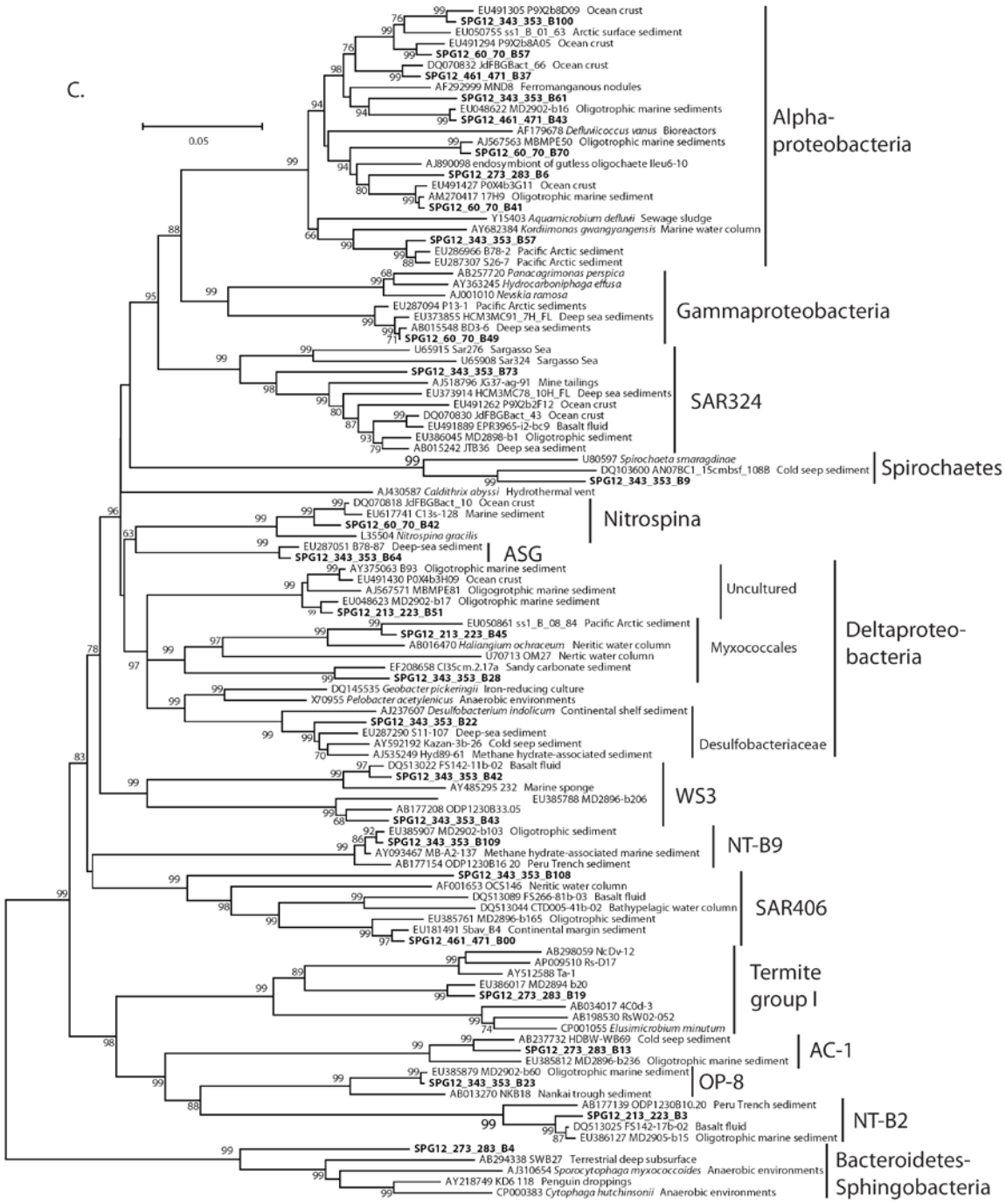


Figure 9. Neighbor-Joining phylogeny of SPG12 Planktomyces (A), Chloroflexi (B), and other bacterial phyla (C), based on a ~1200-bp alignment of bacterial 16S rRNA sequences that passed a 40% conservation filter using the filter tool in ARB. A maximum-composite likelihood substitution model with gamma-corrected site rates was used in MEGA4. Support for clades is provided by 2000-iteration interior branch values.

Archaeal and Bacterial comparative diversity analysis.

Differences between archaeal community structure in different sediment horizons are supported by

different statistical tests. Whole-phylogeny analysis using Treeclimber (Schloss and Handelsman 2006) indicated each Archaeal clone library to be significantly different from all others at $P < 0.00001$. Abundance-based Jaccard index values ranged from ~0.20 to ~0.36, with overlapping standard errors (Table 5A), excepting the 213-223 cmbsf/273-283 cmbsf and 401-411 cmbsf/273-283 cmbsf comparisons, both of which showed low between-library similarity. Because of the small sample size for 163-153 cmbsf, these results should be interpreted with caution, as should all comparisons between depth pair values with overlapping standard errors. As the Jaccard index represents the probability that a given phylotype is present in both communities, without correcting for the relative abundance of shared phylotypes in each community, the archaeal Jaccard values indicate that shared phylotypes form a low but significant portion of all archaeal clone libraries, with a few, non-universal exceptions. Overall community structure similarity, as estimated by ThetaYC, ranged from 0 for several comparisons to 0.3470 (+/- 0.0672) for 343-353 cmbsf/273-283 cmbsf. Notably, even for depths sharing dominant deeply-branching lineages and even 1% OTU phylotypes (e.g., 273-283 and 343-353 cmbsf), Jaccard and ThetaYC indices indicated relatively low community structure similarity.

Analysis using Treeclimber on Bacterial clone libraries found all clone libraries to be significantly different, with 95% confidence, although P values were larger compared to the Archaeal comparisons. For community structure comparisons, the frequently-used operational taxonomic unit (OTU) definition of 3% (Stackebrandt and Goebel 1994) rather than 1% (Stackebrandt and Ebers 2009), due to the high sequence richness of this dataset and the slightly greater comparative power a less stringent OTU definition offers. These diversity indices (Table 5B) confirmed the patterns observed in the percent-composition figures: a distinct community, at both the phylum and subphylum level, for samples from above the oxygen depletion depth (60-70 cmbsf), compared to deeper horizons (Figure 8, Table 5B); a gradual increase in the relative abundance of Chloroflexi with depth (Figure 8); and a diversity maximum among the Chloroflexi and Planctomycetes at 213-223 cmbsf (Figure 8).

Table 5 (A). Archaeal diversity indices, 1% OTU.

Archaeal diversity indices, 1% OTU	Sample indices		Comparative indices				
	# clones	Chao1 species diversity (95% C.I.)	60-70 cmbsf	153-163 cmbsf	213-223 cmbsf	273-283 cmbsf	343-353 cmbsf
60-70 cmbsf	49	11.6 (11.1-18.1)					
153-163 cmbsf	11	4 (4)	1 0.1504 (+/- 0.128) 0.2395 (+/- 0.115)				
213-223 cmbsf	79	10 (10)	0 0 0	1.5 0.05186 (+/- 0.0771) 0.01702 (+/- 0.0193)			
273-283 cmbsf	82	3 (3)	0 0 0	0 0 0	1 0.09788 (+/- 0.0996) 0.08531 (+/- 0.0322)		
343-353 cmbsf	94	4 (4)	0 0 0	1.5 0.06500 (+/- 0.0864) 0.02520 (+/-0.0275)	4 0.2730 (+/- 0.142) 0.1747 (+/-0.0474)	2 0.3601 (+/- 0.222) 0.3470 (+/- 0.0672)	
401-411 cmbsf	57	4 (4)	0 0 0	1.5 0.0869 (+/- 0.117) 0.07256 (+/- 0.0790)	3 0.2048 (+/- 0.122) 0.1305 (+/- 0.0425)	0 0 0	2 0.2907 (+/- 0.187) 0.1551 (+/- 0.0462)

Table 5 (B). Bacterial diversity indices, 3% OTU.

	Sample indices		Comparative indices			
	# clones	Chao1 species diversity (95% C.I.)	60-70 cmbsf	213-223 cmbsf	273-283 cmbsf	343-353 cmbsf
60-70 cmbsf	37	41 (25-106)				
213-223 cmbsf	49	207.5 (88-585)	5 0.05311 (+/- 0.0656) 0.02069 (+/- 0.0199)			
273-283 cmbsf	27	34.3 (21-91)	0 0 0	6 0.1078 (+/- 0.0779) 0.1239 (+/- 0.0724)		
343-353 cmbsf	86	122.3 (82-218)	30.5 0.1468 (+/- 0.0989) 0.03414 (+/- 0.0191)	47 0.5886 (+/- 0.238) 0.1200 (+/- 0.0488)	15 0.3024 (+/- 0.156) 0.2732 (+/- 0.104)	
461-471 cmbsf	50	84 (44-218)	17.5 0.1566 (+/- 0.133) 0.02422 (+/- 0.0171)	67 0.5527 (+/- 0.180) 0.1762 (+/- 0.0729)	5 0.2250 (+/- 0.106) 0.2453 (+/- 0.122)	34.4 0.4982 (+/- 0.158) 0.4378 (+/- 0.116)

Table 5. Archaeal and Bacterial diversity indices for SPG12 clone libraries. (A) ~1% furthest-neighbor operational taxonomic unit (OTU)-based diversity analyses for SPG12 Archaea, and (B) ~3% furthest-neighbor OTU-based diversity analyses for SPG12 Bacteria. Calculated using DOTUR (Schloss and Handelsman 2005) and SONS (Schloss and Handelsman 2006). For comparative indices, first row is the shared Chao1 estimate (Chao 1984), the second is abundance-based Jaccard (Yue et al. 2002), and the third is ThetaYC similarity (Yue and Clayton 2005).

This last was evident in the uniformly low ThetaYC values for all cross-comparisons with this depth (Table 5B). For comparisons between other depths, the highest ThetaYC similarity was generally found for immediately adjacent depths, although errors overlap somewhat, possibly due to relatively low sample size. This is consistent with a gradual, transitional change in community composition, a trend perhaps driven by the increasing proportion of Chloroflexi (Figure 8). This trend was not as clear in abundance-based Jaccard values, which sometimes had high values for non-adjacent depths; this likely reflects the uneven clone library size and richness and small sample sizes overall, as well as possibly shared haplotypes abundant only one of the two clone libraries under comparison. Notably, the ThetaYC similarity values tend to increase with depth, suggesting a stabilizing community composition, although the present dataset offers only a few points to support this.

Archaeal and Bacterial species diversity and rarefaction analysis.

Rarefaction analysis using either a 1% OTU (Figure 10) or 3% (not shown) OTU definition revealed that for all depths, Bacteria diversity was significantly higher than Archaeal diversity. Slopes of 1% OTU bacterial rarefaction curves approached 1 and did not show signs of leveling, suggesting sampling of the PCR reaction did not near saturation. By contrast, all archaeal rarefaction curves stabilized to a particular value, suggesting nearly all diversity amplified by PCR was sequenced (Figure 10). Using a 1% OTU definition, this value for Archaea was ~10 species or lower, a result also reflected in Chao1 diversity estimates (Table 5A). Chao1 diversity estimates for bacteria were much higher, and ranged from 34-208 with 95% confidence intervals ranging from 21-586 for 3% OTUs (Table 5B); it should be noted that such measures underestimate true diversity (Hong et al. 2006, Hughes et al. 2001). In addition to low depthwise archaeal diversity, several 1% phylotypes were recovered from widely separated depths. The DSEG-2 1% OTU phylotype SPG12_401_411_A1, comprising ~38% of clones from 401-411 cmbsf, also accounted for 5% of clones at 210-220 cmbsf; the MBG-A2 1% OTU phylotypes were found in abundance at both 273-283 and 340-350 cmbsf; and finally, MG-V phylotype SPG12_401_411_A20 was found as a single representative (~9%) at 153-163 cmbsf, but also made up ~38% and ~53% respectively of clones from the 343-353 cmbsf and 401-411 cmbsf intervals (Table 5).

.

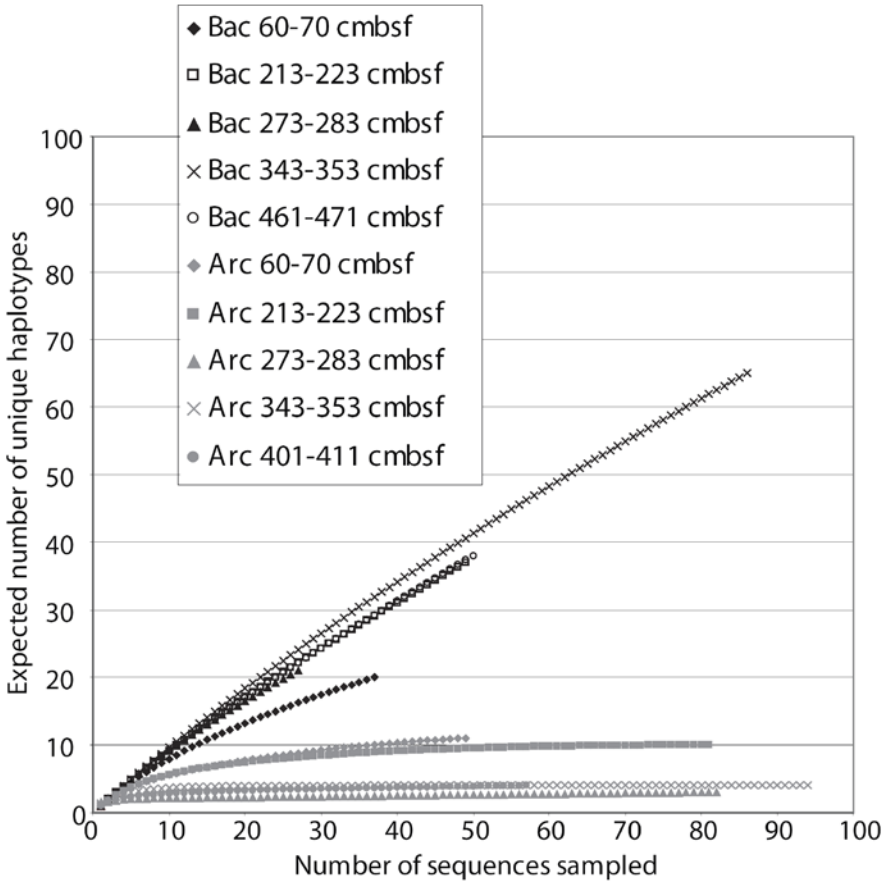


Figure 10. Rarefaction analysis at 1% OTU resolution for Archaea (gray symbols) and Bacteria (black symbols).

Discussion.

Several factors likely differentiate the oligotrophic sediment environment from both the oxic/suboxic and anoxic strata of organic-rich margin sediments. High energy electron acceptors are available over a wide depth range in oligotrophic sediments, in contrast to eutrophic margin sediments; this would generally favor organisms that can take advantage of the additional Gibbs free energy offered by higher-energy oxidants compared to organisms performing sulfate reduction, methanogenesis, or fermentations. However, the low availability of electron donors in oligotrophic sediments counteracts the advantages of electron acceptor abundance. The relatively high redox state of oligotrophic sediments may also impose an energetic cost in the pathways available for biomass synthesis, which is exacerbated by the low availability of reducing equivalents (McCollom and Amend 2005). Finally, oligotrophic sediments are also often found in deep waters and receive extensively degraded organic carbon (Wakeham et al. 1997, Hedges et al. 2001, Lee et al. 2004**),** which may be associated with protective mineral phases.

Additional suboxic redox processes may be possible in non-sulfidic, anaerobic sediments such as SPG12, compared to eutrophic sediments. In highly reduced, sulfidic sediments, metal oxides may be unavailable to respiration due to metal sulfide coatings. Reduced iron in the form of pyrite (FeS_2) is refractory to oxidation by Mn(IV) or nitrate (Schippers and Jorgensen 2001, 2002), unless the FeS_2 is coated by Mn(IV) or the iron oxide product is solubilized by chelation with organic ligands, since insoluble Fe(III) is the attacking agent (Schippers and Jørgensen 2001, 2002). Additionally, oxidation of metal sulfides produces protons, sometimes resulting in highly acidic environments such as acid mine drainages (Edwards et al. 1999); unless this acidity is buffered, for example by oxidation of metal-carbonate minerals (Severman et al. 2006, Nercessian et al. 2005), it impacts the feasibility of the metabolism, since the thermodynamics of microaerophilic metal oxidation become unfavorable at pHs below 5.5-6 for (Severman et al. 2006). Thus, in neutral pH sediments without metal sulfides, additional metal cycling pathways are feasible compared to acidic and/or sulfidic, eutrophic sediments (e.g., Schippers and Jørgensen 2001, 2002; Severman et al. 2006). Although SPG12 is below the carbonate compensation depth (CCD) and does not contain significant carbonate content (D'Hondt et al. 2009), the slow substrate input fluxes, and hence slow respiration rates, may allow the buffering capacity of seawater to maintain a near-neutral pH. Together, the different types of electron acceptors, higher reaction diversity of those electron acceptors and inorganic electron donors, and recalcitrance of deep-sea carbon might be expected to shape a phylogenetically distinct microbial community at SPG12 compared to eutrophic margin sediments, either suboxic or anoxic.

Archaeal diversity and distribution.

This study appears to confirm a distinct microbial community at SPG12 for the Archaea. Some of the archaeal lineages recovered from SPG12 sediments are mostly found in oligotrophic marine sediments, such as the MBG-A-related Crenarchaeotal lineages (Figure 5, Figure 6). Other lineages, for example the Euryarchaeota DSEG 1-4, show a broader habitat preference, including various marine sediments and hydrothermal vent environments, but are consistent members of the SPG12 archaeal community (Figure 5, Figure 7). Widespread archaeal groups commonly found in organic-rich subsurface sediments are conspicuously absent (e.g., Biddle et al. 2006, Inagaki et al. 2006, Sørensen and Teske 2006, Parkes et al. 2005). The composition of archaeal clone libraries at SPG12 appears to fluctuate widely with sediment

depth, with neighboring sediment horizons having different phylum-level archaeal lineages. The resulting depth distribution appears to reflect the redox zonation of the sediment to some extent (Figure 5): The MG-I archaea predominate in clone libraries from the oxic portion of the sediment column, and disappear as oxygen nears the detection limit. The MBG-A-related lineages are most abundant in the 273-283 cmbsf horizon just below the nitrate depletion depth, possibly associated with a redox interface involving metals and NO_X species, which occur in largest spatial extent in oligotrophic sediments (D'Hondt et al. 2004). Both nitrate and manganese (IV) oxides are viable electron acceptors for dissolved, reduced iron (Schippers and Jorgensen 2001, 2002; Straub et al. 1996), the form in which Fe(II) would be expected to occur in, given the lack of sulfide production suggested by an unchanged sulfate profile (D'Hondt et al. 2009). If the clone libraries at 273-283 and 343-353 cmbsf represent active populations of MBG-A related lineages, the fact that they are most abundant in clone libraries below the nitrate depletion depth suggests nitrite or manganese(IV) as a possible terminal electron acceptor.

By contrast, the DSEG-4 occur solely in the 213-223 cmbsf clone library, just above the nitrate depletion depth, and thus nitrate is available to these organisms; along with their absence from more surficial clone libraries in the denitrification zone, this would be consistent with a reliance on neutrophilic, non-heterotrophic denitrification using iron(II) diffusing up from below (e.g., Straub et al. 1996, Emerson and Moyer 1997). Thermodynamics suggest that manganese(II) can be oxidized by both nitrate and oxygen, although available evidence suggests that only oxygen is used in nature (Clement et al. 2009). The detection limit of the instrumentation commonly used to measure oxygen in sediments precludes verification of the possibility that microaerophilic manganese oxidation occurs at SPG12, and indeed the measured oxygen values never drop below the minimum required for significant manganese(II) oxidation to occur (Clement et al. 2009). Mn(III) may also be important in the extended suboxic zone of SPG12. This soluble metal species, indistinguishable from Mn(II) by standard methods, has recently been shown to be a key component of the manganese redox cycle in the Black Sea water column, where manganese cycling maintains a wider suboxic zone (without either oxygen or sulfide) than would otherwise be the case (Trouwborst et al. 2006). As Mn(III) can act as either an electron acceptor or donor, this complicates the network of redox pathways possible in the suboxic zone or oxic interface at SPG12.

Unlike the DSEG-4 or MBG-A-related lineages, MG-V archaea do not appear to have a specific

association with the nitrate profile, but rather appear to increase with depth, from just above the nitrate-depletion zone to the deepest sediment analyzed here, 401-411 cmbsf, excepting the 273-283 cmbsf interval where only MBG-A1/A2 were recovered. While this would appear to rule out denitrification, this distribution in clone libraries is consistent with a metal-reducing niche, along with other anaerobic heterotrophic metabolisms.

Bacterial diversity and distribution

A key difference between bacterial and archaeal clone library assemblages at SPG12 are the gradual depth transitions within the bacterial clone libraries (Figure 8), in contrast to the more discontinuous changes of archaeal communities from one depth layer to another (Figure 5). A continuously transitioning bacterial community structure as revealed in clone libraries is supported by ThetaYC values (Table 5B). ThetaYC compares overall community structure similarity and is not biased by shared haplotypes abundant in only one community, as are Jaccard indices. The highest bacterial clone library Theta YC values are found for comparisons between 'adjacent' sediment horizons, supporting a transition from one horizon to the next. This trend may be driven principally by a steadily increasing proportion of the Chloroflexi IV subphylum with depth (Figure 8).

The SPG12 bacterial community consisted largely of recognized phylum-level clades, such as the Planctomycetes and Chloroflexi, which comprised 60-90% of clones below the (putative) oxygen depletion depth. In contrast to the Archaea, subgroups of the Planctomycetes and Chloroflexi abundant in SPG12 clone libraries are also dominant in clone libraries from eutrophic margin sediments; for example, the dominant Chloroflexi subgroup at SPG12, Chloroflexi IV (Figure 9B), is also abundant at several continental margin sites (e.g., Inagaki et al. 2006), and while the dominant Planctomycete clades I and II (Figure 9A) found at SPG12 include numerous sequences from another oligotrophic abyssal sediment environment (Li et al. 2008), interspersed with these were clones from methane hydrate-associated sediments and eutrophic continental shelf sediments. Thus, it appears that these Planctomycete and Chloroflexi clades are metabolically flexible, using a wide array of electron donors and/or acceptors. It is also possible that they specialize in particular substrates that are abundant in marine environments, particularly less-labile deposits such as deep marine sediments. For example, the dominant Chloroflexi

clade, IV, is closely related to two obligate reductive dehalogenators, *Dehalobium chlorocoercia* and *Dehalococcoides ethenogenes*, and is the clade most closely related to these dehalogenators among uncultured clades identified from the eutrophic continental margin deep subsurface (Inagaki et al. 2006). It is unclear whether halogenated substrates are particularly abundant in ocean sediments, and especially so in recalcitrant sediments such as those at SPG12; it may also be that the dehalogenation pathway is to some extent shared by pathways used to degrade recalcitrant carbon.

The greatest subphylum diversity of Planctomycetes, including the Anammox clade, was found at the 213-223 cmbsf depth, just above the nitrate depletion horizon (Figure 8); these results are similar to a previously discovered peak in Planctomycete diversity and abundance at the intersection of nitrate and ammonia profiles in the Black Sea water column (Kirkpatrick et al. 2006), and suggest a key N cycling role for Planctomycetes in SPG12 sediments. The deeply divergent anammox clade of Planctomycetes has a well-documented role in mediating anaerobic ammonium oxidation in marine environments (e.g., Schmid et al. 2007). Evidence from genomics also implicates another Planctomycete isolate, more closely related to the uncultured clades found at SPG12, in nitrogen cycling. Nitrite/nitrate import and respiration genes were found to be predicted highly expressed (PHX) in the genome of the planctomycete, *Rhodopirellula balthica* (Glöckner et al. 2003), suggesting the possibility of denitrification. Planctomycetes have also been shown to be capable of organic fermentations (Schlesner and Stackebrandt 1986), consistent with their appearance below the nitrate depletion depth. Finally, *Rhodopirellula balthica* was found to host an unusually large number of sulfatases, likely involved in allowing more efficient access to sulfated carbon skeletons (Glöckner et al. 2003). Thus, the uncultured Planctomycete lineages abundant in suboxic SPG12 sediments may also be targeted towards sulfate-ester-linked glycopolymers, abundant in marine snow (Passow 2002). The highly diversified, potentially N-cycling Planctomycetes-dominated bacterial community at 213-223 cmbsf was echoed in the highly diversified Archaeal community at this depth horizon (Figure 8).

In addition to Fe(II) oxidation using Mn(IV)/ nitrate, or metal oxidation with microaerophilic c[O_2], an additional thermodynamically favorable metabolism for organisms at the nitrate depletion depth is ammonia oxidation using manganese (Luther et al. 1997, Hulth et al. 1999). Although evidence for this occurring in nature is lacking (Thamdrup and Dalsgaard 2000), it is possible that the differing pH and

sulfide contents of eutrophic continental margin sediments and oligotrophic SPG12 allows anammox with manganese to proceed.

Although no significant sulfate drawdown is evident from porewater geochemical data (D'Hondt et al. 2009), the appearance of Deltaproteobacterias sequences at two depths (213-223 and 343-353 cmbsf), including a clone from the sulfate-reducing Desulfobacteriaceae, which has the sulfate-reducing cultured representative *Desulfobacterium indolicum* (Bak and Widdel 1986) as a close relative, suggest the possibility of sulfate reduction near and below the nitrate depletion depth at SPG12. Additionally, a clone from 343-353 cmbsf was relatively closely related to *Spirochaeta smaragdinae,* a sulfur- and thiosulfate-reducing organotroph (Magot et al. 1997; Figure 9C), and some spirochaetes have been found to belong to a syntrophic sulfur-cycling community in gutless worms (Blazejak et al. 2005). The presence of members of the Alphaproteobacteria and Gammaproteobacteria in the 60-70 cmbsf, 213-223 cmbsf, and 343-353 cmbsf intervals also suggests the presence of an oxidative sulfur cycle: several clones are closely related to a putative sulfur-oxidizing Alphaproteobacterial oligochaete endosymbiont, AJ890098 (Blazejak et al. 2006) (Figure 9C). Thus, it is possible that a cryptic or low-activity sulfur cycle may be present alongside electron-accepting processes with higher standard Gibbs free energy, perhaps exploiting an unused set of substrates, or a substrate pool that is not associated with readily available metal oxide electron acceptors.

Archaeal and Bacterial species richness

The Archaeal and Bacterial clone libraries differed widely in the species diversity as estimated by rarefaction and Chao1 indices (Table 5; Figure 10). Archaea were found to be represented by 4-10 1% OTUs for all depths. In contrast, rarefaction curves of Bacteria did not reach saturation, and Chao1 estimates, which likely underestimate true diversity (Hughes et al. 2001), indicated several tens to several hundreds of bacterial 3% OTU for all depths (Table 5B). There are three possible explanations for this conspicuous difference in bacterial and archaeal OTU diversity:

1) If these results are a reasonable approximation of the true diversity of the Archaea and Bacteria in situ, highly restricted archaeal diversity might indicate that the oligotrophic sediment biome represents an extreme habitat for Archaea, and where only a few specialists thrive. This is consistent with the observation that the SPG archaea are distinct from the archaeal lineages abundant at organic-rich margin

sites (e.g., Biddle et al. 2006, 2008; Inagaki et al. 2006; Fry et al. 2008).

2) A second explanation is stochastic PCR bias. If SPG12 Archaea occur in significantly lower abundance than Bacteria, their low DNA and target gene concentration leads to stochastically biased PCR reactions, diminishing the species diversity detectable in clone libraries. Since bacterial and archaeal cell numbers and 16S rRNA gene copy numbers were not quantifiable due to the use of carrier DNA, this question remains open in this study.

3) A third possibility, not mutually exclusive of the previous two, is that the archaeal groups recovered from SPG12 may have substitutions in key priming regions, with only rare phylotypes being amenable to PCR with 8f/1492r primers. Several of the novel or uncultured lineages described in this work, specifically DSEG-1 through -4 and DHVE-5, are deeply divergent and highly diverse (Table 5). Along with DHVE-4 and DHVE-6, these lineages were previously found to form a long-branched monophyletic group distinct from other Euryarchaeota, termed DHVEG-II (Takai and Horikoshi 1999, Moussard et al. 2006). Deeply-branching archaeal lineages have mismatches to internal 16S rRNA gene primers; for example, a lineage in DHVEG-II clade (DHVE-6) was found to have the second-highest mismatch prevalence with a commonly used primer (Teske and Sorensen 2008). Highly selective PCR amplification of a few matching archaeal phylotypes from otherwise mismatched groups would produce nearly identical archaeal sequences multiple times from different depths, possibly under widely varying geochemical regimes (Figure 5; Appendix C). For example, the DSEG-2 phylotype SPG12_401_411_A1, comprising ~38% of clones from 401-411 cmbsf, also made up 5% of clones at 210-220 cmbsf; and the MBG-A2 phylotypes were found in abundance at both 273-283 and 343-353 cmbsf. The frequent recovery of identical archaeal phylotypes from different samples and PCR reactions is hard to explain by stochastic PCR bias only, but suggests the likelihood of primer bias.

Mismatched PCR primers might be one of the reasons why extant Genbank representatives of SPG12 lineages are rare: for example, there are relatively few MG-V, DSEG-2, DSEG-3, and DSEG-4 sequences in public databases (Table 4); the DSEG and DHVE-5 are also relatively uncommonly encountered. The range of habitats from which these sequences have been recovered, from hydrothermal vents to methanogenic coastal sediments (Figure 7), also suggests that sequence rarity is not simply a function of a lack of sampling of the appropriate habitat, but is consistent with poor detection in PCR. A

recently described novel bacterial phylum, the WWE3 division, illustrates the effects of PCR bias. Despite representing a significant fraction of the bacterial community in a highly-studied habitat, a sludge reactor, the WWE3 was represented by only a single sequence in public databases until its detection by metagenomic methods (Guermazi et al. 2008).

Nevertheless, it is possible that the DHVEG-II and MG-V euryarchaeota form a rare component of various habitats, but are most predominant under conditions such as those at SPG12. Ultimately, the explanation for the low Archaeal diversity may in fact be a combination of all three alternatives described above. More research is needed to explore whether the novel lineages detected at SPG12 are indeed abundant community members with substitutions in conserved regions of 16S, as suggested here by their high abundance in SPG12 clone library studies and phylogenetic placement. Newly developed, multiple PCR primer sets, as well as PCR-independent methods such as whole-genome amplification (Biddle et al. 2008), are needed to characterize the unique microbial communities in oligotrophic marine deep sediments.

SPG12 represents an oligotrophic, oxic/suboxic sediment environment that has so far received little attention as a distinct microbial habitat. Indeed, while the extended oxic and suboxic zones of SPG12 differentiate it geochemically from the anoxic conditions that generally characterize organic-rich sediments, there exist oxic/suboxic zones in organic rich sediments, albeit highly compressed. It remains to be seen whether the extended oligotrophic oxic/suboxic zones are microbiologically analogous to the compressed high-energy electron acceptor zones of organic rich sediments, which can be a key control on carbon burial efficiency (Hartnett et al. 1998).

Methods.

Site description and sample collection.

Samples were obtained during the Knox02-RR cruise in January 2007 from site SPG12, located in the margin of the South Pacific Gyre, 45 7.855 S, 163 11.051 W, at a water depth of 5306m, using a ~10 m gravity coring device. Total core recovery was 401 cm for core 1 (TG1) and 498 cm for core 4 (TG4). Trigger cores were split into approximately 1m-long sections, which were then further sectioned into ~30 cm rounds. Cores were sampled by pushing autoclaved cut-off 60 mL syringes into the core cross section, avoiding an approximately 1-2 cm margin inside of the core barrel edge, and then extruding sediment 1 cm

at a time while slicing off subsamples using a flame-sterilized spatula. Subsamples were frozen at -80 until analysis.

Geochemical and sedimentation data.

The sedimentation rate, calculated as a function of the age of the underlying crust and the thickness of the sediment column, was estimated to be ~1.8 m/Myr for SPG12. Cell counts and porewater chemistry measurements, including alkalinity, oxygen, DIN, and dissolved metals, were done as described elsewhere (D'Hondt et al., 2009).

Nucleic acid extraction, amplification and sequencing.

Genomic DNA for SPG12 samples was extracted from samples collected from 60-70 centimeters below the sediment surface (cmbsf), 153-163 cmbsf, 213-223 cmbsf, 343-353 cmbsf from core 1, and 401-411 cmbsf and 461-471 cmbsf from core 4, using a standard phenol-chloroform extraction protocol (Zhou et al. 1996) with some modifications. Blank extractions were also performed for all steps. To begin, 7-8 g of sediment was slurried in 7.5 mL of extraction buffer (100 mM phosphate, 250 mM sodium acetate, 50 mM EDTA, 50 mM NaCl, pH 8), and then added to pre-baked bead-beating canisters containing zirconium beads with a sediment-to-beads mass ratio of ~0.6-0.7. Beads used for all SPG12 extractions were a 60/40 mix of 0.1 mm and 0.2 mm low-protein-binding zirconium silicate beads (Shi 2005) (OPS Diagnostics). Approximately 30 µg of a carrier DNA, polydeoxyinosine-polydeoxicytosine (poly[dIdC]) (Barton et al. 2006), was then added. No carrier DNA was added to the SPG12 60-70 cmbsf sample, as higher cell densities nearer the sediment surface reduced the need for a carrier DNA to act as a blocking agent and to soak up nucleases to protect sample DNA.

Next, 15 mL of pH-8 buffered phenol and 875 µL of 20% sodium dodecyl sulfate was added, and the samples were subjected to bead-beating in a Braun MSK Cell Homogenizer for 20 s at high speed, followed by centrifugation for 10 minutes at 3500 times gravity (xg). The remainder of the extraction protocol followed a standard phenol-chloroform extraction and cleanup protocol (Zhou et al. 1996), with the exception that the sediment pellet was reextracted with a second volume of extraction buffer with no SDS added, and this second aqueous phase was processed through the same organic phases as the first, thus re-extracting the organic phases and reducing DNA loss. The crude DNA extract was precipitated

overnight at -20° C, after the addition of, in order, 2.5 ul/mL of UVC-irradiated GenElute linear polyacrylamide coprecipitant (LPA) (Sigma) to maximize precipitation of dilute nucleic acids, 0.5 volume of 7.5 M ammonium acetate (NH_4CH_3COO), and 1 volume of isopropanol. The precipitation was carried out at -20 ℃ overnight. Further purification of crude DNA extract was done by processing extract through a MoBio PowerSoil kit as per manufacturer's instructions, with the exception that Solution C1, used in the cell-disruption step, was omitted.

The primers A8f/A1492r and B8f/B1492r (Teske et al. 2002) were used to amplify an approximately 1500-base pair (bp) fragment of the 16S gene of Archaea and Bacteria, respectively, for both sample and blank extractions. Sample PCRs were carried out in triplicate and then combined to minimize "jackpot" amplification bias and to maximize yield. Each 25-µL PCR reaction contained 1 µL (SPG12 60-70 cmbsf, SPG12 343-353 cmbsf), 5 µL (SPG12 213-223 cmbsf, SPG12 401-411 cmbsf) or 15 µL (SPG12 461-471 cmbsf) of DNA template, 2.0 µL of dNTP solution (2.5 mM each dNTP), 0.250 µL (1.25 units) Takara Bio Inc. SpeedSTAR hot-start *Taq* enzyme, 2.5 uL of Takara Fast Buffer I (30mM Mg^{2+}), 1 µL of 100 mg/mL bovine serum albumin (BSA), and 1.5 uL each of 10 mM forward and reverse primers. The conditions for PCR were as follows: denaturation at 94 ℃, followed by 35 amplification cycles, each consisting of 5 s denaturation at 94 ℃, 15 s at the annealing temperature, and 20 s of elongation at 72 ℃, followed by one 10-minute elongation cycle at 72 ℃. Triplicate PCR reactions were then combined, target bands extracted and purified from a 2.0% low-melting-point (LMP) agarose (Promega) gel using SYBR Gold and the Invitrogen S.N.A.P. Miniprep kit, following the manufacturer's instructions. A gel fragment at approximately 1500 bp was also extracted and purified for the PCR and extraction negative controls, whether or not a band was present. All purified PCR products were further concentrated by precipitation with 0.5 µL of GenElute linear polyacrylamide (LPA) (Sigma). The pellets were then dried, and resuspended in 4-20 uL of PCR H_2O.

All SPG sequences, as well as closest GenBank relatives identified via BLAST searches, were aligned in ARB (Ludwig et al. 2004) with final adjustments made by eye. Initial NJ trees were made in ARB, containing the GenBank and NCBI sequences that would be included in the final phylogenies. These datasets were then checked for chimeric sequences using the GreenGenes pipeline, using a window size of 200 bp, a "similarity to core set threshold" set to 99%, and the divergence-ratio set to 1.10, the last

indicating that a sequence will be identified as a chimera if fragments derived from it have 'parent' sequences in the database that are more than 10% divergent from each other (Huber et al., 2004). MG-I Crenarchaeota sequences were additionally submitted to three rounds of a more stringent chimera-check, with the same settings as above but a divergence ratio of 1.02.

Phylogenetic analysis was conducted separately for the MG-I Crenarchaeota, other Crenarchaeota, and Euryarchaeota sequences. Bacteria phylogenies were estimated first in ensemble in ARB and then split into three datasets. Operational taxonomic units (OTU) were defined using DOTUR (Schloss and Handelsman 2005), and a single representative sequence was then selected from each OTU to be included in phylogenies. Final phylogenetic analysis used 99% identity OTUs for Archaea, while 97% OTUs were used for Bacteria. Modeltest 3.7 (Posada and Crandall 1998) was used to estimate the best model of sequence evolution according to the Akaike information criterion (AIC), and was always a general time-reversible (GTR) model with a proportion of invariant sites and a gamma distribution of evolutionary rates (GTR+I+G). For the non-MGI crenarchaeotes and the euryarchaeotes, the maximum likelihood tree was then estimated using a GTR+I+G model of sequence evolution in TreeFinder version October 2008 (Jobb et al. 2004), which automatically calculates gamma distribution and other parameters. For these two trees, ML bootstrap support was estimated using 1000 replicates, and values above 50% were inscribed in the ML phylogeny. The Bacterial phylogenies were estimated using the neighbor-joining algorithm (Saitou and Nei 1987), a maximum-likelihood-estimated model of evolution and a gamma-corrected rate distribution, in MEGA 4.0 (Kumar et al. 2008). Branch support estimated using a 2000-replicate interior branch (IB) test (Nei and Kumar 2000), which uses a hypothesis-testing approach to determine the probability that a particular interior branch has a length greater than 0, i.e., that the node in question should not be collapsed into a polytomy. Values greater than 95% should be considered strong support for a particular node (Nei and Kumar 2000).

Treeclimber (Schloss and Handelsman 2004) was used to test the hypothesis that the overall structure of two communities is statistically the same. P values indicate the probability that a similar association between phylogeny and sample origin would occur by chance (Schloss and Handelsman 2006). DOTUR was used to calculate rarefaction curves, while SONS (Schloss and Handelsman 2006) was used to estimate the Chao1 (Chao 1984) species diversity of each depth, as well as cross-comparisons of

community structure using abundance-based Jaccard (Smith et al. 1996, Yue et al. 2001) and Theta (Yue and Clayton 2005) diversity indices. Both individual community Chao1 estimates and estimates of the Chao1 diversity shared between two communities (shared Chao1) were calculated. The abundance-based Jaccard index measures the intersection of the combined assemblage of phylotypes, and can be considered the probability that given a phylotype present in one assemblage, it is present in both assemblages (Schloss and Handelsman 2006). By contrast, ThetaYC corrects for the differences in relative abundance of phylotypes in each assemblage, taking into account both the incidence of shared phylotypes and the similarity of their relative abundances in each assemblage. Thus, it can be considered a percentile measure of the similarity of community structures.

Chapter 3

Abyssal sediment microbial communities: what lives in the other 90% of the seafloor?

Introduction

Marine sedimentary microbial communities are key mediators of global biogeochemical cycles (e.g., Wellsbury et al.2002, D'Hondt et al. 2002, 2004). The Domain Archaea accounts for a large portion, perhaps the majority, of deep subsurface prokaryotic cells and biomass (Biddle et al. 2006, Lipp et al. 2008), which in turn constitutes $1/10^{th}$ to $1/3^{rd}$ of global biomass (Parkes et al. 1994, Whitman et al. 1998). The majority of studies thus far have focused on relatively organic-rich deep subsurface sediments (e.g., Parkes et al. 1994, Wellsbury et al. 2002, Reed et al. 2002, D'Hondt et al. 2004, Parkes et al.2005, Sørensen and Teske 2006, Biddle et al. 2006, Inagaki et al. 2006, Kendall et al. 2007, Heijs et al. 2008). However, abyssal sediments >2000m water depth cover a much larger extent of the ocean floor (~89%; Dunne et al. 2007) and, in contrast to margin or coastal sediments, are generally oligotrophic, with very low organic carbon content (<1%) and slow rates of deposition (Seiter et al. 2004, Dunne et al. 2007). Electron acceptors such as oxygen or nitrate penetrate these oligotrophic sediments on a scale of meters (D'Hondt et al. 2004) or tens of meters (Gieskes and Boulègue 1986, D'Hondt et al. 2009), in contrast to organic-rich continental margin or shelf sediments where these strong electron acceptors are used up within centimeters. This expansion of the oxic/suboxic zone in oligotrophic sediments is a function of the slow rates of carbon deposition. The combination of higher-energy electron-acceptor type and slower flux of electron donor substrates likely imposes distinct constraints on life in oligotrophic marine sediments, which cover the majority of the surface of Earth.

Several phylum-level uncultured archaeal lineages have been identified as typical, in some cases possibly endemic, deep-subsurface sediment-associated groups (Inagaki et al. 2003, 2006; Parkes et al. 2005; Biddle et al. 2006, 2008; Teske and Sørensen 2006, 2008; Fry et al. 2008). These include the Marine Benthic Group B (MBG-B, Vetriani et al. 1999), a deeply-branching phylum-level lineage; Miscellaneous

Crenarchaeotal Group (MCG, Inagaki et al. 2003), a diverse crenarcheotal lineage without close affiliation to other Crenarchaeota; South African Gold Mine Euryarchaeotal Group (SAGMEG, Takai et al. 2001a); and Marine Benthic Group D (MBG-D, Vetriani et al. 1999), a euryarchaeotal group affiliated with the Thermoplasmatales. All of these are approximately phylum-level in divergence, with the exception of the MBG-D, which groups along with the MG-II (DeLong et al. 1992) and MG-III (Furhman and Davis 1997) in a well-supported phylum-level clade also containing the Thermoplasmatales (e.g., Durbin 2009). However, our current datasets on archaeal community composition in deep marine sediments are biased towards organic-rich continental margin sediments (Teske and Sørensen 2008). Relatively few studies have surveyed the archaeal diversity of abyssal or ocean gyre sediments to date (a more or less comprehensive list includes Vetriani et al. 1999, Inagaki et al. 2001, Sørensen et al. 2004, Wang et al. 2004, Nercessian et al. 2005, Wang et al. 2005, Xu et al. 2005, Gillan and Danis 2007, Li et al. 2008, Tao et al. 2008), with even fewer examining the oligotrophic deep subsurface; thus, the available database for archaeal communities in oligotrophic marine subsurface sediments has not yet reached the same coverage as eutrophic sediments. Nonetheless, initial datasets from the South Pacific Subtropical Gyre (Durbin, 2009) and a few other datasets in the literature point to profoundly different archaeal communities, with little or no overlap at the phylum and subphylum level.

The highly divergent geochemical setting of oligotrophic compared to eutrophic sediments, with a large expanse of high-energy electron acceptor zones, additional possible metal-cycling reactions (see Durbin 2009 for discussion), and slow rates of input of recalcitrant organic carbon (e.g., Wakeham et al. 1997, Hedges et al. 2001), also presents an opportunity to investigate the association of diversity and function in marine sedimentary Archaea. In particular, the redox metabolisms pursued by heterotrophic or chemotrophic organisms in oligotrophic sediments are expected to be different. Since the most-energetic available electron acceptor (MEEA) defines the maximum Gibbs free energy available from oxidizing a given substrate, lineages that specialize in the MEEA will have more energy to devote to growth and reproduction, and are expected to predominate over competing organisms that specialize in less-energetic electron-accepting processes. Thus, the appearance of certain phylogenetic groups in sediments of a particular redox state may inform on how subsurface archaeal lineages assort according to metabolism.

To approach questions regarding the nature and function of archaeal diversity in the marine

subsurface, the current study examines the available molecular and geochemical data from marine sediments of a range of trophic states. Based on research to date, the expectation would be that the archaeal phylum- or order-level lineages identified as dominant subsurface lineages in margin sediments (Teske and Sørensen 2008) will also dominate in abyssal, oligotrophic sediments, albeit performing different metabolisms.

Sediment sampling depths and degree of nutrient limitation vary among oligotrophic sediment studies, while dissolved constituent data is not measured in some cases. Several sites that might be characterized as oligotrophic also appear to have hydrothermal characteristics (e.g., Inagaki et al. 2001, Nercessian et al. 2005, Li et al. 2008), lessening their utility in defining the oligotrophic sediment habitat. Additionally, oligotrophic sediments often share oxygen- or nitrate-respiring activity with the overlying water column, blurring the distinction between oligotrophic sediments and the water column as independent microbial habitats. Several studies of oligotrophic abyssal sediments (Wang et al. 2004, Xu et al. 2005, Gillan and Danis 2007) sampled only 3-12 cm below the seafloor (cmbsf) of the sediment column and recovered the same archaeal phylum, the Marine Group I Crenarchaeota, that is presumed dominant in the overlying water column (e.g. Karner et al. 2001, Church et al. 2003, Agogué et al. 2008). Because of the high potential for cross-contamination and the phylogenetic similarity between oxic sediments and the overlying water column, to confidently label MG-I clones collected from oxic sediments as indigenous requires stringent contamination controls and/or investigations of the diversity of contamination sources (e.g., Durbin, 2009). Due to these difficulties, MG-I sequences were excluded from the current analysis.

Of the remaining publications, studies of sedimentary archaeal communities from the central South China Sea (Tao et al. 2008), the eastern margin of the South Pacific Gyre (Sørensen et al. 2004), the equatorial upwelling zone west of the Galapagos (Teske and Sørensen 2008, Teske et al. 2005), the Mid-Atlantic Ridge (Nercessian et al. 2005), and the East Pacific Rise (Li et al. 2008) were included. To provide a eutrophic sediment frame of reference, ODP Leg 201 sites 1227 and 1228 from beneath the highly productive Peruvian upwelling zone and the methane-clathrate-bearing Peru Trench site 1230 (Biddle et al. 2006, Sørensen and Teske 2006) were considered. In addition, methane seep sediments from the Mediterranean provide an example of a pelagic, deep-water, yet organic-rich shallow sediment environment (Heijs et al. 2008).

1. The sedimentary trophic state spectrum

In the following section, we summarize some of the most informative biogeochemical and microbiological parameters for oligotrophic marine subsurface sediments.

Sedimentation rates. Sedimentation rates can approximate sediment trophic states, albeit with some exceptions for high-carbonate or turbidite-associated sedimentation and changing oceanographic conditions in the surface ocean, which can all enrich or impoverish deep marine sediments relative to the expected trophic state based on surface ocean productivity, or if a significant fraction of sedimentation is inorganic carbonate or silicate. Among ODP Leg 201 sites, Site 1231 is decidedly the slowest-accumulating site, followed by 1225, while 1230 followed by 1229 were calculated to have the highest sedimentation rates (D'Hondt et al. 2003; Table 6). There is overlap in the sedimentation rates for sites 1225, 1226, 1227 and 1228, although their trophic states differ. Sedimentation rates at SPG Sites 11 and 12, at 0.9m/My and 1.8m/My, respectively, are lower than those at site 1231 by a factor of 4-10 for SPG11 and 2-5 for SPG12 (Table 7). The sedimentation rate at SPG11 is 200-fold slower than that at Site 1230, dramatically illustrating the different constraints shaping eutrophic margin environments and ultraoligotrophic gyre sediments.

Porewater DIC and organic carbon availability. Concentrations of porewater DIC above or below the mean seawater concentration are an indicator of the magnitude and direction of net metabolism (D'Hondt et al. 2003). Net heterotrophy due to remineralization of organic matter to CO_2 increases DIC concentrations, at least at the more oligotrophic end of the scale. In highly organic-rich sediments, a large fraction of organic carbon may be remineralized not to CO_2 but to acetate and/or methane. Comparisons of maximum DIC for Leg 201 sites suggest a hierarchy of trophic states, from highly eutrophic Site 1230, to moderately eutrophic sites 1227, 1228, and 1229, to mesotrophic Site 1226, to oligotrophic Sites 1231 and 1225 (D'Hondt et al. 2003; Table 6). Alkalinity and not DIC was measured at SPG sites 11 and 12, but since DIC composes ~96% of total alkalinity at seawater pH, these values are included for comparison in milliequivalents/liter (meq/L; Table 6; D'Hondt et al. 2009). By this measure, SPG sites 11 and 12 are more oligotrophic than sites 1231 and 1225, even taking into account the ~4% discrepancy between total alkalinity and DIC. Alkalinity remains at seawater values throughout the sediment column at SPG11, while

SPG12 displays a depthwise increase in DIC comparable in magnitude and rate of change to the rate of oxygen depletion.

Organic carbon availability, like DIC, is also not a direct measure of sediment trophic state, even though it represents the available pool of reduced substrates. Substrate lability and organic carbon residence time can vary between sediments with similar organic carbon contents, thus affecting the trophic state of the sediment without necessarily changing the sediment organic-carbon percent weight. This leads to some degree of overlap in organic carbon content between sites of possibly different trophic state. Nevertheless, organic carbon percent weight can be taken as a rough indicator of trophic state. Among the nutrient-limited sites, organic carbon content values were available only for SPG11at the time of writing. They ranged from 0.59% to 0.45% over the upper 9 cmbsf, with little change after the upper 2 cmbsf (Table 6). Abyssal sediments, which include the remaining oligotrophic sites considered here, were modeled to have 1% or less organic carbon by weight (Seiter et al. 2004). All eutrophic sites varied between ~1-10%, with ~4% most typical for the majority of the sediment column (Table 6), and typically decreasing with depth.

Cell densities.

Cell densities at the SPG sites are the lowest yet recorded for any equivalent depth horizon (D'Hondt et al., 2009). For all sites, cell counts declined with depth, with the lowest value typically being the deepest. Cell counts ranged from $10^{3.9}$ to $10^{6.6}$ per cm^3 over the upper 2.8 m at SPG11, and $10^{5.4}$ to $10^{6.3}$ per cm^3 over the upper ~5 m at SPG12, again setting SPG11 apart (Table 6). Sites SPG12, 1231, and 1225 had similar cell counts, over a much broader depth range, with a lowest value of ~10^5 for both ODP sites and a deepest measurement of 81.6 mbsf for 1231 and 320 mbsf for 1225. Thus, cell counts were lower for the oligotrophic ODP sites than SPG12, but they were higher at the equivalent depth horizon. Cell counts tended to range from 10^6-10^{10} for the ODP mesotrophic and eutrophic sites over a similar depth range as 1231 and 1225, with no trend in lowest recorded value between the meso- and eutrophic sites.

Oxygen, nitrate and sulfate gradients. Most-energetic available terminal electron acceptor is a key defining parameter of sediment trophic state, since it determines the maximum energy yield per mole of substrate respired. Typically, electron acceptors either diffuse in from overlying seawater (oxygen,

nitrate, sulfate), or are deposited via sedimentation (metal oxides; also buried porewater dissolved constituents such as oxygen, nitrate and sulfate). As such, the pattern of electron acceptor depletion generally occurs depthwise in order of declining energy yield, and at a rate proportional to the rate of organic substrate delivery (D'Hondt et al. 2002, 2004). Thus, the most nutrient-limited sites are expected to have the broadest availability of the highest-energy electron acceptors. At site SPG12, oxygen declines exponentially from 114.76 µM at 5 cmbsf to ~2 µM at 179 cmbsf, while nitrate declines from 43.8 µM at 10-15 cmbsf to 0 µM at 253-258 cmbsf, suggesting drawdown of oxygen followed by depletion of nitrate (Table 7, Figure 5). Nitrate drawdown is not evident in geochemical profiles for the ultraoligotrophic SPG11, and in fact, nitrate may be produced in sediments there (Durbin et al. 2009), while oxygen remains at ~160 µM at 280 cmbsf. This likely indicates oxygen is not depleted throughout the sediment column, unless there are deep biotic or abiotic sediment oxygen sinks. Although oxygen data are not available for ODP sites, only sites 1225 and 1231 have nitrate depletion on the scale of centimeters: 10-40 cmbsf for 1231, and ~150 cmbsf for site 1225, a site that includes a significant amount of carbonate bulk not found at site 1231; oxygen is expected to be depleted at shallower depths than nitrate.

By contrast, most of the Mediterranean eutrophic deep-sea methane seep sediments had no measurable oxygen penetration, except that minute concentrations were available above 6 cmbsf at the Kazan mud volcano (Table 6). Likewise, nitrate was not measurable in two of the three mud volcano sites, while nitrate depletion occurred between 6 and 22 cmbsf for the Kazan mud volcano (Table 6). It is expected that the mesotrophic and eutrophic ODP sites have similar oxygen and nitrate penetration depths as the methane seep sites (D'Hondt et al. 2002, 2004). Finally, sulfate was not measurably drawn down at either SPG12 site, nor at the oligotrophic ODP sites 1231 and 1225 (Table 6). Mesotrophic 1226 displayed sulfate drawdown but not depletion over the entire sampled sediment column, along with apparent metal redox cycling to at least ~70 mbsf and methanogenesis throughout nearly the entire sediment column. All Peru Margin and Peru Trench sites (1227, 1228, 1229, 1230) displayed sulfate depletion within meters to tens of meters of a diffusive source (i.e., water column or basement basalt; D'Hondt et al. 2004).

Dissolved constituent data was not collected for the South China Sea, Mid-Atlantic Ridge, or East Pacific Rise sites. However, solid-phase chemical data is available for the hydrothermal-precipitate sediments at the Mid Atlantic Ridge (Nercessian et al. 2005). Notably, solid-phase reduced iron and

manganese, in the form of mangosanite and siderite, increase downcore over the upper ~40 cmbsf. This suggests reduced metals are oxidized towards the sediment surface (e.g. Severmann et al. 2006), which would require a high-energy electron donor such as nitrate or oxygen and indicating a relatively deep penetration of these high-energy electron donors.

The sites reviewed above appear to fall into several natural groups. In terms of electron acceptor gradients, the tens-of-centimeters to meters scales of oxygen and nitrate depletion, coupled with a lack of any sulfate drawdown, unite the oligotrophic sites SPG12, 1231, and 1225, with oligotrophicity decreasing in that order; the lack of oxygen depletion at SPG11 sets this ultraoligotrophic site apart. Mesotrophic site 1226 displayed sulfate drawdown but not depletion over an ~400 m sediment column, while eutrophic sites have oxygen and nitrate penetration depths of centimeters to millimeters, and sulfate depletion depths of tens of meters or less. DIC tracks sediment trophic state closely, as concentrations above background represent heterotrophic activity (D'Hondt et al. 2003); maximum DIC values range from seawater concentrations for ultraoligotrophic, to >10 mM for oligotrophic sites, and significantly more for eutrophic sites. Other parameters are often somewhat noisier and reflect a less direct relationship to sediment trophic state.

2. Sampling methods.

Archaeal 16S rRNA gene sequences from SPG12 were extracted and amplified from sediments subsampled from gravity core sections using a sterilized cut-off syringe, at a sampling resolution of 10 cm. Archaeal 16S sequences for the Peru Margin ODP sites were derived from the upper sulfate-methane transition zone (SMTZ) at site 1227 (37.8 mbsf, Biddle et al. 2006; 35.35, 34.25, 37.75 and 40.35 mbsf, Sørensen and Teske 2006), as well as intervals above and below the SMTZ (6.55mbsf, 7.35 mbsf, 21.35 mbsf, 45.35 mbsf, and 49.85 mbsf; Sørensen and Teske 2006) (Table 7). Sequences from sites 1229 and 1230 were derived from the upper and lower SMTZs at 29.4 and 86.8 mbsf respectively for site 1229, and from the only SMTZ at site 1230, at 11.0 mbsf (Biddle et al. 2006). Finally, 1226 was sampled from approximately 1.3, 7.2, 26.2, and 45.2 mbsf, all intervals where methane and sulfate profiles suggest both methanogenesis and sulfate reduction occurs, while dissolved manganese is present until ~100 mbsf. ODP

Table 6.Comparison of parameters relating to trophic state of sediments.

Site	Depth, mbsf	Sedimentation rate, m/Ma [DSDP/ODP site]	TOC, % weight of sediment	Ammonia, µM	Oxygen depletion depth	Nitrate depletion depth	Sulfate depletion depth	Max. DIC, mM	Log[10] cell counts, range	mbsf for cell counts
SPG11 [a]	0.01 – 3.07	0.9	0.594 – 0.45 (0.0 – 0.09 mbsf only) [a]	BD	No depletion	No depletion	No depletion	2.44 meq (Alkalinity)	3.87 – 6.63 [d]	2.96 – 0.1
SPG12 [a]	0.60 – 4.98	1.8	N/A	ND	~1.79 mbsf	2.55 mbsf	No depletion	3.23 meq (Alkalinity)	5.37 – 6.3 [d]	4.9 – 0.25
ODP Leg 201 1231 [b]	N/A	Eocene-Quaternary: 4-10 [DSDP 321][e]	N/A	7.64 – 32.72	n.a.	0.1 – 0.4 mbsf	No depletion	3.64	5.14 – 8.24	81.64 – 0.01
ODP Leg 201 1225 [b]	N/A	Miocene-Quaternary: 15-65 [ODP 851][f]	N/A	6.15 – 38.48	n.a.	~1.5 mbsf	No depletion	3.98	5.26 – 6.87	320 – 0.85
ODP Leg 201 1226 [b]	0.0 – 422	Miocene-Quaternary: ~40 [ODP 846][g]	0.5– 2	n.a.	n.a.	n.a.	No depletion	7.03	~6 – 8.67	420 – 0.01
ODP Leg 201 1227 [b]	4.5 – 106.9	Miocene-Quaternary: 20-50 [ODP 684][g]	1.16– 10.57	3168.2 – 8875.5	n.a.	n.a.	~41 mbsf	25.78	6.25 – 7.15	151 – 0.15
ODP Leg 201 1228 [b]	4.5 – 78	Miocene-Quaternary: 22-30 [ODP 680][g]	3.01– 8.78	n.a.	n.a.	n.a.	~40 mbsf	19.81	n.a.	n.a.
ODP Leg 201 1229 [b]	3 – 126	Quaternary: ~80 [ODP 681][g]	0.83– 3.97	5335.69 – 6746.53	n.a.	n.a.	~32 mbsf	21.14	6.43 – 9.97	185.7 – 90.45
ODP Leg 201 1230 [b]	3 – 450.5	Miocene-Quaternary: 100-250 [ODP 685][g]	1.27– 3.98	n.a.	n.a.	n.a.	~9 mbsf	162.95	6.1 – 8.62	257.7 – 0.01
Mediterranean cold seep Napoli [c]	0.0 – 0.28	N/A	3.95 – 1.22	217.2 – 40.6	0.0 mbsf	0 mbsf	n.a.	n.a.	n.a.	n.a.
Mediterranean cold seep Kazan [c]	0.0 – 0.34	N/A	2.59 – 174	30.0 – 158.5	0.0– 0.6 mbsf	0.06 – 0.22 mbsf	~0.35 – 0.40 mbsf	n.a.	n.a.	n.a.
Mediterranean cold seep Amsterdam [c]	0.0 – 0.31	N/A	4.31 – 3.9	50.6 – 152.5	0.0 cmbsf	0.0 cmbsf	n.a.	n.a.	n.a.	n.a.
East Pacific Rise [d]	0.05 – 0.40	N/A	1.1 +/– 0.23 – 0.64 +/– 0.20	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

Table 6. Comparison of parameters relating to trophic state of sediments, ordered using oxygen/nitrate/sulfate penetration depth, except for carbonate-rich 1225, where max DIC is used as a trophic state indicator. Sites included are an ultraoligotrophic site (SPG11), defined as sediments without depletion of oxygen; oligotrophic, with a nitrate depletion depth of tens of centimeters to meters; and eutrophic, with extremely shallow nitrate and oxygen depletion depths on the order of millimeters or centimeters, relatively high maximum DIC, sediment accumulation rates, and cell counts. Note that the surficial 1– 2m is often lost or disturbed during coring in the case of the DSDP/ODP samples, so nitrate depletion depths were often not measured, and oxygen profiles were not yet practical to measure in marine sediments; however, oxygen depletion typically occurs within millimeters to a few centimeters for these sites, with nitrate extending slightly deeper. Depths are approximate if depletion depth was in between measurements or if values never reach zero due to measurement artifact. Alkalinity, which is ~96% of DIC at seawater pH, is included instead of DIC for SPG11 and SPG12. DSDP/ODP sediment accumulation rates are from locations also targeted during ODP Leg 201, and are based on ages of boundaries identified biostratigraphically; SPG rates were calculated by dividing depth of column by age of underlying basalt. [a]D'Hondt et al. 2009 [b]ODP Leg 201: D'Hondt et al. (2003). [c]Heijs et al. (2008). [d]Li et al. (2008). [e]Quilty et al. (1976) [f]Mayer et al. (1992) [g]Suess et al. (1988)

sediments were collected and subsampled while avoiding edge contamination and using sterile methods (Biddle et al. 2006, Lever et al. 2006).

Archaeal 16S rRNA gene sequences for the Mediterranean cold seeps were derived from the upper 20-30 cm sediments, which were subsampled from a box core using aluminum cores and divided into 2 or 3 subsections before freezing (Heijs et al. 2007). Sediments for the Rainbow Vent Field, Mid-Atlantic Ridge site were cored via submersible-operated push core, and then extruded and subsectioned into three parts; archaeal 16S gene sequences were amplified from the middle, 7.5-15 cmbsf interval, in a carbonate (siderite and mangosanite)-rich zone (Nercessian et al. 2005). Sequences from the South China Sea were amplified from an extraction of a 10-cm subsample of a box-core (Tao et al. 2008), while sediments from the East Pacific Rise were initially sampled via box-core and subsectioned into 5-10cmbsf, 15-20 cmbsf, 25-30 cmbsf, and 35-40 cmbsf intervals (Li et al. 2008).

Diversity survey methods.

While all studies used slightly different nucleic acid extraction protocols (Table 8), most involved chemical cell-membrane disruption with sodium dodecyl sulfate (SDS), mechanical membrane disruption with bead beating, and phenol-chloroform extraction (e.g., Zhou et al. 1996). The sites with the largest sample size and most extensive geochemical data, i.e. SPG12, 1231, the Peru Margin, and the Mediterranean cold seeps, all used some variation of a SDS/bead beating/phenol-chloroform based extraction protocol, although some sequences from 1231 were only amplified using an enzymatic-lysis+SDS based extraction (Table 7). Primers used for 16S rRNA gene amplification (Table 7) differed between the oligotrophic abyssal-sediment and the organic-rich endmember environments of the Peru Margin and Mediterranean cold seeps, although some of the cold seep clone libraries shared a reverse primer with the clone libraries from the mid-ocean ridge and South China Sea sites, and 8f (Teske et al. 2002) was the forward primer for most studies. Finally, although amplicon size has been shown to influence clone library composition, all amplicons considered here (Table 7) fell in a size range within which amplicon size variation minimally impacts clone library composition (Huber et al. 2009).

3. Deep phylogeny of oligotrophic Archaea.

Previous efforts to reconstruct the phylogeny of lineages abundant in oligotrophic sediments have suffered from several difficulties. One, the problem of long-branch attraction: the extremely deep or long branches of some oligotrophic sediment lineages are difficult to place in relation to recognized uncultured lineages from organic-rich sediments. This leads to long-branch attraction (LBA) artifacts, wherein even well-supported, recognized lineages that have long branches (e.g., the DSEG-#/DHVE-# lineages) can cause incorrect placement of other long branches with a less-clear branch point. Hyperthemophilic lineages can also attract each other, and are themselves attracted to the base of the phylogenetic tree, due to their convergence to high G+C content of rRNA (Boussau and Gouy 2006, Brochier-Armanet et al. 2008). Secondly, the commonly used 16S rRNA gene is simply not long enough to reliably identify the branching order of deep lineages, a problem exacerbated when short, partial sequences are used in phylogenetic analysis (e.g., Takai et al. 2001a, Takai et al. 2001b). Many of the deeply branching Archaeal groups have multiple large insertions in the 16S rRNA gene, which sometimes causes conserved regions to be misaligned in alignments provided in the SILVA database reference tree. Finally, the definition of a given clade may vary between studies and variable levels of statistical support are provided, a problem often accompanied by the proliferation of different names for the same group, as exemplified by the DHVEG-II lineages originally identified by Takai and Horikoshi (1999) (Figure 11). As an attempt to address some of these difficulties, we constructed phylogenies of the Euryarchaeota (Figure 11) and Crenarchaeota (Figure 12), using a long alignment guided by conserved sequence motifs and secondary structure, and a 60% base consensus filter so that only relatively conserved sites are considered in analysis.

4. Archaeal occurrence trends across sites.

Relative abundances of deeply-branching Archaeal lineages in 16S rRNA gene clone libraries from oligotrophic and organic-rich sites are included in Figure 13. All clades found at the organic-rich endmember Peru Margin and Mediterranean mud volcano sites are presented in shades of red or orange in Figure 13, with the remainder colored blue.

The Peru Margin (ODP sites 1227-1229) and Peru Trench (ODP site 1230), and the Mediterranean methane seep sites represent the eutrophic endmember sites for this study (Table 6), with roughly overlapping geochemical parameters, albeit with a more active methane cycle at the methane seep site. A

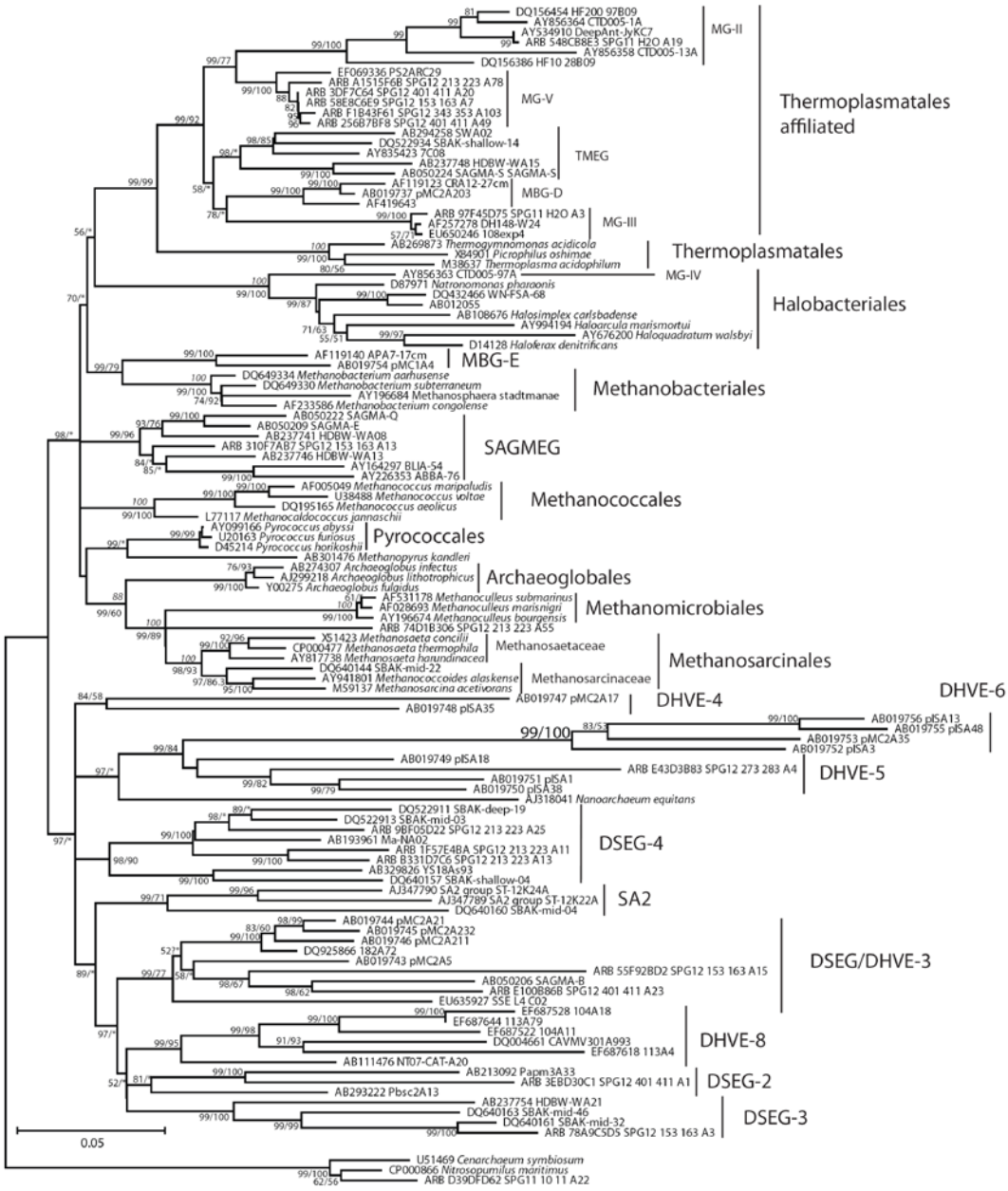


Figure 11. Neighbor-joining (NJ) 16S rRNA gene phylogeny of Euryarchaeota, using a 972-bp alignment originally derived from a 1247-bp alignment that was subjected to a 60% consensus base filter in ARB, so that only sites with at least a 60% conserved base were analyzed phylogenetically. A maximum-composite likelihood distance model with gamma rate correction was used for the NJ analysis, with 1000 iterations of the interior branch algorithm as clade support values (Sitnikova et al. 1995), in MEGA4 (Tamura et al. 2007). This alignment was also subject to ML analysis using TreeFinder (Jobb et al. 2004), using a general time-reversible substitution model with a proportion of invariant sites and gamma-distributed site rates (GTR+I+G). 1000 iteration maximum-likelihood bootstrap values are the rightmost values above. Additionally, ML bootstrap support values from a phylogenetic analysis of 50+ concatenated genes per taxon are provided above, in italics, when the same monophyletic group is supported in both NJ and concatenated-gene ML phylogenies (Brochier-Armanet et al. 2008).

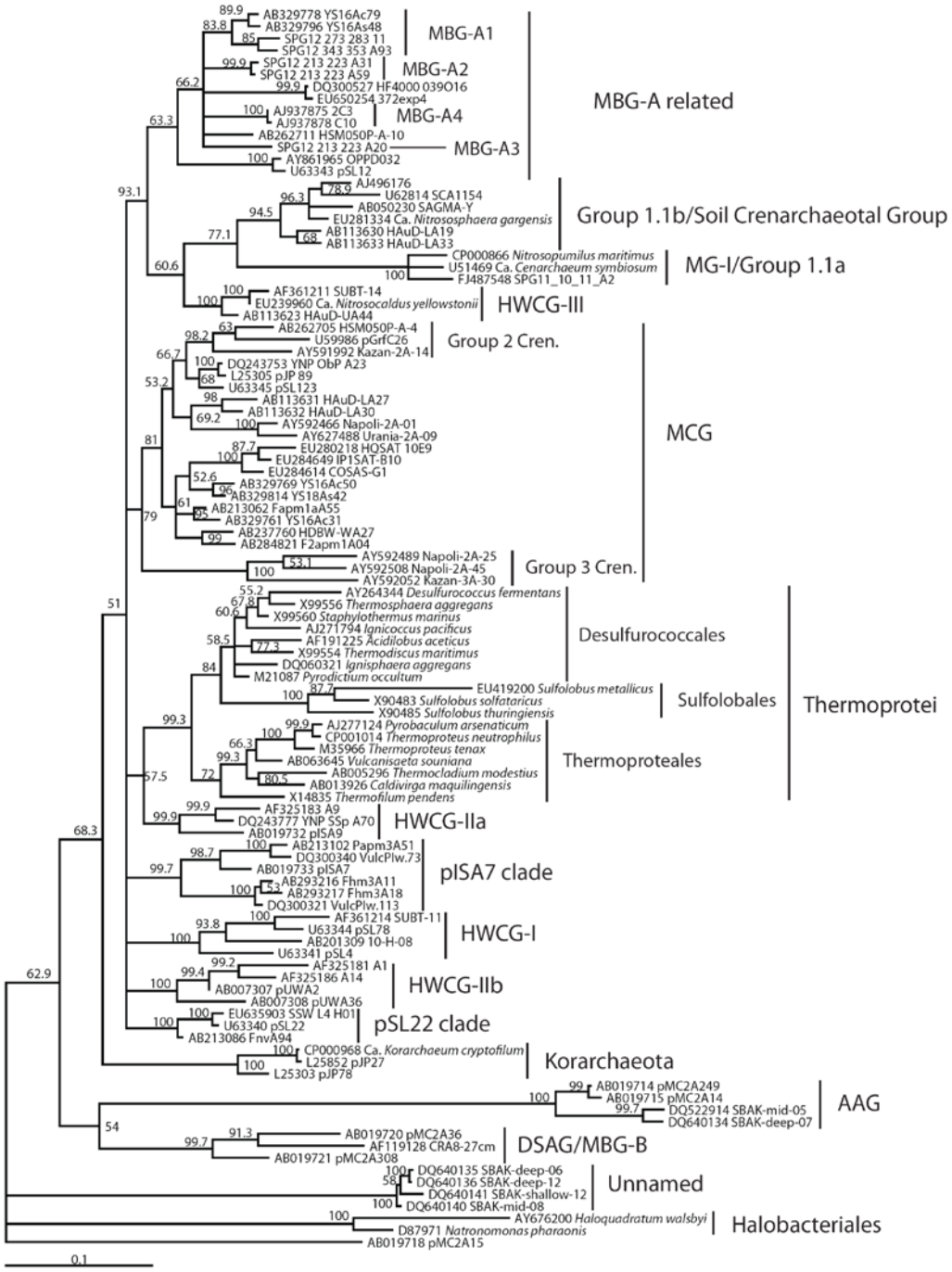


Figure 12. Maximum-likelihood phylogeny of cultured and major uncultured Crenarchaeotal lineages found in oligotrophic sediments, using a GTR+I+G substitution model in TreeFinder (Jobb et al. 2004). Support values are based on 1000 maximum-likelihood bootstrap iterations. Alignment is originally 1247-bp long, and was trimmed to 968 sites by imposing a 60% consensus-base filter, as in Figure 11. Hot Water Crenarchaeotal Group I, II and III (Nunoura et al. 2005) are additional distinct uncultured phylogroups within the crown radiation of Crenarchaeota. MBG- A4 and pISA7 clade are novel lineages; see text for other references.

Table 7. Relevant environmental, sampling and methodological features for the archaeal clone library studies compared in this review.

Study site [reference]	# sampling intervals (# different cores sampled)	Sampled depth range, mbsf	Electron accepting environment	Extraction protocol	DNA or RNA?	Forward Primer	Reverse Primer	Primer annealing, °C	Pre-screening?	Number of clones, MG-I excluded
SPG12 [a]	6 (1)	0.60 — 4.11	Oxic/suboxic	S, BB, PC	DNA	8f	1492r	55	No	235
ODP Site 1231 [b]	3 (1)	1.8 — 43	Suboxic	S, BB, PC; S, EL, PC	DNA	8f 8f	1492r 915r	58	No	64
ODP Site 1225 [c]	2 (1)	9 — 16	Suboxic	S, EL, PC	DNA	8f 8f	1492r 915r	58	No	18
ODP Site 1226[j]	4 (1)	1-45 mbsf	Suboxic/anoxic	S, BB	DNA	21f	915r	58	No	33
Peru Margin/Trench [d], [e]	13 (4)	6.55 — 86.8	Anoxic	S, BB, PC	RNA	8f	915r	58	No	572
Medit. Cold Seeps [f]	9 (3)	0.10 — 0.34	Anoxic	S, BB, EL, PC	DNA	2f	1406r, 958r	57.5	No	237
S. China Sea [g]	1 (1)	0 — 0.10	Unknown	S, HT, HS, EL, PC	DNA	21f	958r	55	T-RFLP	61
Mid-Atlantic Ridge [h]	1 (1)	0.075 — 0.15	Unknown	S, BB	DNA	21f	958r	53	No	60
East Pacific Rise [i]	4 (1)	0.05 — 0.40	Unknown	S, BB	DNA	21f	958r	58	No	67

Table 7. Relevant environmental, sampling and methodological features for the archaeal clone library studies compared in this review. Number of different sampling intervals is total number of independent intervals for all cores. Extraction protocol column lists key features of nucleic acid extractions: S, use of sodium dodecyl sulfate as a membrane disruption agent; HT, high temperature (membrane disruptor); HS, high salt (membrane disruptor); BB, using of bead-beating as means of cell lysis; EL, enzymatic cell lysis; PC, phenol chloroform extraction and purification. If nested amplification was used, nested primer set is indicated by carrot; multiple primer sets used in parallel separated by comma. Pre-screening indicates whether entire clone libraries were sequenced, or phylotype abundances in clone libraries were extrapolated based on T-RFLP screening and sequencing of unique T-RFLP profiles. Primer references: 8f, Teske et al. (2002); 2f, 21f, 915r, 958r, DeLong (1992); 1406r, 1492r, Lane (1991).[a]Chapter 2, this thesis [b]Sorensen et al. (2004)[c]Teske and Sorensen (2008)[d]Biddle et al. (2006)[e]Sorensen and Teske (2006)[f]Heijs et al. (2008)[g]Tao et al. (2008)[h]Nercessian et al. (2005)[i]Li et al. (2008)[j]Klingensmith et al., unpublished

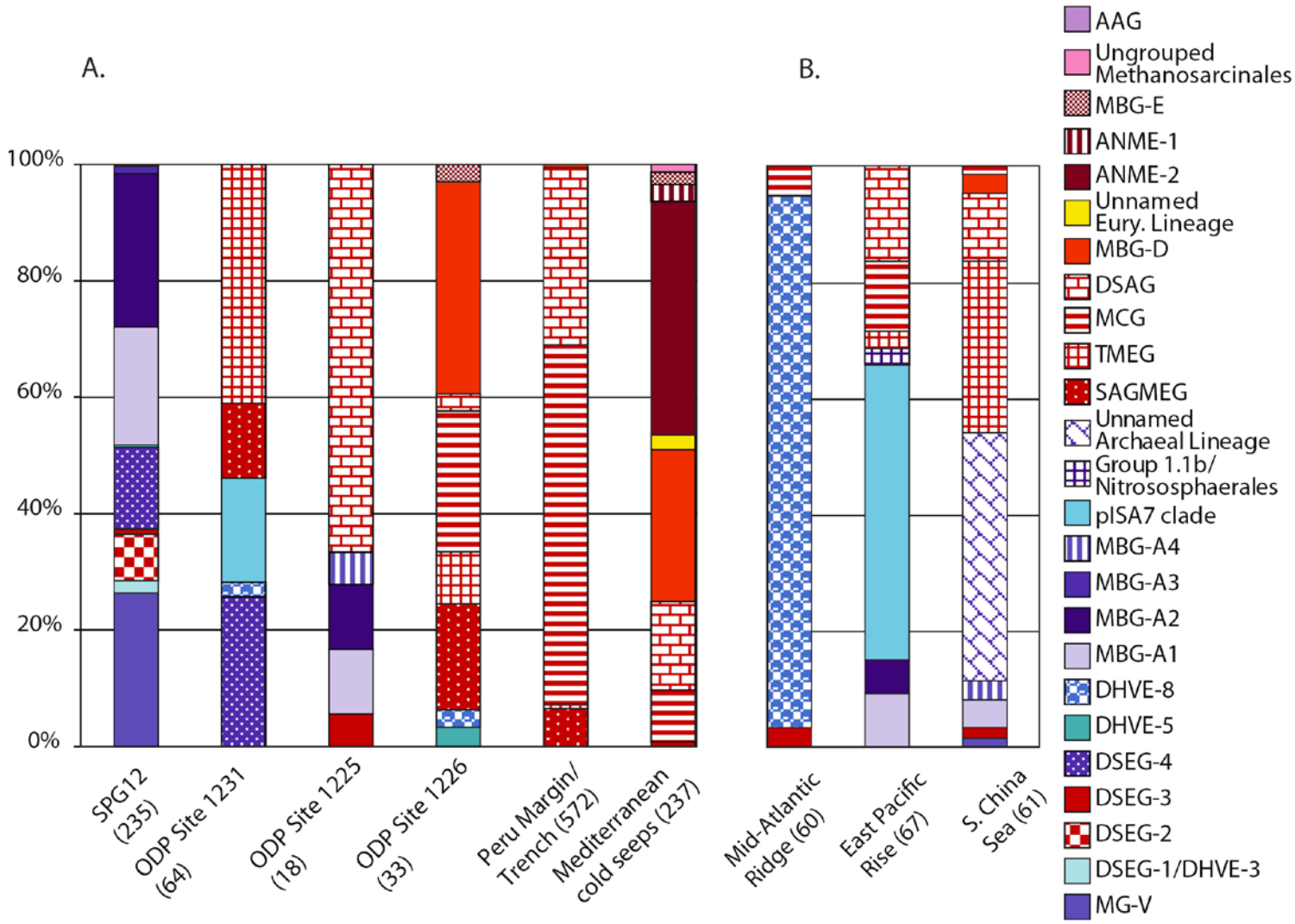

A.
B.
100%
80%
60%
40%
20%
0%
SPG12 (235)
ODP Site 1231 (64)
ODP Site 1225 (18)
ODP Site 1226 (33)
Peru Margin/ Trench (572)
Mediterranean cold seeps (237)
Mid-Atlantic Ridge (60)
East Pacific Rise (67)
S. China Sea (61)
AAG
Ungrouped Methanosarcinales
MBG-E
ANME-1
ANME-2
Unnamed Eury. Lineage
MBG-D
DSAG
MCG
TMEG
SAGMEG
Unnamed Archaeal Lineage
Group 1.1b/ Nitrososphaerales
pISA7 clade
MBG-A4
MBG-A3
MBG-A2
MBG-A1
DHVE-8
DHVE-5
DSEG-4
DSEG-3
DSEG-2
DSEG-1/DHVE-3
MG-V

Figure 13. Percent-abundance of archaeal phyla, excluding MG-I, in clone libraries from all depths at oligotrophic, mesotrophic, and eutrophic sites described in Table 6 (A), plus additional putatively oligotrophic sites with little or no available geochemical information (B). The sites included are oligotrophic sites SPG12 (Durbin, 2009), ODP Site 1231 (Sørensen et al. 2004), 1225 (Teske and Sørensen 2008); mesotrophic site 1226 (Klingensmith et al., unpublished); eutrophic sites ODP Leg 201 Peru Margin sites (Biddle et al. 2006, Sørensen and Teske 2006), Mediterranean cold seeps (Heijs et al. 2008), and putatively oligotrophic sites Rainbow Vent Field (Mid-Atlantic Ridge; Nercessian et al. 2005), East Pacific Rise at 13° N (Li et al. 2008), and Xishu Trough, South China Sea (Tao et al. 2008). All clades found at the organic-rich endmember Peru Margin and Mediterranean mud volcano sites were presented in red or orange colors, with clades not found at these sites colored blue. (A) represents all sites with available geochemical data indicative of sedimentary trophic state, while (B) consists of sites without such data, but which underlie pelagic systems with low surface productivity; additionally, at least one site, the Rainbow Vent Field, Mid Atlantic Ridge, shows hydrothermal influences in the form of vent fluid precipitates.

"big four" collection of uncultured archaeal lineages predominate in organic-rich sediment clone libraries (e.g., Inagaki et al. 2006, Biddle et al. 2006, Fry et al. 2008) and TAG sequences (Biddle et al. 2008): MCG, MBG-B, SAGMEG, and MBG-D. MCG is the dominant group at the Peru Margin sites, accounting for 353/572 (62%) clones, while DSAG/MBG-B (Inagaki et al. 2003, Vetriani et al. 1999), which occurs in high abundance in clone libraries from the SMTZs (Sørensen and Teske 2006), formed a second dominant archaeal clade, at 174/572 (30%) of sequenced clones. Also present were a small number of SAGMEG clones (37/572, 6%) (Takai et al. 2001a), as well as MBG-D (3/572) (Vetriani et al. 1999), and TMEG (4/572) (Takai et al. 2001a). Dominant archaeal phyla recovered from the Mediterranean mud volcano sites included ANME-2 (95/237, 40%) (Orphan et al. 2002), MBG-D (62/237, 26%) (Vetriani et al. 1999), MBG-B/DSAG (36/237, 15%) (Vetriani et al. 1999, Inagaki et al. 2003), and MCG (19/237, 8%) (Takai et al. 2001a, Inagaki et al. 2003) (Figure 11). By contrast, no ANME 16S rRNA gene sequences were recovered from the Peru Margin SMTZ sites in Biddle et al. (2006) or Sørensen and Teske (2006), although a small proportion of ANME sequences were found in a later TAG sequencing effort (Biddle et al. 2008).

The most oligotrophic site featured in Table 7 is SPG11, where oxygen is present throughout the measured sediment column and DIC (alkalinity) does not vary from seawater values. However, all archaeal sequences recovered for this site were MG-I, a group that is not considered in the current study due to the lack of appropriate comparisons with the requisite water column contamination controls. Thus, for the current study, SPG12 represents the oligotrophic endmember site. No MCG, DSAG, MBG-D, or ANME clones, and only a single SAGMEG clone, were recovered from SPG12; conversely, the clones abundant in the SPG12 library, including MG-V, the DSEG, and DSEG-4 phyla, and the MBG-A-related lineages, were not recovered from either of the organic rich environments (Figure 13). The DHVEG-II lineages (DSEG,

DSEG-2, DSEG-3, DSEG-4, DHVE-5 and DHVE-8) are almost entirely absent from the eutrophic endmember sites, with two exceptions: DSEG-3 is a sparsely populated lineage uncommon in oligotrophic or eutrophic datasets but represented in eutrophic endmember clone libraries by a few clones; and DSEG-2, which is somewhat abundant in the SPG12 dataset and is also represented by a few clones from eutrophic endmember clone libraries, although a deep phylogenetic division separates it into two lineages, with only one containing clones from eutrophic endmember sites (Figure 11). Geochemically, ODP sites 1225 and 1231 are more similar to SPG12 than to the eutrophic endmember sites, but are nonetheless intermediate between the two trophic endmembers. Likewise, they display an intermediate clone library composition, including a significant proportion of groups found at eutrophic sites as well as groups only found at oligotrophic sites. 1226, a mesotrophic site, displayed a small proportion of lineages not found in the organic-rich endmember sites, although a smaller proportion than the oligotrophic sites, which is in line with the trend of decreasing 'blue' groups with decreasing oligotrophicity.

Sampling depths varied between sites: SPG12 was not sampled deeper than 4 m; neither 1225 nor 1231 were sampled with high resolution at the surface; and the sediment-water interface and most surficial sediments from the eutrophic sites were not sampled at all, hindering cross comparisons. Presence or absence of phyla may simply reflect the lack of samples of equivalent geochemical strata at the different sites. However, as defined in this paper, "oligotrophic sediments" are equated with oxic or suboxic conditions, while "eutrophic sediments" are equated with anoxic conditions, since those are the redox conditions of the sediment samples from each trophic state, and oxic/suboxic conditions or anoxic conditions characterize the bulk of the oligotrophic or eutrophic sediment column. Additionally, we consider here as "organic-rich" associated phyla those groups that are found in any number at the organic-rich sites. This simplifies the dataset to phyla that are found in the well-sampled anoxic sediments, and phyla that are not. Assuming that methodological and sampling biases mainly work to reduce the relative extent or diversity of a 'difficult' phylum but not eliminate it from clone libraries, this conservative approach avoids some of the pitfalls of comparing studies that used different experimental methods, and avoids making judgment calls over whether, for example, the low relative abundance of a given phylum in a eutrophic-site clone library is due to sampling bias or accurately approximates in-situ abundance.

The initial expectation of this investigation was that the same archaeal phyla recovered in clone

libraries from the well-characterized organic-rich sites, such as the MBG-B/DSAG, MCG, and MBG-D, would also compose clone libraries from oligotrophic sites. Clearly, this hypothesis is not supported by the data presented in Figure 13. Eutrophic-endmember archaeal phyla constituted a decreasing proportion of membership and abundance along an environmental gradient of indicators, such as maximum DIC (alkalinity), sedimentation rate, and organic carbon content, that indicate increasing nutrient limitation and increasing extent of oxic and suboxic strata (Figure 13, Table 5). Thus, some dominant benthic archaeal groups, such as the DSAG and MCG, which so far have appeared to characterize the marine subsurface based on sampling in organic-rich sediments, may in fact not be universal.

5. Association of archaeal lineages with high-energy electron acceptors.

The proportion of archaeal 16S lineages that do not appear in the two eutrophic endmember clone libraries increases almost linearly with increasing oligotrophicity (as revealed by maximum DIC values and other measures; Table 6). There are three likely drivers of a change in community composition with change in trophic state:

(1) Oligotrophic sediments are primarily characterized by slow sedimentation rates, leading to a low substrate flux for microbes and hence energy stress. Microbes in oligotrophic sediments might be adapted to a slow input of organic matter, which would lower the maximum Gibbs free energy of reaction for a given redox process.

(2) Since slower substrate input flux leads to slower depletion of electron acceptors, higher energy electron acceptors extend over a larger portion of the sediment column. Thus, a second driver of the observed pattern may be that subsurface Archaea specialize in the electron acceptor that maximizes the energy yield for a given mole of substrate, leading to different lineages predominating in oligotrophic sediments.

(3) Finally, substrate type might be expected to differ for sediments underlying a 3-5km water column, as organic matter changes significantly in composition during sinking (Wakeham et al. 2004, Hedges et al. 2001), and is expected to become more refractory overall (Lee et al. 1997). Microbes that can metabolize the large fraction of highly refractory substrates in oligotrophic sediments where competition for electron donors is strong would have an advantage over microbes

not able to use these substrates.

These variables are collinear to some degree, and distinguishing the primary driver is not possible in the current analysis. It also seems likely that some combination of these three factors determines community composition: while electron acceptor specialization may be a determinative factor for one group, another group may thrive due to the abundance of particular difficult substrate.

A key argument in favor of a higher-energy electron acceptor association for the 'blue' lineages (Figure 13) is that these lineages are completely absent from the eutrophic endmember clone libraries; high-energy electron acceptors are also absent from the depths sampled from eutrophic sites. By contrast, difficult substrates that are hypothesized to be more abundant in oligotrophic sediments would still be expected to be present in eutrophic sediments, particularly in the deep subsurface where sediments may be millions of years old. It could also be argued that the increasing abundance of 'blue' lineages with increasing oligotrophicity is consistent with adaptation to slow substrate fluxes. However, electron-donor limitation affects the community as a whole, and so if this were the driver behind the appearance of oligotrophic lineages, one might expect them to account for a much larger proportion of the non-eutrophic sites. A recent review examining similar questions in soil microbes focused mostly on the substrate side of metabolism (examining carbon remineralization rate, organic C % weight, silt+clay % weight, soil pH, mean annual temperature, and annual soil moisture deficit), and found that only carbon remineralization rate, an index of carbon availability to microbes, accounted for a significant proportion (25-36%) of the variance in abundance of three dominant groups (the oligotrophic Acidobacteria and the copiotrophic Betaproteobacteria and Bacteroidetes; Fierer et al. 2007). However, the redox environment of soil can be highly heterogenous on a small scale (Conrad 1996), whereas the redox environment of marine sediments is highly zonated, and so redox specialization, instead of substrate flux specialization, may be more paramount in marine sediments. The linear decrease in 'blue' lineages with decreasing oligotrophicity could thus be interpreted as due to the decreasing suboxic/oxic proportion of the sediment column as carbon flux increases.

Primer bias limits the robustness of these arguments. For example, the relative proportions of oligotrophic groups in less oligotrophic sediments may be distorted in favor of eutrophic groups if these groups are less susceptible to primer bias, as has been suggested (Durbin 2009). Similarly, more primer-

friendly lineages may swamp 'oligotrophic' lineages in eutrophic sediments, preventing their detection, and a similar argument could be made for why a phase-shift to oligotrophic groups isn't observed in non-eutrophic sediments. However, it is expected that primer bias might lower the overall relative abundance and diversity of deeply-branching oligotrophic associated lineages in clone libraries, and not eliminate them altogether; as long as the same general trend is preserved, this does not materially affect the interpretation of the data in the current study.

The lineages depicted in blue in Figure 13 are phylum or class-level lineages (Figure 11, Figure 12) that are found only in oxic to suboxic oligotrophic sediments and not in eutrophic endmembers, despite extensive sampling. Interestingly, these lineages are often associated with other oxic or suboxic environments, and have few or no close relatives found in purely anoxic environments. This is illustrated in Figure 14, which shows the number of independent studies that have recovered a given lineage from a particular redox environment, for all relevant sequences in the SILVA SSU Ref v.95 database. Studies of oxic or suboxic environments, including oxic/suboxic sediments or water column, and the typically suboxic/oxic environments of hydrothermal fluids and chimney surfaces, account for a majority or plurality of groups associated with 'blue' lineages (Figure 13), such as the pISA7 Crenarchaeotal and the DHVEG-II Euryarchaeotal lineages. Although at their source hydrothermal fluids are often highly reduced and anoxic, they are principally oxic or suboxic when sampled in the thermally habitable region of a mixing gradient between an anoxic, hyperthermal source fluid and an oxic endmember (seawater, groundwater or atmosphere) (e.g., Amend and Shock 1998, Spear et al. 2005, Dias and Barriga 2006, Rogers and Amend 2006; but see Kelley et al. 2001). The second largest group of studies for these 'blue' lineages is most often studies of environments that have heterogeneous or uncertain redox states, such as soil (e.g., Conrad 1996), ground water (e.g., Jakobsen 2007), and hydrothermal sediments (including hydrothermal deposits and sediments bathed in hydrothermal fluid; Teske et al. 2002, Severmann et al. 2006, Nercessian et al. 2005, Dias and Barriga 2006), where 'blue' lineages are hypothesized here to mostly occupy high-energy acceptor niches.

Another large portion of studies for these lineages are studies of shallow, organic rich sediments. While typically these environments are thought of as anoxic, they can be distinguished from the purely anoxic environments such as eutrophic deep subsurface sediments or anoxic bioreactors because they often

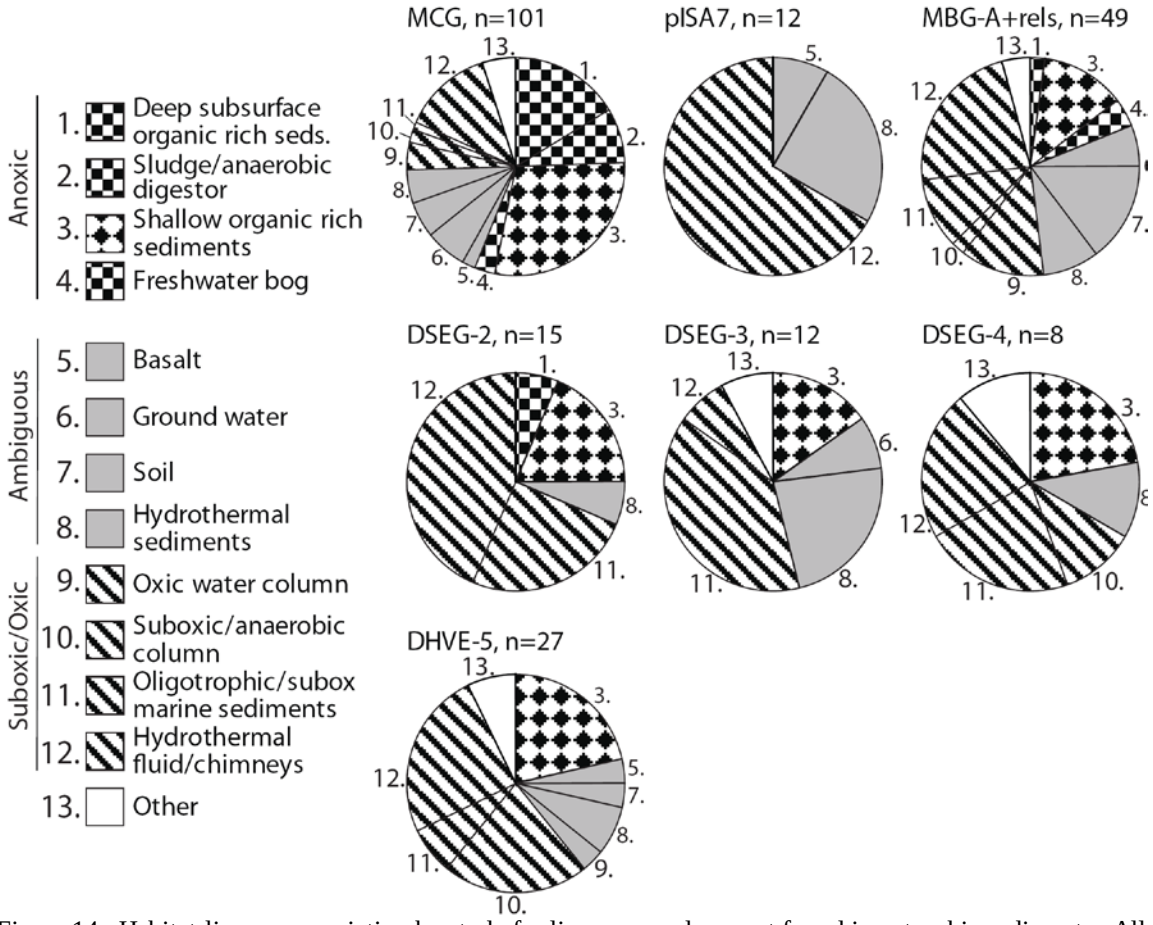


Figure 14. Habitat-lineage association by study for lineages rarely or not found in eutrophic sediments. All sequences belonging to the relevant clades in SILVA SSU Ref V.95 were examined, with each instance of a clade being recovered by a unique study counting as a sample point, and sample points were categorized by habitat association. Habitats were then categorized as being either oxic/suboxic, anoxic, or ambiguous if a redox status is uncertain or heterogeneous within a given habitat. Note that hydrothermal fluids are generally sampled not at the 'source,' but somewhere along a mixing continuum with an aerobic endmember and so are assumed to be primarily suboxic or oxic. Shallow organic sediments are primarily anoxic, although very thin surface layers or microenvironments may be oxic/suboxic. Sample points with example sequence accession numbers are tabulated in Appendix F.

have suboxic or oxic layers (e.g., Jørgensen 1977, Mäkelä and Tuominen 2003, Jørgensen et al. 2005, Glud 2008, Ku et al. 2008), particularly for near-shore marine sediments where mineral input may introduce significant metal oxides (Thomas and Bendell-Young 1999, Poulton and Raiswell 2002). In contrast, the typically eutrophic-associated MCG is recovered in a large proportion of studies of purely anoxic sites, such as the marine deep subsurface, anaerobic digestors, and freshwater bogs, and a lower proportion of suboxic or oxic sites. They are hypothesized here to occupy less-energetic metabolic niches in habitats of heterogeneous or ambiguous redox state.

These patterns may integrate over some inconsistencies, such as potential minority clades within lineages that engage in metabolisms that are non-typical for that lineage, and also reflect uncertainty in identifying the exact redox state of the environment from which the clone was recovered, such as the case of shallow organic-rich sediments; this is particularly true for unpublished datasets with only brief sequence-source descriptions. Nevertheless, these results are again consistent with a largely suboxic/oxic metabolism for the oligotrophic-associated lineages, and suggest an avenue for future research may be verifying whether oligotrophic lineages in apparently anoxic environments are actually inhabiting suboxic/oxic microniches.

6. Problems for future research.

The biological meaning of different deeply-branching archaeal lineages predominating in oligotrophic and eutrophic sediments is unclear. Of course, the correlation between phylogenetic composition of clone libraries and sediment trophic state may change as more of the oligotrophic sediment column is sampled, particularly deep suboxic sediments, for example. However, if the correlation between deep archaeal phylogeny and electron-accepting environment is supported in further study, why is this the case? Why do deep phylogenetic divisions exist between archaeal clone libraries from subsurface sedimentary environments of different trophic state? Finding that groups originally identified in oligotrophic sediments do in fact exist in eutrophic sediments, but only in suboxic or oxic microniches, would be consistent with the idea that metabolic specialization is the defining characteristic of these deep-branching 16S lineages. A glance at the phylogenetic trees (Figure 11, Figure 12) and the habitat-lineage associations (Figure 14) suggests this may be possible. For example, some representatives of the DHVEG-II groups have been recovered from shallow methanogenic sediments of Skan Bay, Alaska (the 'SBAK' clones; Kendall et al. 2007; Figure 11), sediments that are shallow enough and close enough to land that faunal bioturbation and terrigenous input of minerals may lead to a heterogeneous sediment environment containing oxic or suboxic microniches, as well as suboxic/oxic surficial microlayers.

The complement to this test is to examine at what hierarchical level phylogenetic lineages assort between different 'versions' of the same electron accepting environment, comparing, e.g., the spatially expanded suboxic/oxic strata of abyssal sediments and the suboxic/oxic strata perched on top of deep

anoxic sediments. This would indicate whether factors other than electron acceptor, such as rate of substrate flux or substrate lability, can be linked with phylum-level divergences, or only with lower-order divergences. For example, while electron-acceptor specialization may play a key role in the adaptive landscape of Archaea, other factors may play a role in niche differentiation in marine subsurface sediments: low substrate flux increase biosynthentic costs as a proportion of cellular metabolism in higher-free-energy electron-accepting environments, an effect exacerbated in oligotrophic sediments by the severe lack of reducing equivalents; however, the importance of this difference relative to total cell energy expenditure is unclear (McCollom and Shock 2005). Such factors may differentiate deep oligotrophic sediment environments from even the oxidatively equivalent zones of margin sediments, albeit hypothetically at a lower taxonomic rank. Other factors such as dispersal contingencies or isolation by distance or physical barrier will be particularly important to consider when looking at finer-scale biogeography and phylogenetic patterns (Hughes-Martiny et al. 2006). Ultimately, however, even if biogeographical approaches can reveal electron acceptor type to be the primary correlate of phylum level divergences, this may indicate either exploitation of a given electron acceptor, or adaptation to the conditions associated with its use. Comparative genomics coupled to 16S identification may shed light on those questions, as may more extensive, fine-scale geochemical comparisons of suboxic/oxic vs suboxic/oxic environments.

Several problems exist with standard approaches to assaying microbial diversity of oligotrophic sediments. Because cell densities are low to extremely low, standard in-matrix cell lysis approaches may recover inadequate DNA amounts due to adsorption to clay or other mineral particles, which can be highly efficient in binding up negatively charged DNA molecules (e.g., Webster et al. 2003). Other threats to DNA yield include nuclease attack and shearing (Roose-Amsaleg et al. 2001). Whole-sediment extraction methods may also co-extract PCR inhibitors (e.g., Barton et al. 2006), which become problematic when attempts are made to concentrate DNA extracts to minimize jackpot amplification bias. A further issue when using PCR-based approaches is primer bias, as reviewed by Teske and Sørensen (2008). DHVE-6, a lineage within the DHVEG-II superphylum, had severe mismatch problems with commonly used primers, with nearly all primers having one or more mismatches with all DHVE-6 sequences examined (Teske and Sørensen 2008). Further, PCR using primers designed to target ANME euryarchaeota recovered a novel deeply-branching Thermoplasmatales clade basal to the MBG-D/TMEG/MG-II/MG-III cluster, from

oligotrophic North Pacific Gyre sediments (Wang et al. 2005). This suggests that new general primers, perhaps Euryarchaeota-specific, may uncover significantly greater diversity in energy-limited sediments than standard primers. Existing primer sets likely uncover only a small fraction of diversity even within known phyla, a hypothesis that finds support in the extremely low species diversity within the novel archaeal groups uncovered at SPG12.

7. Conclusions.

The results examined in this review suggest that Archaeal communities undergo a marked shift along gradients that determine sediment trophic state. Archaeal lineages found in oligotrophic, primarily oxic to suboxic sediments significantly expand the higher-order taxonomic diversity within the Euryarchaeota and Archaea. These lineages have most often been found in other environments that are suboxic or oxic, with some proportion occupying habitats of ambiguous redox state, or a habitat that is primarily anoxic with the potential for oxic/suboxic microniches (organic-rich shallow sediments). Such a lineage distribution is consistent with electron acceptor specialization determining in large part the distribution of archaeal lineages in the marine subsurface. Since the phylogenetic groups associated with this diversity shift were deeply-branching (possibly class or phylum level), if this premise holds, it is also consistent with electron acceptor specialization being a fundamental correlate with archaeal diversification. However, phylogenetic assortment by habitat may also reflect most-energetic terminal electron acceptor only indirectly, via specialization in the geochemical milieu characteristic of a given electron accepting zone, perhaps by parasitism or endosymbiosis. Further studies are needed to resolve the phylogenetic placement of these novel lineages, as well as to explore the relationship between two defining characteristics of oligotrophic habitats, nutrient limitation and availability of high-energy electron acceptors, and archaeal 16S rRNA gene evolution.

8. Methods.

This study involved analyzing Archaeal 16S rRNA and 16S rRNA gene sequences from public databases, as well as data generated from Cruise Knox02RR South Pacific Gyre site 12 (SPG12). Sequences from SPG12 were amplified and sequenced as described in Durbin (2009). A search for all

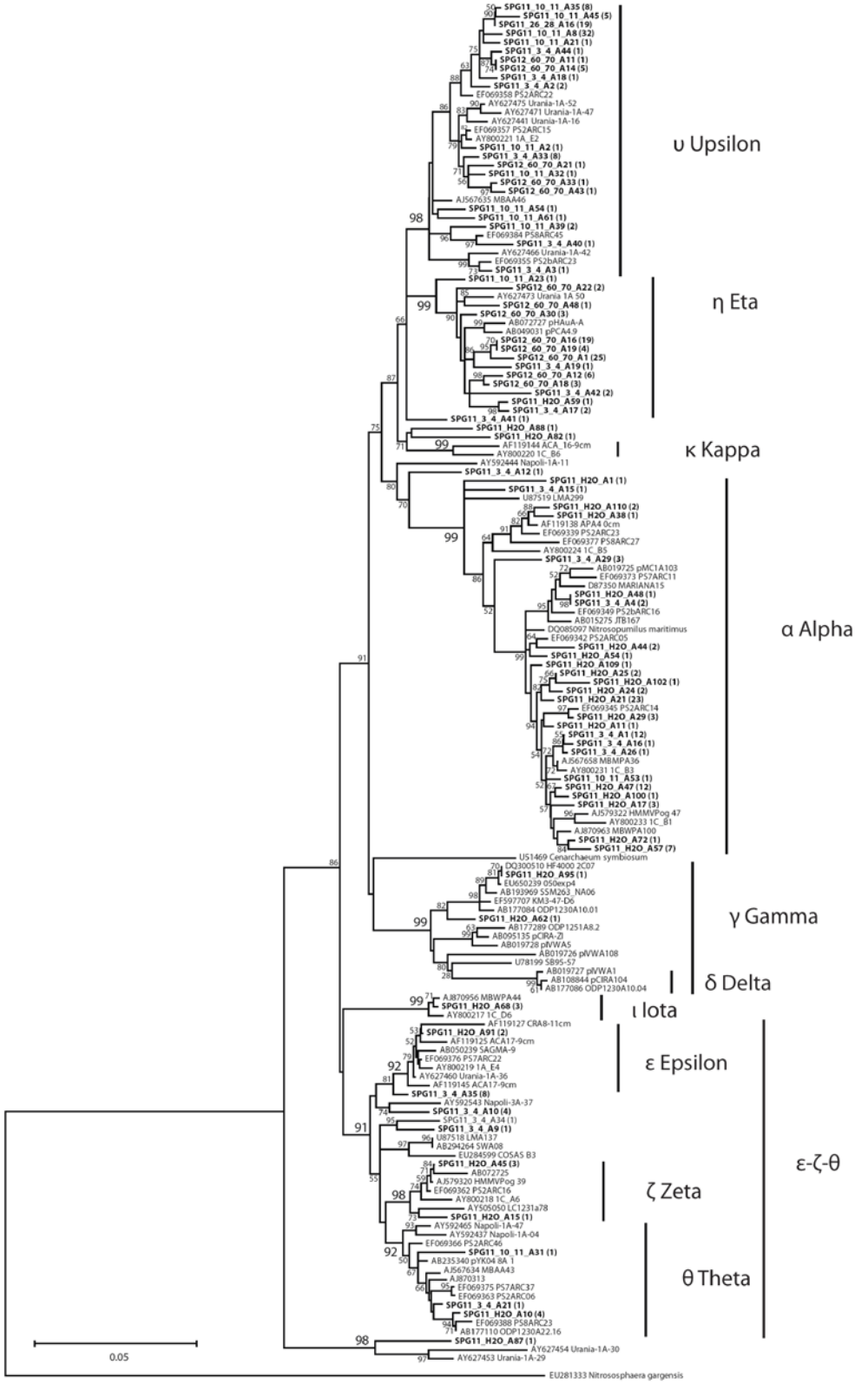
available Archaeal 16S sequences from oligotrophic marine sediment sites was conducted first by identifying closest relatives to SPG12 Archaeal sequences via BLAST searches and searches within the ARB v.95 REF 16S/18S database. If a sequence was derived from an oligotrophic sediment environment, defined as any marine benthic environment not situated on a continental slope or shelf, all sequences from the associated publication were imported and aligned in ARB. Any closest relatives presented in the associated publication were also imported and aligned. Additionally, all sequences from abyssal sites 1231 (Sørensen and Teske 2004) and 1225 (Teske and Sørensen 2008) of ODP Leg 201 were included. Further internet searches using keywords yielded no additional publications. For comparison, sequences representative of major phylogroups from the sulfate-methane transition zones of sites 1227 (Sørensen and Teske 2006, Biddle et al. 2006), 1229 (Biddle et al. 2006), and 1230 (Biddle et al. 2006) were aligned manually in ARB, as well as representative sequences from the upper ~40 cm of a Mediterranean sapropel/cold seep site (Heijs et al. 2008).

Phylogenies were estimated separately for the Crenarchaeota and Euryarchaeota. After manual alignment adjustments, sequences from the assembled oligotrophic dataset longer than 1200 nt, or closest relatives longer than 1200 nt, were exported from ARB for phylogenetic analysis. A base-frequency filter excluding sites with a base shared by less than 60% of sequences was estimated separately for the Euryarchaeota and Crenarchaeotal datasets to create a more conserved alignment for analysis. The final analyzed alignment with all gaps or missing data excluded consisted of 972 bp for the Euryarchaeota, and 968 bp for the Crenarchaeota and deeply-branching archaeal lineages. A maximum likelihood phylogeny was then estimated using Treefinder (Jobb et al. 2004). Support for clades was estimated using 1000 maximum-likelihood bootstrap replications. For the Euryarchaeota, the same alignment was used to construct a phylogeny using the neighbor-joining algorithm, with a maximum-composite-likelihood-estimated substitution model and complete deletion of gaps or missing data, in MEGA 4.0 (Tamura et al. 2007). Branch support for this phylogeny was estimated using the interior-branch support algorithm (Sitnikova et al. 1995), as bootstrap is vulnerable to a conservative bias when large numbers of sequences are considered (Sitnikova et al. 1995, Sanderson and Wojciechowski 2000). The neighbor-joining tree is presented inscribed with both interior bootstrap support from the neighbor-joining algorithm, as well as the maximum-likelihood bootstrap support values for the monophyly of the lineages contained with the clade.

Additionally, bootstrap support values from the concatenated-gene dataset of Brochier-Armanet et al. (2008) are included when they support the monophyly of the clades shown in the neighbor-joining tree. Phylum affiliations of sequences shorter than 1200 bp were then determined by constructing a NJ tree in ARB using a smaller alignment, and/or adding short sequences to the ~1200 nt alignment reference tree via the parsimony-add tool in ARB. Relative abundances of sequences for the different major clades was then determined by either counting sequences in ARB, or by using the abundances of representative sequences as described in the original publication.

To calculate distances between groups, and to align sequences for ML phylogenetic analysis, all available sequences greater than 900 bp and found in the SILVA v.95 database (Pruesse et al. 2007), current as of June 2008, were included for analysis, with the exception of the DHVE-5 (Takai and Horikoshi 1999), a very large and diverse group. Chimeras were excluded using Bellerophon3 (Huber et al. 2004). The Silva v. 95 SSU REF phylogeny was used to identify those lineages that branch nearest to clades containing sequences from Takai and Horikoshi 1999 and sequences previously found to be closest relatives of SPG12 deep-branching sequences. These lineages were then re-aligned as necessary, using conserved segments as anchors and by optimizing the alignment to reflect likely secondary structure base-pairings. It was found that many lineages had lengthy insertions. These may be additional loops not accounted for in the secondary structure model supplied in ARB, based on the appearance of typical stem-loop sequences (Woese et al. 1990) and conserved sequence motifs between insertions found in closely related species. Often these insertions caused misalignment of more conserved regions in the alignment supplied in the SILVA database. The complete set of DHVEG-II phylogroups found in Takai and Horikoshi (1999), Nercessian et al. (2005), and SPG12 (Durbin, 2009), as well as their closest neighboring clades were checked and re-aligned in their entirety with the exception of the DHVE-5. Only representative sequences from DHVE-5 clades that contained or were most closely related to SPG12 and Takai and Horikoshi (1999) sequences were examined in this way, due to the large number of sequences from the DHVE-5-related groups in the SILVA database. This re-aligned sequence set was then used to construct a reference tree, and from that representative or all available sequences >1200 bp were selected for further analysis using MEGA4 or Treeclimber.

Appendix A. 1% OTU NJ phylogeny for MG-I.



Appendix A. NJ phylogeny, using a maximum-composite-likelihood model of sequence evolution with gamma-corrected site rates and calculated for an 800-base-pair alignment, for SPG11 and SPG12 MG-I 1% representative phylotypes (operational taxonomic units; OTUs).

Appendix B. 1% OTU representative phylotype table for MG-I.

1% OTU phylotype name	1% OTU phylotype accession number	Best BLAST max %ID	Best BLAST match accession number	Best BLAST clone	MG-I subgroup designation	SPG11 overlying water	SPG11 3-4 cmbsf	SPG11 10-11 cmbsf	SPG11 26-28 cmbsf	SPG12 60-70 cmbsf	SPG12 153-163 cmbsf	3% OTU representative phylotype accession number
SPG12_60_70_A12	FJ487512	97	AY627473	Urania-1A-50	eta					6		FJ487503
SPG11_3_4_A17	FJ487528	95	EF069355	PS2bARC23	eta		1	1				FJ487503
SPG12_60_70_A18	FJ487515	97	AY627473	Urania-1A-50	eta					3		FJ487503
SPG12_60_70_A16	FJ487514	98	AY627473	Urania-1A-50	eta			2		13	4	FJ487503
SPG12_60_70_A1	FJ487509	97	AY627473	Urania-1A-50	eta			5	8	12		FJ487503
SPG12_60_70_A30	FJ487520	97	AY627473	Urania-1A-50	eta			1		2		FJ487503
SPG11_H2O_A59	FJ487503	98	EF069357	PS2ARC15	eta	1						FJ487503
SPG12_60_70_A48	FJ487522	97	AY627473	Urania-1A-50	eta					1		FJ487503
SPG11_3_4_A19	FJ487530	97	AY627473	Urania-1A-50	eta		1					FJ487503
SPG12_60_70_A19	FJ487516	97	EF069359	PS2ARC17	eta			2		2		FJ487503
SPG11_10_11_A23	FJ487550	97	EF069359	PS2ARC17	eta			1				FJ487550
SPG11_3_4_A29	FJ487535	96	EF069352	PS2ARC11	alpha		1					FJ487529
SPG11_10_11_A8	FJ487551	98	EF069358	PS2ARC22	upsilon			2	30			FJ487551
SPG11_10_11_A45	FJ487563	98	EF069358	PS2ARC22	upsilon			5				FJ487551
SPG11_10_11_A32	FJ487553	98	EF069357	PS2ARC15	upsilon			1				FJ487517
SPG12_60_70_A21	FJ487517	98	EF069357	PS2ARC15	upsilon					1		FJ487517
SPG12_60_70_A33	FJ487521	98	EF069357	PS2ARC15	upsilon					1		FJ487517
SPG11_3_4_A2	FJ487531	99	EF069357	PS2ARC15	upsilon		1	1				FJ487517
SPG11_3_4_A33	FJ487537	99	EF069357	PS2ARC15	upsilon		1	2	5			FJ487517
SPG11_10_11_A35	FJ487562	98	EF069360	PS2ARC24	upsilon			8				FJ487551
SPG11_10_11_A21	FJ487549	98	EF069360	PS2ARC24	upsilon			1	0			FJ487551
SPG12_60_70_A14	FJ487513	99	EF069360	PS2ARC24	upsilon		1			4		FJ487551
SPG11_3_4_A18	FJ487529	97	EF069360	PS2ARC24	upsilon		1					FJ487529
SPG11_10_11_A61	FJ487561	98	EF069357	PS2ARC15	upsilon			1				FJ487517
SPG11_10_11_A54	FJ487560	97	EF069357	PS2ARC15	upsilon			1				FJ487550
SPG11_H2O_A91	FJ487481	98	EF069376	PS2ARC22	epsilon	1	1					FJ487481
SPG11_10_11_A39	FJ487554	96	EF069359	PS2ARC17	upsilon			2				FJ487529
SPG11_3_4_A44	FJ487545	97	EF069385	PS8ARC12	upsilon		1					FJ487551

SPG12_60_70_A22	FJ487518	97	EF069360	PS2ARC24	eta				2	FJ487518
SPG12_60_70_A11	FJ487511	98	EF069360	PS2ARC24	upsilon				1	FJ487551
SPG11_10_11_A2	FJ487548	98	EF069357	PS2ARC15	upsilon			1		FJ487536
SPG12_60_70_A43	FJ752632	97	EF069357	PS2ARC15	upsilon				1	FJ487536
SPG11_3_4_A3	FJ487536	97	EF069355	PS2bARC23	upsilon		1			FJ487536
SPG11_3_4_A40	FJ487541	97	EF069384	PS8ARC45	upsilon		1			FJ487541
SPG11_3_4_A41	FJ487542	97	EF069355	PS2bARC23	ungrouped		1			FJ487529
SPG11_H2O_A82	FJ752631	98	AY505049	LC1231a76	ungrouped	1				FJ752631
SPG11_H2O_A106	FJ487482	98	EF069352	PS2ARC11	alpha	2	1			FJ487473
SPG11_H2O_A47	FJ487475	98	EF069352	PS2ARC11	alpha	11		1		FJ487473
SPG11_H2O_A17	FJ487484	98	EF069353	PS2bARC07	alpha	3				FJ487473
SPG11_H2O_A57	FJ487486	98	EF069353	PS2bARC07	alpha	7				FJ487473
SPG11_H2O_A11	FJ487489	98	EF069353	PS2bARC07	alpha	1				FJ487473
SPG11_H2O_A100	FJ487500	98	EF069352	PS2ARC11	alpha	1				FJ487473
SPG11_3_4_A1	FJ487523	99	EF069352	PS2ARC11	alpha		6	5		FJ487473
SPG11_H2O_A102	FJ487505	98	AY856365	CTD005-11A	alpha	1				FJ487473
SPG11_H2O_A21	FJ487473	99	FJ150794	116exp4	alpha	23				FJ487473
SPG11_H2O_A25	FJ487491	99	AY856365	CTD005-11A	alpha	2				FJ487473
SPG11_H2O_A24	FJ487490	98	EF069352	PS2ARC11	alpha	2				FJ487473
SPG11_H2O_A29	FJ487474	99	EF069345	PS2ARC14	alpha	3				FJ487473
SPG11_H2O_A54	FJ487502	99	EF069371	PS7bARC04	alpha	1				FJ487473
SPG11_H2O_A44	FJ487477	98	EF069381	PS8ARC18	alpha	2				FJ487473
SPG11_3_4_A4	FJ487540	99	EF069382	PS8ARC06	alpha		2			FJ487540
SPG11_H2O_A48	FJ487497	98	EF069382	PS8ARC06	alpha	1				FJ487540
SPG11_H2O_A72	FJ487478	97	EF069352	PS2ARC11	alpha	1				FJ487473
SPG11_H2O_A42	FJ487543	97	EF069352	PS2ARC11	alpha	1				FJ487473
SPG11_10_11_A53	FJ487557	97	EF069352	PS2ARC11	alpha			1		FJ487483
SPG11_3_4_A26	FJ487534	97	EF069352	PS2ARC11	alpha		1			FJ487483
SPG11_3_4_A16	FJ487527	97	EF069352	PS2ARC11	alpha		1			FJ487483
SPG11_H2O_A62	FJ487493	98	AB193969	SSM263-NA06	gamma	1				FJ487493
SPG11_H2O_A95	FJ487508	99	EU686642	SAT1000-49-D2	gamma	1				FJ487493
SPG11_H2O_A109	FJ487483	97	AY856365	CTD005-11A	alpha	1				FJ487483

SPG11_3_4_A42	FJ487544	97	EF069352	PS2ARC11	eta		1			FJ487543
SPG11_H2O_A1	FJ487499	97	AY856365	CTD005-11A	alpha	1				FJ487499
SPG11_3_4_A15	FJ487526	98	EF069352	PS2ARC11	alpha		1			FJ487499
SPG11_H2O_A38	FJ487492	98	EF069378	PS8ARC36	alpha	1				FJ487472
SPG11_H2O_A110	FJ487472	99	EF069378	PS8ARC36	alpha	2				FJ487472
SPG11_3_4_A12	FJ487525	97	EF069367	PS7ARC40	ungrouped		1			FJ487525
SPG11_3_4_A21	FJ487532	97	EF069374	PS7ARC32	theta		1			FJ487476
SPG11_3_4_A10	FJ487524	96	EF069362	PS2ARC16	ungrouped		1			FJ487524
SPG11_H2O_A10	FJ487476	99	EF069365	PS2bARC09	theta	2	1			FJ487476
SPG11_H2O_A45	FJ487485	99	EF069362	PS2ARC16	zeta	3				FJ487485
SPG11_H2O_A15	FJ487494	97	EF069362	PS2ARC16	zeta	1				FJ487485
SPG11_10_11_A31	FJ487552	96	EF069375	PS7ARC37	theta			1		FJ487538
SPG11_3_4_A34	FJ487538	97	EF069365	PS2bARC09	ungrouped		1			FJ487538
SPG11_3_4_A9	FJ487547	97	EF069374	PS7ARC32	ungrouped		1			FJ487547
SPG11_3_4_A35	FJ487539	97	EF069376	PS7ARC22	ungrouped		1			FJ487524
SPG11_H2O_A68	FJ487496	99	EF069387	PS8ARC01	iota	3				FJ487496
SPG11_H2O_A88	FJ487507	95	FJ002832	T1_35_A2-35	ungrouped	1				FJ487507
SPG11_H2O_A87	FJ487487	92	EF069354	PS2bARC10	ungrouped	1				FJ487487
SPG11_26_28_A16	FJ752633	98	EF069358	PS2ARC22	upsilon			7	14	FJ487551

Appendix B. Table of 1% OTU representative phylotypes, closest NCBI relatives, and depthwise distribution of instances of representative phylotypes. Also included are the corresponding 3% OTU representative phylotypes.

Appendix C. 1% representative phylotype depthwise abundances for SPG12 Archaea.

1% OTU phylotype name	Accession number	BLAST Max %ID	best BLAST hit accession #	best BLAST hit clone name	Phylum designation	60-70 cmbsf	153-163 cmbsf	213-223 cmbsf	273-283 cmbsf	343-353 cmbsf	401-411 cmbsf	Total # clones
SPG12_401_411_A49	FJ487467	94	EF069336	PS2ARC29	MG-V						1	1
SPG12_340_350_A103	FJ487466	93	EF069336	PS2ARC29	MG-V					1		1
SPG12_153_163_A7	FJ487471	94	EF069336	PS2ARC29	MG-V		1					1
SPG12_210_220_A78	FJ487464	96	AY360486	BSA7-305m	MG-V			2				2
SPG12_401_411_A23	FJ487459	95	AB019743	pMC2A5	DSEG-1						5	5
SPG12_153_163_A15	FJ487468	87	EU635927	SSE_L4LC02	DSEG-1		2					2
SPG12_153_163_A3	FJ487470	96	DQ640161	SBAK-mid-32	DSEG-3		3					3
SPG12_210_220_A13	FJ487469	85	AB193961	Ma-NA02	DSEG-4			18				17
SPG12_210_220_A25	FJ487462	88	DQ522913	SBAK-mid-03	DSEG-4			19				18
SPG12_273_283_A4	FJ487465	81	AB019751	pISA1	DHVE-5				1			1
SPG12_153_163_A13	FJ487461	96	DQ301983	1H5_D09	SAGMEG		1					1
SPG12_210_220_A20	FJ487452	91	AB329782	YS16As04	MBG-A3			4				4
SPG12_340_350_A93	FJ487456	93	AB329796	YS16As48	MBG-A1					1		1
SPG12_210_220_A59	FJ487454	99	AJ870340	MBWPA126	MBG-A2			2		23		25
SPG12_210_220_A31	FJ487453	97	AJ870340	MBWPA126	MBG-A2			9	37	13		59
SPG12_273_283_A11	FJ487460	93	AB329796	YS16As48	MBG-A1				44	20		64
SPG12_401_411_A1	FJ487457	99	FJ264821	so4A119	DSEG-2			4			21	25
SPG12_401_411_A20	FJ487458	94	EF069336	PS2ARC29	MG-V			13		36	30	79
SPG12_210_220_A55	FJ487463	96	AY360486	BSA7-305m	near Methanosarcinales			2				2
SPG12_210_220_A11	FJ487455	86	DQ522913	SBAK-mid-03	DSEG-4			8				8
SPG12_60_70_A12	FJ487512	97	AY627473	Urania-1A-50	MG-I eta	6						6
SPG12_60_70_A18	FJ487515	97	AY627473	Urania-1A-50	MG-I eta	3						3
SPG12_60_70_A16	FJ487514	98	AY627473	Urania-1A-50	MG-I eta	13						13
SPG12_60_70_A1	FJ487509	97	AY627473	Urania-1A-50	MG-I eta	12	4					16
SPG12_60_70_A30	FJ487520	97	AY627473	Urania-1A-50	MG-I eta	2						2
SPG12_60_70_A48	FJ487522	97	AY627473	Urania-1A-50	MG-I eta	1						1

SPG12_60_70_A19	FJ487516	97	EF069359	PS2ARC17	MG-I eta	2	2
SPG12_60_70_A21	FJ487517	98	EF069357	PS2ARC15	MG-I upsilon	1	1
SPG12_60_70_A33	FJ487521	98	EF069357	PS2ARC15	MG-I upsilon	1	1
SPG12_60_70_A14	FJ487513	99	EF069360	PS2ARC24	MG-I upsilon	4	4
SPG12_60_70_A22	FJ487518	97	EF069360	PS2ARC24	MG-I eta	2	2
SPG12_60_70_A11	FJ487511	98	EF069360	PS2ARC24	MG-I upsilon	1	1
SPG12_60_70_A43	FJ752632	97	EF069357	PS2ARC15	MG-I upsilon	1	1

Appendix C. 1% OTU representative phylotypes for non-MG-I Archaea, with best BLAST hit sequence information and depthwise phylotype abundances at SPG12.

Appendix D. Bacterial 3% OTU phylotype table.

3% OTU phylotype name	NCBI accession number	Depth, cmbsf 60-70	213-223	273-283	343-353	461-471	Phylogenetic designation
SPG12_343_353_B100	FJ746084				1		Alphaproteobacteria-Rhodospirillaceae
SPG12_343_353_B101	FJ746085				3	1	Chloroflexi IV
SPG12_343_353_B103	FJ746086				1		Chloroflexi-ungrouped
SPG12_343_353_B104	FJ746087			4	5	2	Chloroflexi IV
SPG12_343_353_B107	FJ746089				1		OP3
SPG12_343_353_B108	FJ746090				2		SAR406
SPG12_343_353_B109	FJ746091				1		NT-B9
SPG12_343_353_B110	FJ746092			1	4		Chloroflexi IV
SPG12_343_353_B114	FJ746094				1		OP3
SPG12_343_353_B116	FJ746095				1		Planctomycetes I
SPG12_343_353_B1	FJ746097		5		1	2	Chloroflexi IV
SPG12_343_353_B10	FJ746098			2	6	10	Chloroflexi IV
SPG12_343_353_B11	FJ746099				1		Chloroflexi IV
SPG12_343_353_B117	FJ746100	3			1		Actinobacteria-Acidomicrobiae
SPG12_343_353_B119	FJ746101				1	1	Planctomycetes I
SPG12_343_353_B14	FJ746104			1	4		Chloroflexi IV
SPG12_343_353_B16	FJ746105		1	2	6	6	Chloroflexi IV
SPG12_343_353_B17	FJ746106	3			1	1	Chloroflexi VIII
SPG12_343_353_B18	FJ746107				1		Chloroflexi IV
SPG12_343_353_B2	FJ746108				1		Planctomycetes-Planctomycetaceae
SPG12_343_353_B22	FJ746111				2		Deltaproteobacteria-Desulfobacterales
SPG12_343_353_B23	FJ746112		1		2	3	OP8
SPG12_343_353_B24	FJ746113				1		Planctomycetes I
SPG12_343_353_B28	FJ746115				1		Deltaproteobacteria-Myxococcales
SPG12_343_353_B29	FJ746116				1		OP3
SPG12_343_353_B3	FJ746117				1		Chloroflexi VI
SPG12_343_353_B30	FJ746118		1		3	3	Chloroflexi IV
SPG12_343_353_B31	FJ746119		1		1		Chloroflexi IV
SPG12_343_353_B35	FJ746122	1			1		Planctomycetes II
SPG12_343_353_B38	FJ746124				1		Gemmatimonadetes
SPG12_343_353_B41	FJ746127				1		Chloroflexi IV
SPG12_343_353_B42	FJ746128				1		WS3
SPG12_343_353_B43	FJ746129				1		WS3
SPG12_343_353_B44	FJ746130	1	4		1	2	Chloroflexi-Anaerolineae
SPG12_343_353_B48	FJ746132				1		Planctomycetes II
SPG12_343_353_B5	FJ746133				1	1	Planctomycetes I
SPG12_343_353_B51	FJ746134				1		Chloroflexi IV
SPG12_343_353_B52	FJ746135		1	1	1	1	Chloroflexi IV
SPG12_343_353_B54	FJ746136				1		Chloroflexi VIII
SPG12_343_353_B57	FJ746138				2		Alphaproteobacteria-Kordiimonadales
SPG12_343_353_B61	FJ746141	2			1		Alphaproteobacteria-Rhodospirillaceae
SPG12_343_353_B64	FJ746143	1			2	1	ASG
SPG12_343_353_B66	FJ746144		1		2		Chloroflexi IV
SPG12_343_353_B69	FJ746146				1		Planctomycetes IV

SPG12_343_353_B72	FJ746149				1	1	Planctomycetes I
SPG12_343_353_B73	FJ746150				1		SAR324
SPG12_343_353_B75	FJ746152				1		Chloroflexi IV
SPG12_343_353_B76	FJ746153			1	1		Chloroflexi V
SPG12_343_353_B80	FJ746156				1		Planctomycetes I
SPG12_343_353_B82	FJ746158		6		2	2	Planctomycetes I
SPG12_343_353_B86	FJ746162		1		1		Gemmatimonadetes
SPG12_343_353_B9	FJ746164				2		Spirochaeta-Spirochaetaceae
SPG12_343_353_B92	FJ746167				1		Chloroflexi IV
SPG12_343_353_B99	FJ746169				1		Chloroflexi IV
SPG12_461_471_B36	FJ746171		1			1	Actinobacteria-Acidomicrobiae
SPG12_461_471_B37	FJ746172					1	Alphaproteobacteria-Rhodospirillaceae
SPG12_461_471_B43	FJ746177	1	1			1	Alphaproteobacteria-Rhodospirillaceae
SPG12_461_471_B51	FJ746181					1	Chloroflexi IV
SPG12_461_471_B00	FJ746183					1	SAR406
SPG12_461_471_B56	FJ746187					1	OP11
SPG12_461_471_B69	FJ746196		1			1	Planctomycetes II
SPG12_213_223_B25	FJ746203		1				Planctomycetes II
SPG12_213_223_B26	FJ746204		2				Planctomycetes III
SPG12_213_223_B27	FJ746205		1				Chloroflexi IV
SPG12_213_223_B29	FJ746207		1				Planctomycetes II
SPG12_213_223_B3	FJ746208		4	4			NT-B2
SPG12_213_223_B31	FJ746209		1				Chloroflexi IV
SPG12_213_223_B33	FJ746210		1				Planctomycetes-anammox
SPG12_213_223_B37	FJ746211		1				Chloroflexi IV
SPG12_213_223_B45	FJ746215		1				Deltaproteobacteria-Myxococcales
SPG12_213_223_B46	FJ746216		1				Planctomycetes II
SPG12_213_223_B49	FJ746217		1				OP11
SPG12_213_223_B51	FJ746219		1				Deltaproteobacteria-Uncultured 1
SPG12_213_223_B56	FJ746221		1				Planctomycetes II
SPG12_213_223_B77	FJ746230		1				Planctomycetes I
SPG12_213_223_B89	FJ746237		1				Planctomycetes-Planctomycetaceae
SPG12_213_223_B9	FJ746238		1			1	Planctomycetes-Planctomycetaceae
SPG12_213_223_B91	FJ746240		1				Planctomycetes-Planctomycetaceae
SPG12_213_223_B94	FJ746242		1				OP3
SPG12_213_223_B80	FJ746244		1				Chloroflexi-Anaerolineae
SPG12_461_471_B12	FJ746248					1	Planctomycetes I
SPG12_461_471_B2	FJ746251					1	Chloroflexi IV
SPG12_461_471_B20	FJ746252					1	Chloroflexi VII
SPG12_461_471_B8	FJ746256					1	Chloroflexi IV
SPG12_273_283_B13	FJ746261			1			AC1
SPG12_273_283_B14	FJ746262			2			OP3
SPG12_273_283_B19	FJ746265			1			Termite group I
SPG12_273_283_B26	FJ746271			1			Planctomycetes IV
SPG12_273_283_B3	FJ746274			1			Planctomycetes I
SPG12_273_283_B30	FJ746275			1			Planctomycetes-Planctomycetaceae
SPG12_273_283_B34	FJ746278			1			Chloroflexi IV
SPG12_273_283_B4	FJ746281			1			Bacteroidetes: Sphingobacteria
SPG12_273_283_B5	FJ746282			1			Chloroflexi IV

SPG12_273_283_B6	FJ746283			1			Alphaproteobacteria
SPG12_60_70_B29	FJ746287	4					OP11
SPG12_60_70_B36	FJ746290	1					Planctomycetes-Planctomycetaceae
SPG12_60_70_B37	FJ746291	6					Gemmatimonadetes
SPG12_60_70_B41	FJ746295	5					Alphaproteobacteria-Rhodospirillaceae
SPG12_60_70_B42	FJ746296	1					Nitrospina-Deltaproteobacteria
SPG12_60_70_B49	FJ746299	1					Gammaproteobacteria
SPG12_60_70_B56	FJ746306	1					Chloroflexi VIII
SPG12_60_70_B57	FJ746307	2					Alphaproteobacteria-Rhodospirillaceae
SPG12_60_70_B65	FJ746315	1					Actinobacteria-Acidomicrobiae
SPG12_60_70_B69	FJ746318	1					Planctomycetes II
SPG12_60_70_B70	FJ746319	1					Alphaproteobacteria-Rhodospirillaceae
SPG12_60_70_B71	FJ746320	1					Verrucomicrobia
SPG12_461_471_B31	FJ746326					1	Chloroflexi IV
SPG12_213_223_B2	FJ746331		1				Chloroflexi IV
		37	49	27	86	50	
						249	

Appendix D. 3% OTU phylotype table for Bacteria. Listed for each clone library is number of individual representatives of a given phylotype, with accession number and phylogenetic classification. Phylotype SPG12_343_353_B54 belongs to the Abyssal Sediment Bacterial Cluster, a novel designation, while other abbreviations refer to previously published uncultured clades. Subclades of Planctomycetes and Chloroflexi, some introduced in this study, are included, as are the recognized subclades of other major phyla, when possible.

Appendix E. Supplementary Methods.

In addition to performing blank extractions, several procedures were followed to minimize contamination from exogenous DNA or PCR products for these low-biomass sediments. Poly[dIdC] in UVC-permeable 1.5 mL tubes (MoBio Laboratories) was UV-irradiated for 15 minutes approximately 15 cm from 15-watt UVC bulbs in a PCR hood to reduce exogenous 16S gene contamination and to reduce nonspecific amplification potential during downstream PCR applications. With the exception of the SPG12 60-70 and 343-353 cmbsf samples, all steps except for bead-beating and centrifugation were undertaken in a dead-air PCR hood that was intermittently UVC-irradiated. For all samples, the extractions, including centrifugation steps, were carried out at room temperature to minimize SDS precipitation. Blank extractions with only buffer were also carried out under the same conditions as the appropriate sample. For the SPG12 60-70 cmbsf and 343-353 cmbsf sample and blank extractions, all aqueous solutions were made with nuclease-free PCR-grade water, and were then filtered using an 0.1- µm pore size filter into a flask baked at 180 °C overnight. The aqueous solutions used for the 350 cmbsf sample was additionally

submitted to ~30 minutes of irradiation in a plastic bag ~30 cm from a UVC light source. For all other samples and blank extractions, all aqueous solutions were additionally passed through a 10 kDa-molecular weight cutoff (MWCO) ultrafiltering device, either an Amicon Plus-15 or Amicon Plus-70 (Millipore), to remove exogenous DNA. In the case of the Amicon Plus-70 apparatus, which is not shipped sterile, the receiving cup and filter section were exposed to UVC radiation for 30 minutes to 1 hour. 20% sodium dodecyl sulfate solution proved difficult to pass through a 10 kDa MWCO device and so for all samples this was irradiated for 30 minutes approximately 30 cm from a UVC PCR hood light source in a plastic bag.

Appendix F. Habitat-lineage association sample points.

Clade	Habitat	Example Acc.
DSEG-4		
	Suboxic marine sediments	FJ487460
	Coastal shallow methanogenic sediments	DQ522902
	Hypersaline cyanobacterial mats	EU585938
	Hydrothermal precipitates	AB329826
	Suboxic freshwater pond plankton	DQ676273
	Hydrothermal vent water	AB193961
	Hydrothermal vent water	AB301878
	Cold seep sediments	AB189390
DSEG-3		
	Hydrothermal sediments	AB302043
	Oligotrophic marine sediments	EU385997
	Coastal shallow methanogenic sediments	DQ640163
	Suboxic marine sediments	FJ487470
	Oligotrophic marine sediments	AY800211
	Iron- and sulfur-oxidizing sediment	EF687554
	Hydrothermal sediments	AB301978
	Hot spring water	EF444628
	Mediterranean cold seep	AY592520
	Hydrothermal metal-carbonate sediments	AY354110
	Deep-sea microbial mat	AB426447
	Deep subsurface ground water	AB237754
DSEG-2		
	Hot spring water	AF402988
	Hot spring water	AY555816
	Hydrothermal vent water	AB301873
	Sandy shallow sediment	EF208735
	Suboxic marine sediment	FJ487457
	Hydrothermal vent water	DQ910086

	Hydrothermal vent water	AB213092
	Mediterranean cold seep	AY592549
	Hypersaline lake shallow sediments	EU329799
	Hydrothermal precipitates	AB329763
	Eutrophic margin sediments	AB177269
	Tropical estuary sediments	EF367558
	Estuarine sediments	EU681930
	Hydrothermal vent chimney	AB247870
	Hydrothermal vent chimney	AB293222
DHVE-5		
	Estuarine sediments	AY454684
	Hot spring water	EF444658
	Hydrothermal chimneys and worm tubes	DQ082961
	Suboxic freshwater pond plankton	DQ676246
	Hypersaline lake shallow sediments	EU329766
	Evaporitic salt crust	EF106713
	Oxic/suboxic ponds	AY822001
	Basalt	DQ417472
	Shallow Antarctic shelf sediments	AF424534
	Hypersaline cyanobacterial mats	DQ397352
	Hydrothermal vent water	DQ909977
	Saline lake shallow sediment	AF142929
	Hydrothermal precipitates	AB329839
	Anoxic water column	AF224779
	Hydrothermal sediments	AB197216
	Sargasso Sea surface water	AACY023847387
	Fish farm sediment	AY500128
	Sulfidic cold spring water	AJ631246
	Deep subsurface ground water	AB077228
	Hydrothermal vent colonization surface	AB175599
	Seafloor hypersaline lake	DQ340998
	Suboxic marine sediment	FJ487465
	Hydrothermal vent water	AB019751
	Sulfurous lake water	AM0769832
	Arctic continental shelf water column	EU244290
	Hydrothermal vent chimney	AB293233
	Hydrothermal vent water	AF526959
MBG-A rels		
	Aspen forest soil	EF022172
	Soil	EU306957
	Meso/bathypelagic seawater	EU650259
	Hot spring water	AY861939
	Meso/bathypelagic seawater	EF106847
	Anaerobic marine water column	EU369889
	Hot spring water	AY882837
	Hydrothermal precipitates	AB329801
	Hydrothermal sediments	AB301986
	Soil	DQ278143
	Freshwater lake	AY278081
	Forest soil	AY278100

	Oligotrophic marine sediments	EU385976
	Oligotrophic marine sediments	EU048593
	Hydrothermal vent water	DQ270602
	Suboxic marine sediments	FJ487456
	Shallow marine sediments	AF119135
	Hot spring water	DQ300328
	Meso/bathypelagic seawater	DQ300527
	Estuarine sediments	DQ641811
	Estuarine sediments	EF367446
	Hot spring water	EF444663
	Subsurface aquifer	DQ190081
	Rice paddy soil	AB243804
	Shallow peat mire	AB262707
	Hydrothermal vent water	AB213098
	Hot spring water	AY555832
	Deep subsurface ground water	AB294269
	Soil	EU570091
	Boreal forest soil	AJ428044
	Uranium mill tailings	AJ535136
	Acid mine drainage	DQ901270
	Hydrothermal vent water	AB193977
	Eutrophic margin sediments	AB177271
	Bathypelagic seawater	EF597687
	Estuarine sediments	AY454565
	Mangrove sediment	EF125516
	Subsurface aquifer	DQ837283
	Oligotrophic marine sediments	AJ870947
	Estuarine sediments	EF680224
	Hot spring water	U63343
	Hot spring water	DQ243758
	Freshwater lake	AJ937878
	Hydrothermal chimneys	EU559696
	Hydrothermal chimneys	AB293209
	Hydrothermal sediments	EF100626
	Hydrothermal vent water	AB167488
	Hot spring sediment	EU635918
	Minerotrophic fen	EU155999
pISA7		
	Shallow marine sediments	AJ299201
	Hot spring water	DQ300338
	Hydrothermal chimneys	AB293216
	Hydrothermal sediments	AB302002
	Basalt	DQ417463
	Hydrothermal vent water	AB213072
	Hydrothermal chimneys	EU555152
	Hydrothermal chimneys	AB247868
	Hydrothermal sediments	EF100630
	Hydrothermal vent water	AB301879
	Hydrothermal chimneys	AB424721
	Hydrothermal vent water	AB019733

MCG		
	Estuarine sediments	EF367501
	Estuarine sediments	DQ641777
	Marine sediments/coastal	DQ866031
	Estuarine sediments	EU280210
	Eutrophic margin sediments	AB177001
	Estuarine sediments	AY454602
	Estuarine sediments	EU284664
	Oligotrophic marine sediments	EU048654
	Estuarine sediments	DQ363817
	Bog	EU155991
	Cold seep sediments	AJ579325
	Hot spring water	DQ300325
	Estuarine sediments	EF203611
	Hot spring water	AY861924
	Hot spring sediment	EU635929
	Tidal flat sediment	EF125493
	Hydrothermal mat	AB301882
	Hydrothermal vent water	AB301868
	Hydrothermal fluid	AB213062
	Eutrophic marine sediment	AY627509
	Mediterranean cold seep	AY592466
	Basalt	AY704375
	Lake sediment	AB355118
	Hydrothermal precipitates	AB329809
	Estuarine sediments	AB300142
	Estuarine sediments	EF680226
	Cow rumen	AY464796
	Methanogenic coastal sediment	DQ522918
	Eutrophic margin sediments	AY436515
	Cow manure	EU662691
	Eutrophic margin sediments	AB094558
	Shellfish aquaculture sediment	AB116488
	Eutrophic margin sediments	AY093450
	Terrestrial deep subsurface	AF005762
	Eutrophic margin sediments	AB109883
	Deep subsurface ground water	AB237759
	Soda lake water	DQ302468
	Estuarine sediments	AF276442
	Freshwater lake sediment	U59986
	Freshwater lake	AY278094
	Neutral pH mine biofilm	AY082455
	Sewage sludge	AF424768
	Antarctic shelf shallow sediment	AF424526
	Landfill sludge	AJ831142
	Sludge digestor	U81774
	Sludge digestor	EF552166
	Marsh pore water	AM0557704
	Freshwater particles	AF418925
	Ground water	AJ583395

Mine waste	AJ535133
Rice paddy soil	AB243802
Rice paddy soil	AB288602
Bog	AJ548933
Ground water	AB294256
Anaerobic sludge digestor	AY426476
Oil sand tailings pond	EF420188
Oil contaminated soil	AB161339
Anoxic water column	DQ785306
Ground water	AJ567215
Anaerobic digestor	AY835784
Hot spring water	AY555820
Soil	EU307004
Estuarine sediments	EU681927
Estuarine sediments	EF051149
Saline soil	AJ969773
Freshwater lake	AJ937877
Estuarine sediments	AF118657
Cold seep sediments	AJ704633
Bog	AB263705
Freshwater lake	AM076830
Ground water	DQ230924
Eutrophic marine sediments	AB369211
Marine sediments/coastal	AJ294859
Hot spring water	DQ243757
Hot spring sediment	EU239992
Hot spring water	U63345
Hot spring water	AB113634
Hot spring water	L25305
Hydrothermal sediments	AB302001
Hot spring water	U63339
Basalt	AB260058
Eutrophic margin sediments	AF419653
Eutrophic marine sediments	AY345169
Cold seep sediments	AB189389
Cold seep sediments	AB015273
Eutrophic margin sediments	AF119129
Hot spring water	AM039529
Ground water	DQ841220
Hydrothermal fluid	AB284822
Eutrophic margin sediments	AB176996
Hot spring water	EF444630
Oligotrophic marine sediments	DQ984834
Eutrophic marine sediments	EU369873
Hydrothermal vent water	AB247826
Hydrothermal sediments	AY354121
Eutrophic marine sediments	AF328199
Eutrophic margin sediments	AF004344
Cold seep sediments	AB362542
Eutrophic marine sediment	AB426390

	Cold seep sediments	AB188799
	Eutrophic marine sediments	AY499893

Appendix E. Habitat-lineage association sample points. Habitats association was then further categorized into a generalized set of habitats, which were then categorized as principally or entirely oxic/suboxic, anoxic, or ambiguous, with uncertain or heterogeneous redox state.

References.

Agogué, H., Brink, M., Dinasquet, J., Herndl, G.J. (2008) Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic. *Nature* **456**: 788-792.

Acinas, S.G., Klepac-Ceraj, V., Hunt, D.E., Pharino, C., Ceraj, I., Distel, D.L., Polz, M.F. (2004) Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* **430**:551-554.

Amend, J.P., Shock, E.L. (1998) Energetics of amino acid synthesis in hydrothermal ecosystems. *Science* **281**: 1659-1662.

Bak, F., Widdel, F. (1986) Anaerobic degradation of indolic compounds by sulphate-reducing enrichment cultures, and description of *Desulfobacterium indolicum* gen. nov., sp. nov. *Archives of Microbiology* **146**: 170-176.

Bano, N., Ruffin, S., Ransom, B., Hollibaugh, J.T. (2004) Phylogenetic composition of Arctic Ocean Archaeal assemblages and comparison with Antarctic assemblages. *Applied and Environmental Microbiology* **70**: 781-789.

Barns, S.M., Fundyga, R.E., Jeffries, M.W., Pace, N.R. (1994) Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. *Proceedings of the National Academy of Sciences USA* **91**: 1609-1613.

Barton, H.A., Taylor, N.M., Lubbers, B.R., Pemberton, A.C. (2006) DNA extraction from low-biomass carbonate rock: an improved method with reduced contamination and the low-biomass contaminant database. *Journal of Microbiological Methods* **66**: 21-31.

Biddle, J.F., Lipp, J.S., Lever, M.A., et al. (2006) Heterotrophic Archaea dominate sedimentary subsurface ecosystems off Peru. *Proceedings of the National Academy of Sciences USA* **103**: 3846-3851.

Biddle, J.F., Fitz-Gibbon, S., Schuster, S.C., Brenchley, J.E., House, C.H. (2008) Metagenomic signatures of the Peru Margin subseafloor biosphere show a genetically distinct environment. *Proceedings of the National Academy of Sciences USA* **105**: 10583-10588.

Blazejak, A., Erséus, C., Amann, R., Dubilier, N. (2005) Coexistence of bacterial sulfide oxidizers, sulfate reducers, and spirochetes in a gutless worm (Oligochaeta) from the Peru Margin. *Applied and Environmental Microbiology* **71**: 1553-1561.

Blazejak, A., Kuever, J., Erséus, C., Amann, R., Dubilier, N. (2006) Phylogeny of 16S rRNA, ribulose 1,5-bisphosphate carboxylase/oxygenase, and adenosine 5'-phosphosulfate reductase genes from gamma- and alphaproteobacterial symbionts in gutless marine worms (Oligochaeta) from Bermuda and the Bahamas. *Applied and Environmental Microbiology* **72**: 5527-5536.s

Boussau, B. & Gouy, M. Efficient likelihood computations with nonreversible models of evolution. *Syst. Biol.* **55**, 756–768 (2006).

Brochier-Armanet, C., Boussau, B., Gribaldo, S., Forterre, P. (2008) Mesophilic Crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nature Reviews Microbiology* **6**: 245-252.

Cha, R.S., Thilly, W.G. (1993) Specificity efficiency and fidelity of PCR. *Genome Research* **3**: S18-S29.

Chao, A. 1984. Non-parametric estimation of the number of classes in a population. *Scandinavian Journal of Statistics* **11:** 265–270.

Church, M.J., DeLong, E.F., Ducklow, H.W., et al. (2003) Abundance and distribution of planktonic Archaea and Bacteria in the waters west of the Antarctic Peninsula. *Limnology and Oceanography* **48**: 1893-1902.

Clement, B.G., Luther III, G.W., Tebo, B.M. (2009) Rapid, oxygen-dependent microbial Mn(II) oxidation kinetics at sub-micromolar oxygen concentrations in the Black Sea suboxic zone. *Geochimica et Cosmochimica Acta* **73**: 1878-1889.

Conrad, R. (1996) Soil microorganisms as controllers of atmospheric trace gases (H_2, CO, CH_4, OCS, N_2O, and NO). *Microbiological Reviews* **60**: 609-640.

Coolen, M.J.L., Abbas, B., van Bleijswijk, J. et al. (2007) Putative ammonia-oxidizing Crenarchaeota in suboxic waters of the Black Sea: a basin-wide ecological study using 16S ribosomal and functional genes and membrane lipids. *Environmental Microbiology* **9**: 1001-1016.

Corinaldesi, C., Dell'Anno, A., Danovaro, R. (2007) Viral infection plays a key role in extracellular DNA dynamics in marine anoxic systems. *Limnology and Oceanography* **52**: 508-516.

Cragg, B.A., Kemp, A.E.S. (1995) Bacterial profiles in deep sediment layers from the eastern equatorial Pacific Ocean, Site 851. *Proceedings of the Ocean Drilling Program: Scientific results* **138**: 599-604.

Damsté, J.S.S., Rijpstra, W.I.C., Hopmans, E.C., et al. (2002) Distribution of membrane lipids of planktonic Crenarchaeota in the Arabian Sea. *Applied and Environmental Microbiology* **68**: 2997-3002.

DeLong, E.F. (1992) Archaea in coastal marine environments. *Proceedings of the National Academy of Sciences USA* **89**: 5685-5689.

DeLong, E.F., Wu, K.Y. Prézelin, B.B., Jovine, R.V.M. (1994) High abundance of Archaea in Antarctic marine picoplankton. *Nature* **371**: 695-697.

DeLong, E.F., Preston, C.M., Mincer, T., et al. (2006) Community genomics among stratified microbial assemblages in the ocean's interior. *Science* **311**: 496-503.

DeSantis, T. Z., Hugenholtz, P., Keller, K., et al. (2006). NAST: a multiple sequence alignment server for comparative analysis of 16S rRNA genes. *Nucleic Acids Research* **34**: W394-9.

D'Hondt, S., Rutherford, S., Spivack, A.J. (2002) Metabolic activity of subsurface life in deep-sea sediments. *Science* **295**:2067-2070.

D'Hondt, S.L., Jørgensen, B.B., Miller, D.J., et al. (2003). Proceedings of the Ocean Drilling Program, Initial Reports. Volume 201.

D'Hondt, S., Jørgensen, B.B., Miller, D.J., et al. (2004) Distribution of microbial activities in deep subseafloor sediments. *Science* **306**: 2216-2221.

D'Hondt, S., Spivack, A.J., Pockalny, R., Ferdelman, T.G., Fischer, J.P., Kallmeyer, J., Abrams, L.J., Smith, D.C., Graham, D., Hasiuk, F., Schrum, H., Stancin, A.M. (2009) Subseafloor sedimentary life in the South Pacific Gyre. *Proceedings of the National Academy of Sciences USA*. Published online before print June 26, 2009, doi: 10.1073/pnas.0811793106

Dias, A.S., Barriga, F.J.A.S. (2006) Minerology and geochemistry of hydrothermal sediments from the serpentinite-hosted Saldanha hydrothermal field (36° 34' N; 33° 26' W) at MAR. *Marine Geology* **225:** 157-175.

Dunne, J.P., Sarmiento, J.L., Gnanadesikan, A. (2007) A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor. Global Biogeochemical Cycles **21**:

GB4006.

Durbin, A. (2009) Microbial diversity of oligotrophic marine sediments. Master's thesis, University of North Carolina, Chapel Hill.

Edwards, K.J., Goebel, B.M., Rodgers, T.M., Schrenk, M.O., Gihring, T.M. Cardona, M.M., Hu, B. McGuire, M.M., Hamers, R.J., Pace, N.R., Banfield, J.F. (1999) Geomicrobiology of pyrite (FeS_2) dissolution: case study at Iron Mountain, California. *Geomicrobiology Journal* **16**: 155-179.

Elshahed, M.S., Youssef, N.H., Luo, Q., Najar, F.Z., Roe, B.A., Sisk, T.M., Bühring, S.I., Hinrichs, K.-U., Krumholz, L.R. (2007) Phylogenetic and metabolic diversity of the phylum *Planctomycetes* from an anaerobic, sulfide and sulfur-rich spring (Zodletone Spring, OK, USA). *Applied and Environmental Microbiology* **73**: 4707-4716.

Emerson, D., Moyer, C. (1997) Isolation and characterization of novel iron-oxidizing bacteria that grow at circumneutral pH. *Applied and Environmental Microbiology* **63**: 4784-4792.

Fischer, J.P., Ferdelman, T.G., D'Hondt, S., Røy, H., Wenzhöfer, F. (2009) Oxygen penetration deep into the sediment of the South Pacific gyre. *Biogeosciences Discussion*. **6**: 3159-3186.

Frigaard, N.-U., Martinez, A., Mincer, T.J., DeLong, E.F. (2006) Proteorhodopsin lateral gene transfer between marine planktonic Bacteria and Archaea. Nature **439**: 847-850.

Fry, J.C., Parkes, R.J., Cragg, B.A., Weightman, A.J., Webster, G. (2008) Prokaryotic biodiversity and activity in the deep subseafloor biosphere. *FEMS Microbiology Ecology* **66**: 181-196.

Fuhrman, J.A., McCallum, K., Davis, A.A. (1992) Novel major archaebacterial group from marine plankton. *Nature* **356**: 148-149.

Fuhrman, J.A., Davis, A.A. (1997) Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences. Marine Ecology Progress Series **150**: 275-285.

Galand, P.E., Casamayor, E.O., Kirchman, D.L., Potvin, M., Lovejoy, C. (2009) Unique archaeal assemblages in the Arctic Ocean unveiled by massively parallel tag sequencing. *The ISME Journal* 1-10.

Gieskes, J.M., Boulègue, J. (1986) Interstitial water studies, Leg 92. In: Leinen, M., Rea, D.K., et al., Init. Repts. DSDP, 92: Washington (US Govt. Printing Office).

Gillan, D.C., Danis, B. (2007) The archaebacterial communities in Antarctic bathypelagic sediments. *Deep-Sea Research II* **54**: 1682-1690.

Glöckner, F.O., Kube, M., Bauer, M., Teeling, H., Lombardot, T., Ludwig, W., Gade, D., Beck, A., Borzym, K., Heitmann, K., Rabus, R., Schlesner, H., Amann, R., Reinhardt, R. (2003) Complete genome sequence of the marine planctomycete *Pirellula* sp. strain 1. *Proceedings of the National Academy of Sciences USA* **100**: 8298-8303.

Glud, R.N. (2008) Oxygen dynamics of marine sediments. *Marine Biology Research* **4**: 243-289.

Guermazi, S., Daegelen, P., Dauga, C., Riviere, D., Bouchez, T., Godon, J.J., Gyapay, G., Sghir, A. Pelletier, E., Weissenbach, J., Le Paslier, D. (2008) Discovery and characterization of a new bacterial candidate division by an anaerobic sludge digestor: a metagenomic approach. *Environmental Microbiology* **10**: 2111-2123.

Hallam, S.J., Mincer, T.J., Schleper, C., Preston, C.M., Roberts, K., Richardson, P.M., DeLong, E.F. (2006) Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLoS Biology* **4**:e95.

Harris, J.K., Kelley, S.T., Pace, N.R. (2004) New perspective on uncultured bacterial phylogenetic division OP11. *Applied and Environmental Microbiology* **70**: 845-849.

Harrison, B.K., Zhang, H., Berelson, W., Orphan, V.J. (2009) Variations in archaeal and bacterial diversity associated with the sulfate-methane transition zone in continental margin sediments (Santa Barbara Basin, California). *Applied and Environmental Microbiology* **75**: 1487-1499.

Hartnett, H.E., Keil, R.G., Hedges, J.I., Devol, A.H. (1998) Influence of oxygen exposure time on organic carbon preservation in continental margin sediments. *Nature* **39**: 572-574.

Hedges, J.I., Baldock, J.A., Gélinas, Y., Lee, C., Peterson, M., Wakeham, S.G. (2001) Evidence for non-selective preservation of organic matter in sinking marine particles. *Nature* **409**: 801-804.

Heijs, S.K., Haese, R.R., van der Wielen, W.J.J., Forney, L.J., van Elsas, J.D. (2007) Use of 16S rRNA gene based clone libraries to assess microbial communities potentially involved in anaerobic methane oxidation in a Mediterranean cold seep. *Microbial Ecology* **53**: 384-398.

Heijs, S.K., Laverman, A.M., Forney, L.J., Hardoim, P.R., van Elsas, J.D. (2008) Comparison of deep-sea sediment microbial communities in the Eastern Mediterranean. *FEMS Microbiology Ecology* **64**:362-377.

Herndl, G.J., Reinthaler, T., Teira, E., et al. (2005) Contribution of Archaea to total prokaryotic production in the deep Atlantic Ocean. *Applied and Environmental Microbiology* **71**: 2303-2309.

Hershberger, K.L., Barns, S.M., Reysenbach, A.-L., Dawson, S.C., Pace, N.R. (1996) Wide diversity of Crenarchaeota. *Nature* **384**: 420-420.

Hoehler, T.M. (2004) Biological energy requirements as quantitative boundary conditions for life in the subsurface. ***2***: 205-215.

Hong, S.-H., Bunge, J., Jeon, S.-O., Epstein, S.S. (2006) Predicting microbial species richness. *Proceedings of the National Academy of Sciences USA* **103**: 117-122.

Hongoh, Y., Sharma, V.K., Prakash, T., Noda, S., Taylor, T., Kudo, T., Sakaki, Y., Toyoda, A., Hattori, M., Ohkuma, M. (2008) Complete genome of the uncultured Termite Group I bacteria in a single host protist cell. *Proceedings of the National Academy of Sciences USA* **105**: 5555-5560.

Huber, J.A., Morrison, S.M., Neal, P.R., et al. (2009) Effect of PCR amplicon size on assessments of clone library microbial diversity and community structure. *Environmental Microbiology*. doi:10.1111/j.1462-2920.2008.01857.x.

Huber, T., Faulkner, G., Hugenholtz, P. (2004) Bellerophon: a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics* **20**: 2317-2319.

Hugenholtz, P., Stackebrandt, E. (2004) Reclassification of *Sphaerobacter thermophilus* from the subclass *Sphaerobacteridae* in the phylum *Actinobacteria* to class *Thermomicrobia* (emended description) in the phylum *Chloroflexi* (emended description). *International Journal Systematic and Evolutionary Microbiology* **54**: 2049-2051.

Hughes, J., Hellmann, J., Ricketts, T., Bohannan, B. (2001) Counting the uncountable: statistical approaches to estimating microbial diversity. *Environmental Microbiology* **67**: 4399-4406.

Hughes Martiny, J.B., Bohannan, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., Horner-Devine, M.C., Kane, M., Adams Krumins, J., Ruske, C.R., Morin, P.J., Naeem, S., Øvreås, Reysenbach, A.-L., Smith, V.H., Staley, J.T. (2006) Microbial biogeography: putting microorganisms on the map. *Nature Reviews Microbiology* **4**: 102-112.

Hulth, S., Aller, R.C., Gilbert, F. (1999) Coupled anoxic nitrification/manganese reduction in marine sediments. *Geochimica et Cosmochimica Acta* **63**: 49-66.

Inagaki, F., Takai, K., Komatsu, T. (2001) Archaeology of Archaea: geomicrobiological record of Pleistocene thermal events concealed in a deep-sea subseafloor environment. *Extremophiles* **5**: 385-392.

Inagaki, F., Suzuki, M., Takai, K., Oida, H., Sakamoto, T., Aoki, K. et al. (2003). Microbial communities associated with geological horizons in coastal subseafloor sediments from the Sea of Okhotsk. *Applied and Environmental Microbiology* **69**: 7224–7235.

Inagaki, F., Nunoura, T., Nakagawa, S., et al. (2006) Biogeographical distribution and diversity of microbes in methane hydrate-bearing deep marine sediments on the Pacific Ocean Margin. *Proceedings of the National Academy of Sciences USA* **103**: 2815-2820.

Ingalls, A.I., Shah, S.R., Hansman, R.L., et al. (2006) Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proceedings of the National Academy of Sciences USA* **103**: 6442-6447.

Isenbarger, T.A., Finney, M., Ríos-Velázquez, C., Handelsman, J., Ruvkun, G. (2008) Miniprimer PCR, a new lens for viewing the microbial world. *Applied and Environmental Microbiology* **74**: 840-849.

Jakobsen, R. (2007) Redox microniches in groundwater: A model study on the geometric and kinetic conditions required for concomitant Fe oxide reduction, sulfate reduction, and methanogenesis. *Water Resource Research* **43**: W12S12.

Jobb, G., von Haeseler, A., Strimmer, K. (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evolutionary Biology* **4**:18.

Jørgensen, B.B. (1977) Bacterial sulfate reduction within reduced microniches of oxidized marine sediments. *Marine Biology* **41**: 7-17.

Jørgensen, B.B., Glud, R.N., Holby, O. (2005) Oxygen distribution and bioirrigation in Arctic fjord sediments (Svalbard, Barents Sea). *Marine Ecology Progress Series* **292**: 89-95.

Kallmeyer, J., Ferdelman, T.G., Weber, A., Fossing, H., Jorgensen, B.B. (2004) *Limnology and Oceanography Methods* **2**:171-180.

Karner, M.B., DeLong, E.F., Karl, D.M. (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507-510.

Ku, T.C.W., Kay, J., Browne, E., Martini, A.M., Peters, S.C., Chen, M.D. (2008) Pyritization in tropical coastal sediments: Implications for the development of iron, sulfur, and carbon diagenetic properties, Saint Lucia, Lesser Antilles. *Marine Geology* **249**: 184-205.

Jurgens, G., Lindström, Saano, A. (1997) Novel group within the kingdom Crenarchaeota from boreal forest soil. *Applied and Environmental Microbiology* **63**: 803-805.

Kallmeyer, J., Smith, D.C., Spivack, A.J., D'Hondt, S. (2008) New cell extraction procedure applied to deep subsurface sediments. *Limnology and Oceanography:Methods* **6**: 236-245.

Karner, M.B., DeLong, E.F., Karl, D.M. (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507-510.

Kelley, D.S., Karson, J.A., Blackman, D.K., Früh-Green, G.L., Butterfield, D.A., Lilley, M.D., Olson, E.J.,

Schrenk, M.O., Roe, K.K., Lebon, G.T., Rivizzigno, P., & the AT3-60 Shipboard Party (2001) An off-axis hydrothermal vent field near the Mid-Atlantic Ridge at 30°N. *Nature* **412**: 145-149.

Kendall, M.M., Wardlaw, G.D., Tang, C.F., Bonin, A.S., Liu, Y., Valentine, D.L. (2007) Diversity of Archaea in marine sediments from Skan Bay, Alaska, including cultivated methanogens, and a description of *Methanogenium boonei* sp. nov. *Applied and Environmental Microbiology* **73**: 407-414.

Kirkpatrick, J., Oakley, B., Fuchsman, C., Srinivasan, S., Staley, J.T., Murray, J.W. (2006) Diversity and distribution of *Planctomycetes* and related bacteria in the suboxic zone of the Black Sea. *Applied and Environmental Microbiology*, **72**:3079-3083.

Kittelmann, S., Friedrich, M.W. (2008) Novel uncultured *Chloroflexi* dechlorinate perchloroethene to *trans*-dichloroethene in tidal flat sediments. *Environmental Microbiology* **10**: 1557-1570.

Klepac-Ceraj, V., Bahr, M., Crump, B.C., Teske, A.P., Hobbie, J.E., Polz, M.F. (2004) High overall diversity and dominance of microdiverse relationships in salt marsh sulfate-reducing bacteria. *Environmental Microbiology* **6**: 686-698.

Köhler, T., Stingl, U., Meuser, K., Brune, A. (2008) Novel lineages of *Planctomycetes* densely colonize the alkaline gut of soil-feeding termites (*Cubitermes* spp.). *Environmental Microbiology* **10**: 1260-1270.

Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker C.B., Waterbury J.B., Stahl, D.A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543-546.

Kumar, S., Dudley, J., Nei, M., Tamura, K., (2008) MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* **9**: 299-306.

Kuypers, M.M.M., Lavik, G., Woebken, D., et al. (2005) Massive nitrogen loss from the Benguela upwelling system through anaerobic ammonium oxidation. *Proceedings of the National Academy of Sciences USA* **102**: 6478-6483.

Lane, D.J. (1991) 16S/23S rRNA sequencing. In *Nucleic Acid Techniques in Bacterial Systematics*. Stackebrandt, E., and Goodfellow, M. (eds). New York, USA: John Wiley & Sons, pp. 115-148.

Lam, P., Jensen, M.M., Lavik, G., et al. (2007) Linking crenarchaeal and bacterial nitrification to annamox in the Black Sea. *Proceedings of the National Academy of Sciences USA* **104**: 7104-7109.

Lee, C., Wakeham, S., Arnosti, C. (2004) Particulate organic matter in the sea: the composition conundrum. *Ambio* **33**: 559-568.

Lever, M.A., Alperin, M., Inagaki, F., Nakagawa, S., Steinsbu, B.O., Teske, A., (2006) Trends in basalt and sediment core contamination during IODP expedition 301. *Geomicrobiology Journal* **23**: 517-530.

Li, Y., Li, F., Zhang, X. et al. (2008) Vertical distribution of bacterial and archaeal communities along discrete layers of a deep-sea cold sediment sample at the East Pacific Rise (13°N). *Extremophiles* **12**: 575-585.

Lipp, J.S., Morono, Y., Inagaki, F., Hinrichs, K.-U. (2008) Significant contribution of Archaea to extant biomass in marine subsurface sediments. *Nature* **454**: 991-994.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Research* **32**(4): 1363-1371.

Luther III, G.W., Sundby, B., Lewis, B.L., Brendel, P.J., Silverberg, N. (Interactions of manganese with the nitrogen cycle: Alternative pathways to dinitrogen. *Geochimica et Cosmochimica Acta* **61**: 4043-4052.

Magot, M., Fardeau, M.L., Arnauld., O., Lanau, C., Ollivier, B., Thomas, P., Patel, B.K.C. (1997) *Spirochaeta smaragdinae* sp. nov., a new mesophilic strictly anaerobic spirochete from an oil field. *FEMS Microbiology Letters* **155**: 185-191.

Mäkelä, K., Tuominen, L. (2003) Pore water nutrient profiles and dynamics in soft bottoms of the northern Baltic Sea. *Hydrobiologia* **492**: 43-53.

Massana, R., DeLong, E.F., Pedros-Alio, C. (2000) A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Applied and Environmental Microbiology* **66**: 1777-1787.

McCollom, T.M., Amend, J.P. (2005) A thermodynamic assessment of energy requirements for biomass synthesis by chemolithoautotrophic micro-organisms in oxic and anoxic environments. *Geobiology* **3**:135-144.

Mincer, T.J., Church, M.J., Taylor, L.T., Preston, C., Karl, D.M., DeLong, E.F. (2007) Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific Subtropical Gyre. *Environmental Microbiology* **9**: 1162-1175.

Morris, R.M., Rappé, M.S., Urbach, E., Connon, S.A., Giovannoni, S.J. (2004) Prevalence of the *Chloroflexi*-related SAR202 bacterioplankton cluster throughout the mesopelagic zone and deep ocean. *Applied and Environmental Microbiology* **70**: 2836-2842.

Moussard, H., Moreira, D., Cambon-Bonavita, M.-A., López-García, P., Jeanthon, C. (2006) Uncultured Archaea in a hydrothermal microbial assemblage: phylogenetic diversity and characterization of a genome fragment from a euryarchaeote. *FEMS Microbiology Ecology* **57**:452-469.

Nei, M., Kumar, S. (2000) *Molecular Evolution and Phylogenetics*. New York: Oxford University Press.

Nercessian, O., Reysenbach, A.-L., Prieur, D., Jeanthon, C. (2003) Archaeal diversity associated with *in situ* samplers deployed on hydrothermal vents on the East Pacific Rise (13° N). *Environmental Microbiology* **5**: 492-502.

Nercessian, O., Fouquet, Y., Pierre, C., Prieur, D., Jeanthon, C. (2005) Diversity of Bacteria and Archaea associated with carbonate-rich metalliferous sediment sample from the Rainbow vent field on the Mid-Atlantic Ridge. *Environmental Microbiology* **7**: 698-714.

Nunoura, T., Hirayama, H., Takami, H., Oida, H., Nishi, S., Shimamura, S., Suzuki, Y., Inagaki, F., Takai, K., Nealson, K.H., Horikoshi, K. (2005) Genetic and functional properties of uncultivated thermophilic crenarchaeotes from a subsurface gold mine as revealed by analysis of genome fragments. *Environmental Microbiology* **7**: 1967-1984.

Ochsenreiter, T., Selezi, D., Quaiser, A., et al. (2003) Diversity and abundance of Crenarchaeota in terrestrial habitats studied by 16S RNA surveys and real time PCR. *Environmental Microbiology* **5**: 787-797.

Odum, E.P., Barrett, G.W. (2005) Fundamentals of ecology. Belmont, CA: Thomson Brooks/Cole.

Ohkuma, M., Kudo, T. (1996) Phylogenetic diversity of the intestinal bacterial community in the termite *Reticulitermes speratus*. *Applied and Environmental Microbiology* **62**: 461-468.

Orphan, V.J., House, C.H., Hinrichs, K.U., McKeegan, K.D. & DeLong, E.F. (2002) Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proceedings of the National Academy of Sciences USA* **99**: 7663–7668.

Parkes, R.J., Cragg, B.A., Bale, S.J., Getliff, J.M., Goodman, K., Rochelle, P.A., Fry, J.C., Weightman, A.J.,

Harvey, S.M. (1994) Deep bacterial biosphere in Pacific Ocean sediments. *Nature* **371**: 410-413.

Parkes, R.J., Webster, G., Cragg, B.A., et al. (2005) Deep sub-seafloor prokaryotes stimulated at interfaces over geologic time. *Nature* **436**: 390-394.

Passow, U. (2002) Transparent exopolymer particles (TEP) in aquatic environments. *Progress in Oceanography* **55**:287-333.

Posada, D., Crandall K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 817-818.

Poulton, S.W., Raiswell, R. (2002) The low-temperature geochemical cycle of iron: from continental fluxes to marine sediment deposition. *American Journal of Science* **302**: 774-805.

Preston, C.M., Wu, K.Y., Molinski, T.F., DeLong, E.F. (1996) A psychrophilic crenarchaeon inhabits a marine sponge: Cenarchaeum symbiosum gen. nov., sp. nov. *Proceedings of the National Academy of Sciences USA* **93**: 6241-6246.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B., Ludwig, W., Peplies, J., Glöckner, F. O. (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research* **35**: 7188-7196.

Reed, D.W., Fujita, Y., Delwiche, M.E., Blackwelder, D.B., Sheridan, P.P., Uchida, T., Colwell, F.S. (2002) Microbial communities from methane hydrate-bearing deep marine sediments in a forearc basin. *Applied and Environmental Microbiology* **68**: 3759-3770.

Rodrigo, A.G. (1993) Calibrating the bootstrap test of monophyly. *International Journal for Parasitology* **23**: 507-514.

Rogers, K.L., Amend, J.P. (2006) Archaeal diversity and geochemical energy yields in a geothermal well on Vulcano Island, Italy. *Geobiology* **3**: 319-322.

Roose-Amsaleg, C.L., Garnier-Sillam, E., Harry, M. (2001) Extraction and purification of microbial DNA from soil and sediment samples. *Applied Soil Ecology* **18**: 47-60.

Saitou, N., Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406-425.

Sanderson, M.J., Wojcichowski, M.F. (2000) Improved bootstrap confidence limits in large-scale phylogenies, with an example from Neo-Astragalus (Leguminosae). *Systematic Biology* **49**: 671-685.

Schippers, A., Jørgensen, B.B. (2001) Oxidation of pyrite and iron sulfide by manganese dioxide in marine sediments. *Geochimica et Cosmochimica Acta* **65**: 915-922.

Schippers, A., Jørgensen, B.B. (2002) Biogeochemistry of pyrite and iron sulfide oxidation in marine sediments. *Geochimica et Cosmochimica Acta* **66**: 85-92.

Schlesner, H., and E. Stackebrandt. 1986. Assignment of the genera *Planctomyces* and *Pirella* to a new family *Planctomycetaceae* fam. nov. and description of the order *Planctomycetales* ord. nov. *Systematic and Applied Microbiology* **8**:174-176.

Schloss, P.D., Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Applied and Environmental Microbiology* **71**(3): 1501-1506.

Schloss, P.D., Handelsman, J. (2006) Introducing SONS, a tool for operational taxonomic unit-based

comparisons of microbial community memberships and structures. *Applied and Environmental Microbiology* **72**(10): 6773-6779.

Schloss, P.D., Handelsman, J. (2006) Introducing TreeClimber, a test to compare microbial community structures. *Applied and Environmental Microbiology* **72**(4): 2379-2384.

Schmid, M.C., Risgaard-Petersen, N., van de Vossenberg, J., Kuypers, M.M.M., Lavik, G., Petersen, J., Hulth, S., Thamdrup, B., Canfield, D., Dalsgaard, T., Rysgaard, S., Sejr, M.K., Strous, M., op den Camp, H.J.M., Jetten, M.S.M. (2007) Anaerobic ammonium-oxidizing bacteria in marine environments: widespread occurrence but low diversity. *Environmental Microbiology* **9**: 1476-1484.

Seiter, K., Hensen, C., Schroter, J., Zabel, M. (2004) Organic carbon content in surface sediments-defining regional provinces. *Deep-Sea Research I* **51**: 2001-2026.

Severmann, S., Mills, R.A., Palmer, M.R., Telling, J.P., Cragg, B., Parkes, R.J. (2006) The role of prokaryotes in subsurface weathering of hydrothermal sediments: A combined geochemical and microbiological investigation. *Geochimica et Cosmochimica Acta* **70**:1677-1694.

Shi, Y. (2005) Measurement of in situ expression of proteorhodopsin genes at the North Pacific central gyre station Aloha. Master's thesis, University of Maryland, College Park, Maryland.

Sitnikova, T., Rzhetsky, A., Nei, M. (1995) Interior-branch and bootstrap tests of phylogenetic trees. *Molecular Biology and Evolution* **12**: 319-333.

Smith, W., Solow, A.R., Preston, P.E. (1996) An estimator of species overlap using a modified beta-binomial model. *Biometrics* **52**: 1472-1477.

Sørensen, K.B, Lauer, A., Teske, A. (2004) Archaeal phylotypes in a metal-rich and low-activity deep subsurface sediment of the Peru Basin, ODP Leg 201, Site 1231. *Geobiology* **2**: 151-161.

Sørensen, K.B., Teske, A. (2006) Stratified communities of active Archaea in deep marine subsurface sediments. *Applied and Environmental Microbiology* **72**: 4596-4603.

Spear, J.R., Walker, J.J., McCollom, T.M., Pace, N.R. (2005) Hydrogen and bioenergetics in the Yellowstone geothermal ecosystem. *Proceedings of the National Academy of Sciences USA* **102**: 2555-2560.

Stackebrandt, E., Goebel, B.M. (1994) Taxonomic Note: A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic Bacteriology* **44**:846-849.

Stackebrandt, E., Ebers, J. (2009) Taxonomic parameters revisited: tarnished gold standards. *Microbiology Today*, November 6 issue, pp. 152-155.

Straub, K.L., Benz, M., Schink, B., Widdel, F. (1996) Anaerobic, nitrate-dependent microbial oxidation of ferrous iron. *Applied and Environmental Microbiology* **62**:1458-1460.

Takai, K., Horikoshi, K. (1999) Genetic diversity of Archaea in deep-sea hydrothermal vent environments. *Genetics* **152**: 1285-1297.

Takai, K., Moser, D.P., DeFlaun, M., Onstott, T.C., Fredrickson, J.K. (2001a) Archaeal diversity in waters from deep South African gold mines. *Applied and Environmental Microbiology* **67**: 5750-5760.

Takai, K., Komatsu, T., Inagaki, F., Horikoshi, K. (2001b) Distribution of Archaea in a black smoker chimney structure. *Applied and Environmental Microbiology* **67**: 3618-3629.

Takai, K., Oida, H., Suzuki, Y., et al. (2004) Spatial distribution of Marine Crenarchaeota Group I in the vicinity of deep-sea hydrothermal systems. *Applied and Environmental Microbiology* **70**: 2404-2413.

Tamura K, Dudley J, Nei M & Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.

Tao, L., Peng, W., Pinxian, W. (2008) Microbial diversity in surface sediments of the Xisha Trough, the South China Sea. *Acta Ecologica Sinica* **28**: 1166-1173.

Tateno, Y., Takezaki, N., Nei, M. (1994) Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. *Molecular Biology and Evolution* **11**: 261-277.

Teira, E., Lebaron, P., van Aken, H., Herndl, G.J. (2006) Distribution and activity of Bacteria and Archaea in the deep water masses of the North Atlantic. *Limnology and Oceanography* **51**: 2131-2144.

Telford, M.J., Wise, M.J., Gowri-Shankar, V. (2005) Consideration of RNA secondary structure significantly improves likelihood-based estimates of phylogeny: Examples from the Bilateria. *Molecular Biology and Evolution* **22**: 1129-1136.

Teske, A., Hinrichs, K.-U., Edgcomb, V., et al. (2002) Microbial diversity of hydrothermal sediments in the Guaymas Basin: evidence for anaerobic methanotrophic communities. *Applied and Environmental Microbiology* **68**: 1994-2007.

Teske, A., Lauer, A., Sørensen, K. (2005). Prokaryotes in deep marine subsurface sediments: Phylogenetic diversity of a global extreme habitat. Abstract, Biennial Meeting of the NASA-Astrobiology Institute (NAI), April 10-14, Boulder, CO, USA.

Teske, A., Sørensen, K.B. (2008) Uncultured archaea in deep marine subsurface sediments: have we caught them all? *The ISME Journal* **2**: 3-18.

Thamdrup, B., Dalsgaard, T. (2000) The fate of ammonium in anoxic manganese oxide-rich marine sediment. *Geochimica et Cosmochimica Acta* **64**: 4157-4164.

Thomas, C.A., Bendell-Young, L.I. (1999) The significance of diagenesis versus riverine input in contributing to the sediment geochemical matrix of iron and manganese in an intertidal region. *Estuarine, Coastal and Shelf Science* **48**: 635-647.

Thompson, J.R., Pacocha, S., Pharino, C., Klepac-Ceraj, V., Hunt, D.E., Benoit, J., Sarma-Ruparvtarm, Distel, D.L., Polz, M.F. (2005) Genotypic diversity within a natural coastal bacterioplankton population. *Science* **307**: 1311-1313.

Trouwborst, R.E., Clement, B.G., Tebo, B.M., Glazer, B.T., Luther III, G.W. (2006) Soluble Mn(III) in suboxic zones. *Science* **313**: 1955-1957.

Varela, M.M., van Aken, H., Sintes, E., Herndl, G.J. (2008) Latitudinal trends of Crenarchaeota and Bacteria in the meso- and bathypelagic water masses of the Eastern North Atlantic. *Environmental Microbiology* **10**: 110-124.

Vetriani, C., Jannasch, H.W., MacGregor, B.J., Stahl, D.A., Reysenbach, A.-L. (1999) Population structure and phylogenetic characterization of marine benthic archaea in deep-sea sediments. *Applied and Environmental Microbiology* **65**: 4375-4384.

Wakeham, S.G., Lee, C., Hedges, J.I., Hernes, P.J., Peterson, M.L. (1997) Molecular indicators of diagenetic status in marine organic matter. *Geochimica et Cosmochimica Acta* **61**: 5363-5369.

Wang, P., Wang, F., Xu, M., Xiao, X. (2004) Molecular phylogeny of methylotrophs in a deep-sea sediment from a tropical west Pacific Warm Pool. *FEMS Microbiology Ecology* **47**: 77-84.

Wang, P., Xiao, X., Wang, F. (2005) Phylogenetic analysis of Archaea in the deep-sea sediments of west Pacific Warm Pool. *Extremophiles* **9**: 209-217.

Webster, G., Newberry, C.J., Fry, J.C., Weightman, A.J. (2003) Assessment of bacterial community structure in the deep sub-seafloor biosphere by 16S rDNA-based techniques: a cautionary tale. *Journal of Microbiological Methods* **55**: 155-164.

Wellsbury, P., Mather, I., Parkes, R.J. (2002) Geomicrobiology of deep, low organic carbon sediments in the Woodlark Basin, Pacific Ocean. *FEMS Microbiology Ecology* **42**: 59-70.

Wenzhöfer, F., Glud, R.N. (2002) Benthic carbon mineralization in the Atlantic: a synthesis based on in-situ data from the last decade. *Deep-Sea Research I* **49**: 1255-1279.

Whitman, W.B., Coleman, D.C., Wiebe, W.J. (1998) Prokaryotes: the unseen majority. *Proceedings of the National Academy of Sciences USA* **95**:6578-6583.

Woese, C.R., Winker, S., Gutell, R.R. (1990) Architecture of ribosomal RNA: Constraints on the sequence of "tetra-loops." *Proceedings of the National Academy of Sciences USA* **87**: 8467-8471.

Wright, T.D., Vergin, K.L., Boyd, P.W., Giovannoni,, S.J. (1997) A novel δ-subdivision proteobacterial lineage from the lower ocean surface layer. *Applied and Environmental Microbiology* **63**: 1441-1448.

Wuchter, C., Abbas, B., Coolen, M.J.L., et al. (2006) Archaeal nitrification in the ocean. *Proceedings of the National Academy of Sciences USA* **103**: 12317-12322.

Xu, M., Wang, P., Wang, F., Xiao, X. (2005) Microbial diversity at a deep-sea station of the Pacific nodule province. *Biodiversity and Conservation* **14**: 3363-3380.

Yue, J.C., Clayton, M.K., Lin, F.C (2001) A nonparametric estimator of species overlap. *Biometrics* **57**: 743-749.

Yue, J.C., Clayton, M.K. (2005) A similarity measure based on species proportions. *Communications in Statistical Theory and Methods* **34**: 2123-2131.

Zhou, J., Bruns, M.A., Tiedje, J.M. (1996) DNA recovery from soils of diverse composition. *Applied and Environmental Microbiology* **62**:316-322.