Manuscript Draft

The Guaymas Basin Subseafloor Sedimentary Archaeome Reflects Complex Environmental Histories

Gustavo A. Ramírez[1,2*], Luke J. McKay[3,4], Matthew W. Fields[4,5], Andrew Buckley[1], Carlos Mortera[6], Christian Hensen[7], Ana Christina Ravelo[8] & Andreas P. Teske[1]

Author affiliations:
1: Department of Marine Sciences, University of North Carolina at Chapel Hill, NC, USA.
2: College of Veterinary Medicine, Western University of Health Sciences, Pomona, CA, USA.
3: Department of Land Resources and Environmental Sciences, Montana State University, Bozeman, MT, USA.
4: Center for Biofilm Engineering, Montana State University, Bozeman, MT, USA.
5: Department of Microbiology and Immunology, Montana State University, Bozeman, MT, USA.
6: Instituto de Geofisica, Universidad Nacional Autónoma de México, Coyoacán, México.
7: GEOMAR Helmholtz Centre for Ocean Research, Kiel, Germany.
8: Ocean Sciences Department, University of California, Santa Cruz, CA, USA.

*: Corresponding author

Correspondence:
Gustavo A. Ramírez
gara1985@email.unc.edu
gramirez@westernu.edu
zombiephylotype@gmail.com

Abstract

We explore archaeal distribution and environmental niche differentiation in sedimentary subseafloor habitats of Guaymas Basin and the adjacent Sonora Margin, located in the Gulf of California, México. Specifically, we survey diverse subseafloor habitats on the Guaymas Basin flanking regions that are extending from the spreading center, termed here "off-axis" sites. Sampling locations include (i) control sediments without hydrothermal or seep influence, (ii) Sonora Margin sediments underlying oxygen minimum zone water, (iii) compacted, highly reduced sediments from a pressure ridge with numerous seeps at the base of the Sonora Margin, and (iv) sediments impacted by hydrothermal circulation at the off-axis Ringvent site. Generally, archaeal 16S rRNA gene datasets are largely comprised of Bathyarchaeal lineages, members of the Hadesarchaea, MBG-D, TMEG, and ANME-1 groups. The most frequently observed 25 OTUs belong to members of these lineages, and correspond to approx. 40 to 80% of the sequence dataset in each sediment sample. Differential distribution patterns of these archaeal groups in downcore sediments uniquely characterize each major sedimentary environment. Variations in archaeal community composition reflect locally specific environmental challenges throughout the greater Guaymas Basin area. Background sediments are divided into surface and subsurface niches, reflecting increased selection of the archaeal community downcore. In sum, the environmental setting and history of a particular site, not isolated biogeochemical properties out of context, control the subseafloor archaeal communities in Guaymas Basin and Sonora Margin sediments.

Introduction

Guaymas Basin, located in the Gulf of California, México, is a young marginal rift basin where active seafloor spreading generates northeast-to-southwest trending axial troughs surrounded on both sides by extensive flanking regions (Lizarralde et al., 2007). In contrast to mid-ocean spreading centers, axial troughs and flanking regions of Guaymas Basin are covered by thick, organic-rich sediments that represent a combination of terrigenous input and biogenic sedimentation from the highly productive water column (Calvert, 1966). Magmatic intrusions, or sills, are embedded within the thick sediment layers, where they drive hydrothermal circulation (Lonsdale and Becker, 1985) and thermally alter buried organic matter (Seewald et al., 1990), in the process generating complex petroleum compounds (Didyk and Simoneit, 1989), light hydrocarbons and methane (Welhan and Lupton, 1987), carboxylic acids (Martens, 1990), and ammonia (Von Damm et al., 1985). Since the sediments act as a heat-retaining thermal blanket, magmatic activity and organic matter alteration and mobilization are not limited to the spreading center but also occur at considerable distance, up to 50 km off-axis (Lizarralde et al., 2010). Many of these off-axis sites resemble cold seeps, where methane advection is linked to pathways formed by deeply buried magmatic sills (Geilert et al., 2018). If the underlying sill is sufficiently shallow and hot, the hydrothermal underpinnings of these off-axis sites becomes visible; the recently described Ringvent site provides an example (Teske et al., 2019).

In contrast, the Sonora Margin harbors classical cold seeps where sediment compaction drives reducing, methane-rich seep fluids to the surface. Numerous seep sites with carbonate outcrops and cold seep fauna have been observed on an eroding pressure ridge that follows the transform fault at the base of the Sonora Margin (Simoneit et al., 1990;Paull et al., 2007); the seep communities at these sites are largely based on methanotrophy and sulfide oxidation (Portail et al., 2015). Seep communities and sulfide-oxidizing microbial mats are also widespread on the Sonora Margin slopes (Vigneron et al., 2014;Cruaud et al., 2017).

Finally, most of the extensive flanking regions of Guaymas Basin and the Sonora Margin slope are covered by sediments without particular seep or hydrothermal influence; these sediments consist of mixed terrigenous runoff and biogenic components, dominated by diatoms (Calvert, 1966). Sediments on the upper Sonora Margin underlying the oxygen minimum at ca. 600 to 800 m depth, lack bioturbation and show finely laminated, seasonally changing sedimentation patterns of spring diatom blooms and terrestrial runoff during late summer trains (Calvert, 1964).

Here, we survey the distribution of archaea in diverse sedimentary environments located in the greater Northern Guaymas Basin and Sonora Margin regions. Sampling areas include background sediments from the Guaymas Basin flanking regions, Sonora Margin sediment within the oxygen minimum zone, reducing sediment with cold seep characteristics from the base of the Sonora Margin, and sediment from the off-axis Ringvent site where hydrothermal circulation and methane seepage is driven by a gradually cooling, buried shallow sill. We expand a previous limited sequencing survey of these sediments focused on just one of these sites (Teske et al., 2019) by: i) extending the geochemical analyses, ii) increasing the sampling resolution used for molecular sequencing (from one or two samples per site to ~1-meter intervals for all sites), iii) providing a wide breadth of comparative ecological analyses, and iv) discussing the potential implications of our results at a basin-wide scale.

Results

1. Sediment and porewater geochemistry

We surveyed archaeal distribution at six sites on the northwestern and southeastern off-axis regions of Guaymas Basin and on the Upper Sonora Margin (Figure 1, Table S1). These locations represent four different environmental settings: (i) Sediments on the Guaymas Basin flanking regions without hydrothermal or seep activity, represented by cores Cont03, Cont10 and Cont13; (ii) the oxygen minimum zone on the upper Sonora Margin (Calvert, 1964), represented by core OMZP12, (iii) compacted, highly reducing seep sediments from a pressure ridge, running along the transform fault that is cutting across the base of the Sonora Margin (Simoneit et al. 1990, Paull et al. 2007), represented by core Seep06, and (iv) the Ringvent site, characterized by off-axis hydrothermal circulation (Teske et al., 2019), represented by core RNVP11 (Figure 1). At each site, sediment piston cores ranging from 5 to 486 cm below the seafloor (cmbsf) were collected and geochemically characterized (Figure 2). The sediments are geologically young, ranging in age between ~0.05K and ~20K calendar years, as determined by C14 dating (Teske et al., 2019). The different cores show distinct geochemical characteristics.

Core SeepP06 contains sulfide in millimolar concentrations throughout the core. Below the zone of sulfate-dependent methane oxidation at 1 m depth, methane accumulated to the highest concentrations of this survey, > 10 mM. Porewater DIC concentrations were consistently high and increased from 15 mM near the interface to nearly 50 mM with depth. These methane and DIC concentrations reached and in part exceeded the highest concentrations previously measured in Sonora Margin seep fluids (Paull et al., 2007). Sediments of core

SeepP06 yielded only approximately half of the porewater volumes of other cores, indicating porewater loss by pressure-induced compaction. Thus, core SeepP06 represents sulfidic, methane- and DIC-soaked, compacted seep sediments from the pressure ridge aligned with the transform fault at the lower Sonora Margin (Simoneit et al., 1990;Paull et al., 2007).

Cores Cont03 and ContP10 share similar methane, sulfate, sulfide, and DIC profiles indicative of non-reducing conditions where sulfate-reducing and methanogenic activities remain minimal and biogenic sulfide and methane occur only in micromolar trace concentrations. With TOC between 4 and 6 wt %, the sediments of core ContP10 are organic rich and represent the hemipelagic seafloor sediments of Guaymas Basin that receive ample biogenic sedimentation, mostly by diatoms (Calvert, 1966). The $\delta^{13}C$ values ranging from -20.12 to -20.51‰ are consistent with sedimentary organic material resulting predominantly from phytoplankton input (Teske et al., 2002). Slowly increasing $\delta^{15}N$ values ranging from 9.04 to 10.17‰, and gradually increasing C:N ratios downcore are consistent with microbial utilization of nitrogen compounds in sedimentary biomass.

Contrasting with nearby core Cont10, Core RNVP11 shows the biogeochemical signatures of seawater inmixing and previous hydrothermal alteration at Ringvent (Teske et al., 2019). Subsurface-derived porewater methane in high concentrations of 1 to 1.5 mM coexists with porewater sulfate near seawater levels; sulfide is largely absent and reaches 10 to 100 µM only below 3 m depth. Core RNVP11 also stands out by having the lowest DIC concentrations of all cores, approaching seawater DIC in the upper layers. Below 1 mbsf, organic carbon $\delta^{13}C$ values are the lowest for all cores, whereas C/N ratios are the highest, suggesting the influence of isotopically light and nitrogen-depleted fossil carbon sources (Figure 2). In contrast to other cores, Core RNVP11 shows a strong gradient of dissolved silica, increasing from 0.5 to 0.6 mM at the surface (similar to ContP10) towards 0.9 to 1 mM at the bottom (Figure 2 I). Silica dissolution is considered a marker of hydrothermal activity (Peter and Scott, 1988), and leads to elevated concentrations of dissolved silica in the water column of Guaymas Basin (Campbell and Gieskes, 1984).

Core OMZP12 differs from all other cores by its location in the oxygen minimum zone on the Sonora Margin slope (Calvert, 1964). Bottom water anoxia allows millimolar concentrations of porewater sulfide to permeate the entire sediment core including the surface, otherwise only seen in core SeepP06. The sulfate-methane transition zone occurs at approximately 1 and 2m depth for SeepP06 and OMZP12, respectively. Similar to core SeepP06, porewater DIC increases rapidly with depth, with a maximum value of 33mM at 254

cmbsf. Sediment TOC, $\delta^{13}C$, $\delta^{15}N$, and C:N ratio values generally resemble those of other cores in this survey.

Core ContP13, collected on the southeastern flanking region, differs from other cores by terrestrial input from the Yaqui River. Methane, sulfate, sulfide and DIC concentrations for this core follow similar depth profiles as observed for cores ContP03 and ContP10. However, TOC varies between ~1 and ~5 wt % in the first meter of sediment and between ~3 and ~6 wt % below, suggesting sedimentation pulses of varying organic carbon load. Sediment organic matter $\delta^{13}C$, $\delta^{15}N$ and C:N ratios fall within the range of values observed for other cores in this survey.

2. Diversity of the Guaymas Basin archaeome

Rarefaction curves are plotted separately for samples in approx. 1 meter depth intervals to examine potential downcore trends (Figure 3). Starting at 3 m depth, observed species richness based on rarefaction summaries are lower in Ringvent (Core RNVP11) sediment compared to other sediments (Figure 3, Table S2). Substantially more sequence reads, and thus a higher number of observed species, were recovered from the Sonora Margin OMZ sediment (Core OMZP12) relative to the other surveyed sites, below 1 m depth (Figure 3 C-E). To account for different sequencing depths without resorting to rarefying the dataset (McMurdie and Holmes, 2014), we estimated total diversity using a non-linear regression model for ratios of consecutive frequency counts, a state-of-the-art method addressing the issue of heterologous sequencing depths affecting richness estimates (Willis and Bunge, 2015). Results from this model indicate no statistically significant ($P_{val} < 0.05$) differences amongst surveyed sites (Figure S1).

When beta diversity of the archaeal communities was examined for correlations with environmental metadata using two-dimensional principal coordinate analysis, distinct clustering patterns are observed (Figure 4A). Surface communities for all cores except OMZP12 tightly cluster along negative axis 1 values. All OMZP12 samples cluster along axis 1 values greater than 0.01 independently of sediment depth. Separation along axis 2 partitions the SeepP6, RNVP11, and OMZP12 cores (with positive axis 2 values), from subsurface samples of control cores ContP3, ContP10, and ContP13 with negative axis 2 values; the surface samples of these cores cluster separately (Figure 4A). The influence of environmental factors (*i.e.*: methane, sulfate, sulfide, DIC, water depth, and sediment depth) on community ordination is complex (Figure 4B-G), and it appears likely that clustering patterns are not

driven by these environmental parameters alone. Water column depth (Figure 4F) appears to drive core OMZP12 clustering along larger positive values for axis 1, but most likely represents a proxy for the influence of the oxygen minimum zone at this depth.

3. Network analysis

Network analysis based on the co-occurrence of all OTUs in each sample reveals that the deepest communities recovered from Core RNVP11 exhibit the greatest degree of separation (Figure 5). At a maximum ecological (Bray-Curtis) distance of 0.8 (*i.e.*: the maximum distance allowed between two samples to be considered connected in the graphical model), most samples share taxa co-occurrence patterns, except for the deepest communities from core RNVP11 (Figure 5A). Decreasing the minimum ecological distance in the model to 0.5, resolves three independent network clusters (Figure 5B). Here, the two deepest samples from core RNVP11 share similar taxa co-occurrence patterns only with each other and are excluded from the two emergent additional networks. In one of these networks, communities near the seawater interface of all cores, with the exception of core RNVP11, connect at no more than 1-degree of separation. Interface sample SeepP06 5cmbsf connects the near-interface sample cluster to all core SeepP6 subseafloor (depth > 1mbsf) samples. A third independent network shows non-random taxa co-occurrence amongst subseafloor control sediments (cores ContP3, ContP10, ContP13), a subseafloor and a near-interface sample from core RNVP11, and all samples from core OMZP12; the three deepest OMZP12 samples are only peripherally connected (Figure 5B).

4. Community composition

Class-level community descriptions (SILVA132) assigned large membership fractions of the archaeal communities to the Bathyarchaeia, Hadesarchaeaeota, and Thermoplasmata (Figure 6). The Class Methanomicrobia, comprising methane-producing and methane-oxidizing members of the Euryarchaeota, was detected in multiple cores, but appeared most frequently at depth in core SeepP06. An in-depth summary of the Methanomicrobia reveals the presence of methanogenic families (eg: Methanomicrobiaceae) and anaerobic, methane-oxidizing ANME lineages (ANME-1, various ANME-2). Notably, ANME-1 archaea dominate core SeepP06 sequence assignments comprising nearly 40% of the total community at 394 cmbsf in this core (Figure S2, Table S3). Order- or higher-level community taxonomic descriptions for all samples generally contained 60% or greater unclassified community fractions (data not shown) when automated taxonomic assignments were performed. In order to not rely on the

uncertain output of taxonomy pipelines, and to resolve archaeal taxonomy assignments in a manner that is consistent with broadly accepted usage (Spang et al., 2017), we also describe community composition based on phylogenetic placement of dominant sequence variants for the most numerous 25 Operational Taxonomic Units (OTUs).

The majority of high-quality sequences in this study (73.0%) clustered into 25 OTU lineages (Figure 7A). Archaeal communities were largely dominated by OTU lineages related to the Bathyarchaea (16 out of the top 25 OTUs). OTUs 01 to 03, the three most abundant lineages, belong to the MCG-1, MCG-2, and MCG-3 Bathyarchaea subgroups (Kubo et al., 2012) respectively, with close relatives recovered from Guaymas Basin and globally-dispersed subseafloor habitats (Figure 7B). High abundance lineages related to the Marine Benthic Group-D within the Thermoplasmata (MBG-D; OTUs 04, 10, 11, 21, and 23) and the Terrestrial Miscellaneous Euryarchaea Group (TMEG; OTU 08) were recovered from all cores and core depths except the subsurface of Ringvent (Core RNVP11, depth > 1mbsf). Two highly abundant lineages represented by OTUs 05 and 18 (Figure 7C) were recovered from every core except core SeepP06, and identified as relatives of the Hadesarchaea, formerly known as South African Gold Mine Euryarchaea Group (SAGMEG)-1 (Baker et al., 2016). Lastly, OTU 14 clustered within anaerobic ANME-1 methanotrophs and was most closely related to ANME-1 phylotypes from cold seep, hydrate and brine habitats; OTU 14 did not affiliate with the thermophilic ANME-1Guaymas lineage recovered from hydrothermally active, hot sediments in Guaymas Basin (Biddle et al., 2012) (Figure 7C).

5. Differential taxon abundance estimations across ecological niches

Differential abundance analyses (Wald Test, Pval = 0.01) were performed on various ecological-models following potential environmental niches suggested by ordination patterns (Figure 8). Only OTU lineages amongst comparison groups containing more than 100 sequences were used for each test. We tested the influence of sediment depth in the absence of seepage or hydrothermal influence, the impact of the oxygen minimum zone waters on surficial and subsurface sediments, and the effect of hydrothermal disturbance. These analyses have to be qualified by the fact that they are based on patterns of sequence frequencies, which are derived from the archaeal community but do not necessarily represent it in identical proportions.

To test the estimated differential abundance of archaeal community members in near-surface (depth < 1mbsf) and subsurface (depth > 1mbsf) communities under conditions of normal hemipelagic sedimentation, we selected cores ContP3, ContP10, and ContP13; these

cores lack hydrothermal, seepage or OMZ influence and, therefore, show the archaeal community of organic-rich Guaymas Basin sediments in the absence of these selective factors. Here, over 43% of OTUs (n > 100 seqs) are differentially enriched with depth (Figure 8A & B). Most differentially abundant taxa are estimated to be significantly less abundant in near-surface relative to subsurface control sediment (Figure 8B). Subsurface enriched lineages include members of various Bathyarchaeal groups (MCG-1, MCG-3, and MCG-6), a TMEG lineage (OTU 08) closely related to clones previously recovered from Guaymas Basin, and a Hadesarchaea lineage (OTU 18) whose closest relatives are clones from deep Mediterranean waters (Figures 7B & C). Among the top 25 OTUs in this study, the only highly abundant lineage that is significantly (Pval = 0.01) enriched in near-surface relative to subsurface control sediment (Figure 8A), is OTU 07, related to Guaymas Basin and Indian estuary sediment MCG-6 clones (Figure 7B). The same lineage was also found to occur preferentially in surficial estuarine sediments in the White Oak River, while avoiding the sulfide-rich, sulfate-reducing and methanogenic conditions just a few centimeters downcore (Lazar et al., 2015).

When archaeal abundance in all oxygen minimum core OMZP12 samples was checked against the shallow sediment samples of the control cores (ContP03, ContP10, ContP13, see Figure 8C and D), over 50% of archaeal OTUs (n > 100 seqs) present in surficial controls were differentially enriched in core OMZP12. Of the 25 most highly abundant OTUs in this study, those that were differentially abundant in this comparison (14 OTUs) were predominantly enriched in core OMZP12 (eleven of fourteen high abundance OTUs, Figure 8D). Only OTUs 07, 09, and 21, representing the Bathyarchaeal groups MGC-6 and MCG-2, and a MBG-D lineage, respectively, were enriched in surficial control sediment (Figure 8D).

We also tested for differentially abundant taxa between all core OMZP12 samples and subsurface control sediment (Figure 8E). Here, only 31.1% of OTUs (n > 100 seqs) were differentially enriched; of the 25 most abundant OTUs in this study only 4 were differentially enriched. Of these four differentially enriched archaeal OTUs, OTUs 15 and 20, Bathyarchaeal lineages in the MCG-1 and MCG-2, respectively, were enriched in subsurface control sediment. OTU 06 within the MCG-2 group, and OTU 19, an MCG lineage tenuously related (bootstrap value < 70%) to MCG-6, were enriched in core OMZP12. Overall, archaeal types occurring in anoxic subsurface sediments of core OMZP12 resemble those in other subsurface cores to a large extent.

The hydrothermally influenced Ringvent core RNVP11 was compared against control core ContP10, located only 1.6 km further west, at the same depth and local sedimentation regime (Figure 8F-H). In surficial sediment (> 1 mbsf) this comparison revealed only one

differentially enriched phylotype, OTU 18 within the Hadesarchaea (Figure 8G). On the other hand, the comparison of subsurface (>1 mbsf) communities identified 43 OTUs, or 61% of all shared OTUs (n > 100 seqs), as differentially abundant between these two sites (Figure 8H). Focusing on the top 25 highest abundance OTUs in this study, eleven OTUs were differentially enriched. Eight taxa, comprised of Bathyarchaeal, TMEG, MBG-D, and ANME-1 representatives, were significantly enriched in subsurface control sediment relative to Ringvent subsurface sediment (OTUs: 07, 08, 10, 11, 14, 15, 16, and 23). The remaining three taxa (OTUs 06, 12, and 25), enriched in the Ringvent subsurface, were classified as Bathyarchaea related to the MCG-1 and MCG-2 subgroups.

Core SeepP06, from compacted seep sediments, was checked against both shallow and subsurface sediment samples of the control cores (ContP03, ContP10, ContP13, Figure 7I-K). 41% of archaeal OTUs (n > 100 seqs) present in shallow controls and seep sediment were differentially enriched (Figure 8J). Of the 25 most highly abundant OTUs, most were predominantly enriched in seep sediment and included lineages classified as ANME-1, MBG-D, and TMEG (Figure 8J). When comparing differentially abundant taxa between core SeepP06 and subsurface control sediment, almost 53% of archaeal OTUs (n > 100 seqs) were differentially enriched (Figure 8K). Most of the OTUs in this comparison, including the most abundant OTUs in this study (OTUs 01-03), were enriched in control subseafloor sediments rather than in core SeepP06 sediment (Figure 8K).

Discussion

1. Complex determinants of archaeal ecosystem structure

Overall, complex physical and geochemical factors structure sedimentary habitats and depth-related niches for archaea in Guaymas Basin. Archaeal community ordination patterns reveal niche differentiation and some unexpected clustering patterns among the different sites (Figure 4A).

Most notably, surficial communities of background control sediments ContP03, ContP10, and ContP13, at 5 to 10 cm depth, cluster away from their respective subsurface communities near 1 m depth and below (Figure 4A), implying an ecological bifurcation into two distinct niches. Taxa co-occurrence network patterns support this differentiation between shallow and subseafloor control sediment sites (Figure 5B). Possibly, the availability of electron acceptors such as oxygen, nitrate or oxidized metals drives the depth-dependent niche separation observed in background control sediment sites (ContP03, ContP10, and ContP13). Surface archaeal communities continually change as sediment layers accumulate; given high

sedimentation rates of 0.23 to 1 mm/y at Guaymas Basin (Teske et al., 2019), it takes approx. 1000 to 4000 years for background surface communities to transition to subsurface communities at 1 m depth. Interestingly, shallow vs subsurface differentiation is less apparent in seep, OMZ or hydrothermally influenced sites (Figure 5B); parameters other than sediment depth or surficial redox regime are shaping archaeal community composition in seepage- or hydrothermally-influenced habitats, compared to the control sites.

The cores SeepP06 and RNVP11 represent different geochemical regimes (compaction-induced continental margin seepage vs hydrothermal circulation, respectively), yet these two sites cluster tightly in ordination space (Figure 4). Individual geochemical factors, for example, the sulfidic, methane- and DIC-rich conditions in SeepP06 would have indicated that OMZP12 should be its closest equivalent (Figure 2). The unexpected clustering of SeepP06 and RNVP11 suggests that factors beyond current geochemical conditions, for example recent environmental disturbance, can influence archaeal community structure. At Ringvent (RNVP11), sedimentary community diversity may have been reduced during prior episodes of thermal purging or high methane flux (Teske et al., 2019), selecting for a resilient, yet potentially less diverse (Figure 3), "survivor" community.

Lastly, community ordination differentiates OMZ sediment from all other sedimentary habitats (Figure 4). While water depth appears to have a strong influence on OMZ sediment ordination (Figure 4F), we propose that differences in redox potential at the sediment interface due to its direct contact with oxygen-depleted water (Calvert, 1964) rather than water column depth, is the key environmental constraint driving the ordination patterns of OMZ sediment.

2. A "forest view" of Archaea in Guaymas Basin sediments

Archaea observed in this sedimentary habitat survey belong to the Bathyarchaea, the MBG-D and TMEG lineages within the Thermoplasmatales, the Hadesarchaea (SAGMEG), and ANME-1 lineages, as shown previously in a sequencing survey using different archaeal primers (Teske et al. 2019). The uncultured Bathyarchaea and MBG-D archaea pronouncedly dominate the dataset, and cold anoxic marine sediments globally (Kubo et al., 2012;Lloyd et al., 2013). Bathyarchaea play an important role, tantamount to that of the domain Bacteria, in the remineralization of complex organic matter in marine sediment (Lloyd et al., 2013); some of their members (MCG-8 lineage) use lignin, the second most common biopolymer on Earth, as an energy source (Yu et al., 2018). Since Bathyarchaeota harbor the Woods-Ljungdahl pathway, they are implied in acetogenic subsurface metabolism (He et al., 2016), but it remains open whether acetogenic pathways are used for net autotrophy, or derive their substrates from

organic carbon sources (Lever et al., 2010); similar considerations apply to the metabolically versatile MBG-D archaea (Zhou et al., 2019). Some Bathyarchaea harbor genes of the MCR complex, suggesting methylotrophic methanogenic activity and, perhaps, syntrophic interactions with sulfate reducing bacteria leading to the anaerobic oxidation of methane (Evans et al., 2015). However, the dominant Bathyarchaea OTUs recovered in this study (MCGs 1, 2, 3, 6, and 13, Figure 7B) are only distant relatives of methane-cycling marine Bathyarchaea, which fall into MCGs 15 and 16 (Evans et al., 2015). Hadesarchaea, originally described as the South-African Gold Mine Miscellaneous Euryarchaeal Group (SAGMEG), are metabolically versatile anaerobic heterotrophs with the metabolic potential for CO and H_2 oxidation coupled with nitrite reduction to ammonia and are found in environments across broad (4 to 80°C) thermal ranges (Baker et al., 2016). One of two frequently recovered Hadesarchaea lineages (OTU 05) is conspicuously enriched in subsurface sediment in Ringvent (RNVP11), where low observed sequence richness (Figure 3) coincides with evidence (high silica porewater concentrations at depth indicative of a thermal dissolution of sedimentary diatoms, Figure 2I) for a thermal purge in the past (Teske et al., 2019).

Unsurprisingly, the methane-cycling Methanomicrobia are rare in background sediments (ContP03, Cont10, and Cont13) but are strongly enriched in core SeepP6 and, to a much lower extent (slightly over 2%), in core OMZP12 at 204 cm depth (Figures 6 and S2). Interestingly, a single ANME-1 OTU lineage, OTU14, is highly enriched in SeepP06 sediments (Figure 7A and C). This lineage is closely related to ANMEs recovered from cold, anoxic habitats, such as seafloor seep sediments, methane hydrates, and hypersaline anoxic basins, and distinct from previously described ANME-1 phylotypes from Sonora Margin cold seep sediments and potentially thermophilic ANME-1 phylotypes from Guaymas Basin hydrothermal sediments (Holler et al., 2011;Biddle et al., 2012). Although ANME-1 archaea were generally assumed to be obligate methanotrophs, this assumption has been challenged and this lineage has been proposed as potentially methanogenic, based on its occurrence and activity in sulfate-depleted sediments (Lloyd et al., 2011;Kevorkian et al., 2020); thus, the biogeochemical role of these archaea would be modulated by the presence or absence of sulfate, or concomitant changes in electron donors. ANME-2 and cultured methanogenic lineages were observed in low percent abundances in all cores in this study (Figure S2). Interestingly, ANME-2 lineages were extremely rare, representing less than 0.05% of any sample and less than 0.02% of any SeepP6 community (Figure S2, Table S3). The prevalence of ANME-1 over ANME-2 in the El Puma cores is consistent with the ecophysiological preference of ANME-1 archaea

for reducing, sulfidic subsurface sediments, and the preference of ANME-2 for near-surface sediments with intermittently oxidizing conditions (Rossel et al., 2011;Ruff et al., 2015). Previous surveys of mat-covered seep sediments on the Sonora Margin have revealed transitions from ANME-2 towards ANME-1 within short push cores of max. 17 cm depth (Vigneron et al., 2013).

Overall, we hypothesize that the Archaeome in the sedimented flanking regions of Guaymas Basin is generally fueled by heterotrophic processes including the degradation of proteins, polymeric carbohydrates (Ziervogel and Arnosti, 2020), and accumulating lipids (Teske et al., 2002) resulting from high sedimentation rates driven by high levels of primary production in the water column. Diverse niche communities allow the Guaymas Archaeome to adapt to environmental challenges, such as hydrothermalism or methane seepage, that are common in the greater Guaymas Basin area; thus, the site-specific complexity of the Guaymas Archaeome underpins its resilience.

3. Ecological comparisons: Differentially abundant taxa across sedimentary habitats

3.1 Near-surface vs. subsurface sediment niches

When comparing the surficial to the subsurface archaeal populations in background control sediments, the majority of OTUs estimated to be differentially abundant are significantly more enriched in the subsurface relative to the surficial sediment, suggesting that benthic archaea prefer subsurface conditions (Figure 8A, B). This trend may also reflect the impact of electron acceptors; for example, oxygen permeates background sediments in Guaymas Basin for at least one centimeter [(Teske et al., 2016), Figure 8B therein]. Following a recently proposed model for benthic microbial communities (Starnawski et al., 2017) the archaeal community at the oxic water-sediment interface likely undergoes downcore selection, based on site-specific selective pressure, resulting in reduced diversity with depth but a higher prevalence of subsurface-adapted taxa within a few thousand years after burial in anoxic subseafloor sediment. Benthic archaea, predominantly Bathyarchaea and MBG-D lineages, survive on residual carbon sources that remain after burial and microbial degradation in surficial sediments (Lloyd et al., 2013). Interestingly, catabolic activity and electron donor diversity, rather than terminal electron acceptor type or burial time, appear to drive bacterial OTU richness in anoxic subseafloor sediment (Walsh et al., 2016). This niche construction mechanism, driven by the biotic microenvironment as opposed to abiotic environmental filtering (Aguilar-Trigueros et al., 2017), is potentially widespread across the large habitable volume represented by non-hydrothermally influenced subseafloor sediments in Guaymas Basin.

3.2 OMZ vs. control sediment

When comparing OMZ and surficial background control sedimentary communities, 50% of high abundance OTUs found across both habitats are estimated to be differentially enriched (Figure 8D). Two thirds of the differentially enriched taxa were enriched in the OMZ rather than the surficial background sediments. The MCG lineages MGC-6 and MCG-2 (OTUs 9 and 7) and a MBG-D phylotype (OTU 21) are enriched in the surficial background controls relative to the OMZ sediment (Figure 8D). Interestingly, MCG-6 members bear hydrolases that specifically target plant-derived polymeric carbohydrates (Lazar et al., 2016), a potential trait-environment relationship that may differentiate surficial background control from OMZ sediment communities (Figure 8D). When comparing subsurface background control and OMZ sediment communities, the number of OTUs estimated to be differentially enriched was about equal across both environments; however, the majority (68.9%) of high abundance OTUs show no significant differences in their estimated abundances at all (Figure 8E). This implies that the subsurface, rather than surficial, background control communities are more similar to the OMZ communities, a point also corroborated by taxa co-occurrence network analysis (Figure 5B). Thus, oxygen depletion in background subsurface sediment, and oxygen depletion through the overlying oxygen minimum zone of the water column (Calvert, 1964), result in some convergence between archaeal communities across geographically distant and environmentally distinct sedimentary habitats.

3.3 Ringvent vs. control sediment

The surficial archaeal communities of Ringvent (RNVP11) and its nearby control site (ContP10) are similar to each other, as indicated by extensive co-occurrence networks (Figure 5) and by the lack of differential enrichment between the two cores (Figure 8G). A member of the Hadesarchaea, OTU18, is estimated to be significantly enriched in Ringvent surficial sediment relative to the control; otherwise differences in taxon abundance across these habitats are negligible. Both sites are only 1.6 km apart and therefore most likely share recent depositional histories and microbial inoculum sources, which validates core ContP10 as a site-specific control for assessing the environmental determinants structuring subsurface archaeal communities at Ringvent. The reduction in sequence recovery and, potentially, archaeal community richness in subsurface Ringvent (RNVP11) sediment (Figure 3) is attributed to environmental selection via hydrothermal purging or methane seepage driven by recent sill emplacement that continues to drive hydrothermal circulation, selecting against microbes

458 unable to withstand chemical or thermal changes associated with hydrothermal circulation
459 (Teske et al., 2019). Thus, OTUs may be enriched in Ringvent subsurface sediment relative to
460 its nearby control site (Figure 8H) via two possible ecological scenarios; i) surviving resilient
461 microbes could dominate the habitat after their competitors have been removed, and ii) new
462 arrivals after the disturbance could efficiently recolonize the depopulated surface sediment.
463
464 *3.4 Seep vs. control sediment*
465 Differential abundance comparisons show that the ANME-1, MBG-D, and TMEG lineages are
466 significantly enriched in the seep sediments, compared to controls (Figure 8I-K). Generally,
467 methane seeps are specialized microbial benthic habitats where methanotrophic archaea
468 (ANME) and syntrophic Deltaproteobacteria oxidize methane anaerobically exploiting sulfate
469 as an electron acceptor (Lloyd et al., 2010;Ruff et al., 2015). The dominance of these inter-
470 domain syntrophic partners distinguishes seafloor seep habitats (Ruff et al. 2015). Archaeal
471 community structure in SeepP06 sediments differs little with depth; it is most similar, in terms
472 of taxa overlap, to other samples from the same core (Figures 4 & 5). Therefore, the influence
473 of cold seepage drives community selection to a greater degree than the environmental factors
474 associated with depth-dependent niche differentiation observed in background control
475 sediment.
476 Comparison with other Sonora Margin cores highlights the seep characteristics of core
477 SeepP06. Based on the presence or absence of major archaeal lineages, SeepP06 archaeal
478 communities are similar to surficial (<1mbsf) communities from Sonora Margin cold seeps,
479 predominantly comprised of Thermoplasmata (MBG-D), Bathyarchaea, and ANME lineages
480 (Cruaud et al., 2017). The SeepP06 archaeal communities share dominant archaeal lineages –
481 the Thermoplasmatales (MBG-D), Lokiarchaeota and Bathyarchaeota - with Sonora Margin
482 subsurface sediments [core BCK1, (Vigneron et al., 2014)]. Interestingly, the high proportion
483 of ANME-1 archaea in SeepP06 is not shared by the Sonora Margin subsurface core (Vigneron
484 et al., 2014). The Sonora Margin subsurface sediment core has a deeper methane/sulfate
485 interface than Seep06, ca. 4-5 m instead of 1 m, and contains little sulfide above 5 m depth,
486 indicating strongly attenuated seep influence in core BCK1 compared to SeepP06.
487
488 *4. Core-specific features of the benthic Archaeome*
489 Controls that structure microbial communities in hydrothermal sediments of Guaymas Basin
490 have been studied extensively; for example, extreme temperature and porewater gradients
491 shape microbial population structure, genomic repertoire and activities within a few

492 centimeters depth beneath the seafloor (McKay et al., 2012;McKay et al., 2016;Dombrowski
493 et al., 2018). However, ecological factors influencing microbial life in other sedimentary
494 habitats at Guaymas Basin are comparatively unconstrained. By comparing archaeal
495 communities in diverse sedimentary habitats to background controls representative of standard
496 hemipelagic sedimentation, characteristic responses of the archaeal communities to these
497 distinct environmental settings are becoming apparent. Compaction-induced seepage near the
498 base of the Sonora Margin, and the resulting methane- and sulfide-rich porewater conditions
499 in core SeepP06, selected for anaerobic methane-oxidizing archaea (ANME-1) and MBG-D
500 archaea within the Thermoplasmata, and reduced the relative proportion of Hadesarchaea and
501 Bathyarchaeota. Prior disturbances by hydrothermal impact or strong methane seepage,
502 exemplified in the Ringvent sediments (RNVP11), also strongly differentiated sedimentary
503 archaeal communities from those in background controls. Observed community richness in
504 RNVP11 based on rarefaction curves are reduced throughout much of the core; these results
505 resembled the outcome of a parallel study using different archaeal primers, and bacterial
506 primers as well (Teske et al., 2019). Lastly, anoxic bottom waters impinging on the sediment
507 on the upper Sonora Margin (OMZP12) drive similarities between anoxic surficial sediment at
508 this site and anoxic subsurface background control sediments. The anoxic redox state of the
509 water-sediment interface may also enhance archaeal richness estimates in the upper sediment
510 column, potentially by facilitating the pelagic-benthic transition of archaea, or selecting against
511 a bacteria-dominated interface (Xia et al., 2017). In brief, the archaeal communities of different
512 cores respond in different ways to specific local controls.
513
514 *5. Environmental history determines ecological context*
515 The sediment cores shared similar biogeochemical parameters, such as sedimentary TOC, and
516 organic matter $\delta^{13}C$, $\delta^{15}N$ and C:N ratios. Repeatedly, studies of uncultured microbes in the
517 sedimentary subsurface tried to correlate community composition with a wide range of
518 biogeochemical or thermal parameters, in the hope that these linkages provide insights into
519 habitat preference and ecophysiology of uncultured archaea (Durbin and Teske, 2012;Lazar et
520 al., 2015;McKay et al., 2016). While this strategy can yield valuable results, we caution that
521 patterns of archaeal community composition are not deterministically linked to biogeochemical
522 parameters alone, rather, the full context of an ecological interpretation requires that biological
523 and geochemical observations are integrated with the environmental setting and history of a
524 site. For example, the lighter $\delta^{13}C$ values of sedimentary organic matter in RNVP11 (trending

towards -22 ‰ compared to most values clustering between 20 and 21 ‰), the slightly elevated C:N ratios at this site, increased Si concentrations at depth, or the elevated methane content superimposed on seawater-like porewater characteristics, are not in themselves critical factors that determine biological metrics in this core; these factors are significant because they reveal a depositional history of organic-rich sediments overprinted by relatively recent hydrothermalism and methane flux that has left its footprint on the present-day archaeal community. In another example, the archaeal communities of cores SeepP06 and OMZP12 would be assumed to be similar, since both sites are rich in sulfide, methane and DIC, and show sulfate depletion concomitant with methane accumulation. However, the distinct environmental settings and histories of these two cores, at the heavily compacted, seep-influenced base of the Sonora Margin (SeepP06), and under the oxygen minimum zone waters of the upper Sonora Margin (OMZP12), ultimately select for different archaeal communities.

Conclusion

In the greater Guaymas Basin and Sonora Margin area, complex geological and oceanographic processes impose environmental controls on different sedimentary habitats and their archaeal populations relative to background control sites. In background sediments, archaeal communities vary little with depth after the surface/subsurface transition; here, subsurface communities result primarily from long-term survival likely conferred by relatively reduced mortality (Kirkpatrick et al., 2019). In contrast, localized factors, including water column anoxia, methane seepage, and hydrothermal circulation, constrain the biodiversity and potential biogeochemical activity of sedimentary Archaea across our benthic survey in specific ways. Local sediment biogeochemistry has to be viewed in a broader context – within the history and evolution of a particular site – to reveal its influence on selective survival for certain lineages and subsequent shaping of the resident archaeal ecosystem.

Materials and Methods

Sample Collection. All samples were collected using piston coring during R/V *El Puma* (Universidad Nacional Autónoma de México, UNAM) Expedition Guaymas14 to the Gulf of California, October 14-27th, 2014. A 5-m long piston core (RNVP11) was obtained on Oct 21, 2014 from the central basin within the ring (27°N30.5090/111°W40.6860, 1749 m; core length 4.9 m), parallel to a control core (ContP10) approx. 1 mile to the west of Ringvent (27°N30.5193/111°W42.1722; 1731 m depth, 3.93 m core length) collected on the same day. Core SeepP06 was obtained on Oct. 19 from the lower Sonora Margin, near its boundary with

559 the Ridge flanks (27°N38.8367/111°W36.8595; 1681 m depth, 3.95 m core length). Core
560 OMZP12 was taken on Oct. 22 from the upper Sonora Margin (27°N52.1129/111°W41.5902,
561 667 m, 4 m core length) in the oxygen minimum zone as previously determined by water
562 column oxygen profiling (Calvert, 1964). Core ContP03 was collected on Oct. 17 from the
563 northwestern end of the ridge flanks (27°N37.6759/ 111°W52.5740; 1611 m depth, 3.27 m
564 core length. Core ContP13 was obtained on Oct. 22 from the southeastern ridge flank of
565 Guaymas Basin (27°N12.4470/111°W13.7735, 1859m depth, 3.31 m core length).
566
567 *Geochemical Analyses*. Porewater was obtained from freshly collected sediments on RV *El*
568 *Puma* by centrifuging ca. 40 ml sediment samples in 50 ml conical Falcon tubes for ca. 5 to 10
569 minutes, using a Centra CL-2 Tabletop centrifuge (Thermo Scientific) at approx. 1000*g*, until
570 the sediment had settled and produced ca. 8 to 10 ml of porewater. Porewater was extracted
571 from 5 cm thick sediment samples, which are designated by the top of each sample. For
572 example, a "95 cm" geochemistry sample extends from 95 to 100 cm below the sediment
573 surface. Sulfate, sulfate, methane and DIC porewater profiles for cores SeepP06 to OMZP12
574 were previously published (Teske et al., 2019), and are re-plotted here for comparison with
575 unpublished profiles from cores ContP03 and ContP13. Porewater analyses were performed as
576 previously described, using the colorimetric cline assay for sulfide, ion chromatography for
577 sulfate, and GC-IRMS for DIC and methane (Teske et al., 2019). Carbon and nitrogen isotopic
578 and elemental composition was determined at the Stable Isotope Laboratory (SIL) at the
579 University of California, Santa Cruz (UCSC). Bulk sediment $\delta^{15}N$ and elemental ratio data
580 were collected using 20mg samples in Sn capsules; organic $\delta^{13}C$ and elemental
581 composition data were collected using 2.5mg samples of acidified sediment in Sn capsules. All
582 samples were measured by Dumas combustion performed on a Carlo Erba 1108 elemental
583 analyzer coupled to a ThermoFinnigan Delt Plus XP isotope ratio mass spectrometer (EA-
584 IRMS). An in-house gelatin standard, Acetanilide, and an in-house bulk sediment standard,
585 "Monterey Bay Sediment Standard", were used in all runs. Reproducibility of an in-house
586 matrix-matched sediment standard is <0.1‰ VPDB for $\delta^{13}C$ and <0.2‰ AIR for $\delta^{15}N$. Data is
587 corrected for blank, and for drift when appropriate. Carbon and nitrogen elemental composition
588 was estimated based on standards of known composition, for which analytical precision is
589 determined to be better than 1 %. Filtered but unamended porewater samples, stored at 4°C,
590 were used for quantifying multiple stable ions, including silicate, by ion chromatography at
591 GEOMAR, Kiel, Germany (Hensen et al., 2007). All geochemical data in this study are
592 publicly available at the Biological and Chemical Oceanography Data Management Office

(BCO-DMO) under the following dataset IDs: 661750, 661658, 66175 and 661808 for Methane, DIC, Sulfate and Sulfide, respectively.

3. DNA extraction and gene sequencing

Samples for DNA sequencing [approx. 2 cm^3 each] were obtained by syringe coring at the indicated depth [in cm] below the sediment surface. DNA for all survey sites was extracted from ~0.5-1.0 cm^3 sediment sample volumes using the Powersoil DNA extraction kit according to the manufacturer's instructions (QIAGEN, Carlsbad, CA, USA). Archaeal 16S rRNA gene amplicons from DNA extracts were generated using the following primer set: A751F: 5'-CGA CGG TGA GRG RYG AA-3' and A1204R: 5'-TTM GGG GCA TRC NKA CCT-3'(Baker et al., 2003). Amplicons were sequenced on an Illumina MiSeq platform (Illumina, San Diego, CA, USA) at the Center for Biofilm Engineering in Bozeman, Montana. Sequencing run specifications are found in the Visualization and Analysis of Microbial Population Structures (VAMPSs) website (https://vamps.mbl.edu/resources/primers.php) (Huse et al., 2014).

4. Sequence Processing

Sequences were processed with *mothur* v.1.39.5 (Schloss et al., 2009) following the *mothur* Illumina MiSeq SOP (Kozich et al., 2013). Briefly, forward and reverse reads were merged into contigs and selected based on primer-specific amplicon length and the following parameters: maximum homopolymers of 6bp, and zero ambiguities. High quality sequences were aligned against the *mothur*-recreated Silva SEED v132 database (Yarza et al., 2010) and subsequently pre-clustered at 1% dissimilarity. As suggested elsewhere (Kozich et al., 2013), spurious sequences are mitigated by abundance ranking and merging with rare sequences based on minimum differences of three base pairs. Chimeras were detected and removed using UCHIME de novo mode (Edgar et al., 2011). Sequences were then clustered, by generating a distance matrix using the average neighbor method, into operational taxonomic units (OTUs, 97% similarity cutoff). OTU classification was performed on *mothur* using the SILVA v132 database as implemented using the classify.seqs command using the Wang algorithm (kmer assignment with 1/8 kmer replacement as bootstrap) and cutoff=80 (minimal bootstrap value for sequence taxonomy assignment). All sequence data are publically available at the following repository: NCBI under BioProject PRJNA553578 and accession numbers SRX6444849-SRX6444877.

5. Sequence Analyses

627 *5.1 Community Analyses and Visualizations*

628 Community analyses were performed in *RStudio* version 0.98.1091 (Racine, 2012),
629 implemented in R version 3.5.2, using the *vegan* (Oksanen et al., 2015) and *phyloseq*
630 (McMurdie and Holmes, 2013) R-packages. Sample richness analyses used the R package
631 *breakaway* (Willis et al. 2017) for inferring precision of diversity estimations given the
632 heterologous sequencing depth. Data were rlog normalized using *DESeq2* (Love et al., 2014)
633 prior to ordination using Bray-Curtis distances. An identical normalization strategy was used
634 on Bray-Curtis distances for co-occurrence network analysis performed using the
635 makenetwork() *phyloseq* command and visualized using the *igraph* R-package. DESeq2 was
636 also used to perform differential abundance analyses of taxa with low abundance taxa ($n < 100$
637 total reads per OTU) removed for the un-rarefied dataset, as suggested elsewhere (McMurdie
638 and Holmes, 2014).
639
640 *5.2 Phylogenetic Analyses*
641 Sequence alignments were performed using the high speed multiple sequence alignment
642 program MAFFT (Katoh and Standley, 2013) with the command: mafft --maxiterate 1000 –
643 localpair seqs.fasta > aligned.seqs.fasta. Maximum likelihood trees with 100 bootstrap support
644 were constructed using the RAxML (Stamatakis, 2014) program using the following
645 parameters: raxmlHPC -f a -m GTRGAMMA -p 12345 -x 12345 -# 100 -s aligned.seqs.fasta -
646 n T.tree, -T 4 ML search + bootstrapping. Newick trees files were uploaded to FigTree v1.4.2
647 for visualization.
648
649 **Data availability.** Geochemical data are available at the BCO-DCO under these reference links:
650 Porewater methane data: https://www.bco-dmo.org/dataset/661750/data
651 Porewater sulfate data**:** https://www.bco-dmo.org/dataset/661775/data
652 Porewater DIC data: https://www.bco-dmo.org/dataset/661658/data
653 Porewater sulfide**:** https://www.bco-dmo.org/dataset/661808/data
654
655 **Acknowledgements**
656 Sampling in Guaymas Basin was funded by NSF OCE grant 1449604 “Rapid Proposal:
657 Guaymas Basin site survey cruise for IODP proposal 833” to AT; NSF C-DEBI grant
658 “Characterizing subseafloor life and environments in Guaymas Basin” to AT and ACR. This
659 is C-DEBI Publication No. XXX. Thanks go to Zachary Stewart for helping with the

sedimentary organic matter geochemistry analyses. We thank the Science crew of R/V *El Puma* for excellent piston coring skills and a very enjoyable cruise.

Author Contributions

APT conceived the study. LJM, ACR, and APT collected samples. CM led the *El Puma* cruise and the piston coring effort. AB extracted DNA from sediments. LJM performed sequencing in the lab of MWF. ACR, LJM and CH performed biogeochemical and sedimentological analyses. GAR analyzed the sequence data with phylogeny input from APT. GAR and APT wrote the manuscript with input from all authors.

Figure Legends

Figure 1: Continental and bathymetric hybrid map depicting the location of Guaymas Basin and Sonora Margin in the Gulf of California, and relevant coring sites of the *El Puma* cruise. The bathymetry blue scale is annotated with 100-meter isobaths; the deepest areas in the axial valley range to just below 2000 meters.

Figure 2: Geochemical profiles for A) Methane, B) Sulfate, C) Sulfide, and D) DIC porewater concentrations; E) Organic Carbon content in weight %; F) Organic Matter $\delta^{13}C$ values, and G) $\delta^{15}N$ values; H) Carbon to Nitrogen ratios, and I) Silica porewater concentrations. Geochemical data for site ContP03 are not available for the analyses depicted in panels E-H.

Figure 3: Depth mapped rarefaction summaries (color coded to match surveyed sites) for complete high quality sequence dataset depicting richness as number of OTUs (97% similarity clustered) observed per sequences sampled. A) samples near the interface (0-10cmbsf), B) samples from ~100cmbsf, C) samples from ~200cmbsf, D) samples from ~300cmbsf, E) samples from depths greater than or approximately equal 400cmbsf.

Figure 4: A) Two-dimensional Principal Coordinate Analyses of Bray-Curtis dissimilarity distances from r-log normalized sequence count data. Each community plotted is color-coded to the core site and numerical labels indicate sediment depth (cmbsf). The first and second axes explain 29.8 and 14.8% of the variance, respectively. Environmental metadata superimposed on ordination plot are B) methane, C) sulfate, D) sulfide, and E) DIC concentrations, F) water depth, and G) sediment depth.

Figure 5. Network analysis based on the co-occurrence of all OTUs at in each sample. Nodes represent all archaeal communities analyzed in this study. Nodes are color-coded to match descriptions from Figure 1A. Edges are unweighted interactions depicting OTU co-occurrence meeting arbitrary thresholds. A) Co-occurrence network threshold set at a maximum Bray-Curtis distance of 0.8. B) Co-occurrence network threshold set at a maximum Bray-Curtis distance of 0.5.

Figure 6: Class-level community composition of all depths for all cores in this study. Core labels are color-coded to match the collection sites depicted in the bathymetric map in Figure 1.

Figure 7: A) Heatmap depicting percent abundance distribution for the 25 most abundant OTUs, representing 73.0% of all high-quality sequences in this study, for all cores and across all depths. Core labels are color-coded to match the collection sites depicted in the bathymetric map in Figure 1A. The phylogenetic association of each OTU lineage is depicted above each OTU header. The percent of total reads represented by the 25 most abundant OTUs in each community is shown in the column labeled "% of community". Maximum likelihood phylogenetic trees, with 100 bootstrap support, placing the top 25 most abundant OTUs within the following lineages: B) Bathyarchaea, C) the Euryarchaeotal lineages MBG-D, TMEG, SAGMEG, and ANMEs.

Figure 8: Differential Abundance Analyses bases on Wald's test (significance: alpha = 0.01). A) Ordination depicting archaeal community clustering for surface and subsurface samples of control sites ContP3, ContP10, and ContP13. B) Differentially abundant OTUs in near-surface versus subsurface communities from control sites. C) Ordination depicting community clustering in OMZP12, and surficial vs. subseafloor control sites. D) Differentially abundant OTUs in OMZP12 compared to surficial and E) subsurface communities from control sites. F) Ordination depicting community clustering in RNVP11 to Cont10. Differentially abundant OTUs in RNVP11 and Cont10 for G) surface samples and H) subsurface samples. I) Ordination depicting community clustering in SeepP06, and surficial and subseafloor control sites. Differentially abundant OTUs in SeepP06 samples compared to J) surficial and K) subsurface communities from control sites. Note: OTUs 13, 17 and 21 are color-coded as "unclassified"

by SILVA132, but the phylogeny identifies them as members of MCG-13 (OTUs 13 and 17) and MBG-D (OTU 21).

Works cited:

Aguilar-Trigueros, C.A., Rillig, M.C., and Ballhausen, M.-B. (2017). Environmental filtering is a relic. A response to Cadotte and Tucker. *Trends in Ecol. & Evol.* 32**,** 882-884. doi: 10.1016/j.tree.2017.09.013.

Baker, B.J., Saw, J.H., Lind, A.E., Lazar, C.S., Hinrichs, K.U., Teske, A.P., and Ettema, T.J. (2016). Genomic inference of the metabolism of cosmopolitan subsurface Archaea, Hadesarchaea. *Nat. Microbiol.* 1**,** 16002. doi: 10.1038/nmicrobiol.2016.2.

Baker, G.C., Smith, J.J., and Cowan, D.A. (2003). Review and re-analysis of domain-specific 16S primers. *J. of Microbiol. Meth.* 55**,** 541-555. doi: 10.1016/j.mimet.2003.08.009.

Biddle, J.F., Cardman, Z., Mendlovitz, H., Albert, D.B., Lloyd, K.G., Boetius, A., and Teske, A. (2012). Anaerobic oxidation of methane at different temperature regimes in Guaymas Basin hydrothermal sediments. *ISME J.* 6**,** 1018-1031. doi: 10.1038/ismej.2011.164.

Calvert, S.E. (1964). "Factors affecting distribution of laminated diatomaceous sediments in the Gulf of California" In Marine Geology of the Gulf of California, Vol. 3, eds. T.H. van Andel and G.G. *Shor, (Tulsa: American Association of Petroleum Geologists Memorir)* 3**,** 311-330.

Calvert, S.E. (1966). Origin of diatom-rich, varved sedimetns from the Gulf of California. *J. Geology* 74**,** 546-565. doi: https://www.jstor.org/stable/30059298.

Campbell, A.C., and Gieskes, J.M. (1984). Water column anomalies associated with hydrothermal activity in the Guaymas Basin, Gulf of California. *Earth and Planet. Sci. Lett.* 68**,** 57-72.

Cruaud, P., Vigneron, A., Pignet, P., Caprais, J.-C., Lesongeur, F., Toffin, L., Godfroy, A., and Cambon-Bonavita, M.-A. (2017). Comparative study of guaymas basin microbiomes: cold seeps vs. hydrothermal vents sediments. *Front. Mar. Sci.* 4. doi: 10.3389/fmars.2017.00417.

Didyk, B.M., and Simoneit, B.R.T. (1989). Hydrothermal oil of Guaymas Basin and implications for petroleum formation mechanisms. *Nature* 342**,** 65-69.

Dombrowski, N., Teske, A., and Baker, B.J. (2018). Exapnsive microbial metabolic cersatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. *Nature Comm.* 9. doi: https://doi.org/10.1038/s41467-018-07418-0.

Durbin, A.M., and Teske, A. (2012). Archaea in organic-lean and organic-rich marine subsurface sediments: an environmental gradient reflected in distinct phylogenetic lineages. *Front. Microbiol.* 3**,** 168. doi: 10.3389/fmicb.2012.00168.

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

Evans, P., Parks, D.H., Chadwick, G.L., Robbins, S.J., Orphan, V.J., Golding, S.D., and Tyson, G.W. (2015). Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* 350.

Geilert, S., Hensen, C., Schmidt, M., Liebetrau, V., Scholz, F., Doll, M., Deng, L., Fiskal, A., Lever, M.A., Su, C.-C., Schloemer, S., Sarkar, S., Thiel, V., and Berndt, C. (2018). On the formation of hydrothermal vents and cold seeps in the Guaymas Basin, Gulf of California. *Biogeosciences* 15**,** 5715-5731. doi: 10.5194/bg-15-5715-2018.

He, Y., Li, M., Perumal, V., Feng, X., Fang, J., Xie, J., Sievert, S.M., and Wang, F. (2016). Genomic and enzymatic evidence for acetogenesis among multiple lineages of the archaeal phylum Bathyarchaeota widespread in marine sediments. *Nat Microbiol* 1**,** 16035. doi: 10.1038/nmicrobiol.2016.35.

Hensen, C., Nuzzo, M., Hornibrook, E., Pinheiro, L.M., Bock, B., Magalhães, V.H., and Brückmann, W. (2007). Sources of mud volcano fluids in the Gulf of Cadiz—indications for hydrothermal imprint. *Geochim. et Cosmochim. Acta* 71**,** 1232-1248. doi: 10.1016/j.gca.2006.11.022.

Holler, T., Widdel, F., Knittel, K., Amann, R., Kellermann, M.Y., Hinrichs, K.U., Teske, A., Boetius, A., and Wegener, G. (2011). Thermophilic anaerobic oxidation of methane by marine microbial consortia. *ISME J.* 5**,** 1946-1956. doi: 10.1038/ismej.2011.77.

Huse, S.M., Mark Welch, D., Voorhis, A., Shipunova, A., Morrison, H.G., Eren, A.M., and Sogin, M. (2014). VAMPS: a website for visualization and analysis of microbial population structures. *BMC Bioinformatics* 15. doi: doi: 10.1186/1471-2105-15-41.

Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30**,** 772-780. doi: 10.1093/molbev/mst010.

Kevorkian, R., Callahan, S., Winstead, R., and Lloyd, K.G. (2020). ANME-1 archaea drive methane accumulation and removal in estuarine sediment. *bioRxiv* 2020.02.24.963215. doi: 10.1101/2020.02.24.963215.

Kirkpatrick, J.B., Walsh, E.A., and D'Hondt, S. (2019). Microbial Selection and Survival in Subseafloor Sediment. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.00956.

Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79**,** 5112-5120.

Kubo, K., Lloyd, K.G., J.F., B., Amann, R., Teske, A., and Knittel, K. (2012). Archaea of the Miscellaneous Crenarchaeotal Group are abundant, diverse and widespread in marine sediments. *ISME J.* 6**,** 1949-1965. doi: 10.1038/ismej.2012.37.

Lazar, C.S., Baker, B.J., Seitz, K., Hyde, A.S., Dick, G.J., Hinrichs, K.U., and Teske, A.P. (2016). Genomic evidence for distinct carbon substrate preferences and ecological niches of Bathyarchaeota in estuarine sediments. *Environ. Microbiol.* 18**,** 1200-1211. doi: 10.1111/1462-2920.13142.

Lazar, C.S., Biddle, J.F., Meador, T.B., Blair, N., Hinrichs, K.U., and Teske, A.P. (2015). Environmental controls on intragroup diversity of the uncultured benthic archaea of the miscellaneous Crenarchaeotal group lineage naturally enriched in anoxic sediments of the White Oak River estuary (North Carolina, USA). *Environ. Microbiol.* 17**,** 2228-2238. doi: 10.1111/1462-2920.12659.

Lever, M.A., Heuer, V.B., Morono, Y., Masui, N., Schmidt, F., Alperin, M.J., Inagaki, F., Hinrichs, K.-U., and Teske, A. (2010). Acetogenesis in deep subseafloor sediments of the Juan de Fuca Ridge Flank: A synthesis of geochemical, thermodynamic, and gene-based evidence. *Geomicrobiol. J.* 27**,** 183-211. doi: 10.1080/01490450903456681.

Lizarralde, D., Axen, G.J., Brown, H.E., Fletcher, J.M., Gonzalez-Fernandez, A., Harding, A.J., Holbrook, W.S., Kent, G.M., Paramo, P., Sutherland, F., and Umhoefer, P.J. (2007). Variation in styles of rifting in the Gulf of California. *Nature* 448**,** 466-469. doi: 10.1038/nature06035.

Lizarralde, D., Soule, S.A., Seewald, J.S., and Proskurowski, G. (2010). Carbon release by off-axis magmatism in a young sedimented spreading centre. *Nat. Geosci.* 4**,** 50-54. doi: 10.1038/ngeo1006.

824 Lloyd, K.G., Albert, D.B., Biddle, J.F., Chanton, J.P., Pizarro, O., and Teske, A. (2010). Spatial
825 structure and activity of sedimentary microbial communities underlying a *Beggiatoa*
826 spp. mat in a Gulf of Mexico hydrocarbon seep. *PLoS One* 5**,** e8738. doi:
827 10.1371/journal.pone.0008738.
828 Lloyd, K.G., Alperin, M.J., and Teske, A. (2011). Environmental evidence for net methane
829 production and oxidation in putative ANaerobic MEthanotrophic (ANME) archaea.
830 *Environ. Microbiol.* 13**,** 2548-2564. doi: 10.1111/j.1462-2920.2011.02526.x.
831 Lloyd, K.G., Schreiber, L., Petersen, D.G., Kjeldsen, K.U., Lever, M.A., Steen, A.D.,
832 Stepanauskas, R., Richter, M., Kleindienst, S., Lenk, S., Schramm, A., and Jorgensen,
833 B.B. (2013). Predominant archaea in marine sediments degrade detrital proteins.
834 *Nature* 496**,** 215-218. doi: 10.1038/nature12033.
835 Lonsdale, P., and Becker, K. (1985). Hydrothermal plumes, hotsprings, and conductive heat
836 flow in the Southern Trough of Guaymas Basin. *Earth and Planet. Sci. Lett.* 73**,** 211-
837 225.
838 Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and
839 dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15**,** 550. doi:
840 10.1186/s13059-014-0550-8.
841 Martens, C.S. (1990). Generation of short chain organic acid anions in hydrothermally altered
842 sediments of teh Guaymas Basin, Gulf of California. *Appl. Geochem.* 5**,** 71-76.
843 McKay, L., Klokman, V.W., Mendlovitz, H.P., Larowe, D.E., Hoer, D.R., Albert, D., Amend,
844 J.P., and Teske, A. (2016). Thermal and geochemical influences on microbial
845 biogeography in the hydrothermal sediments of Guaymas Basin, Gulf of California.
846 *Environ Microbiol Rep* 8**,** 150-161. doi: 10.1111/1758-2229.12365.
847 McKay, L.J., MacGregor, B.J., Biddle, J.F., Albert, D.B., Mendlovitz, H.P., Hoer, D.R., Lipp,
848 J.S., Lloyd, K.G., and Teske, A.P. (2012). Spatial heterogeneity and underlying
849 geochemistry of phylogenetically diverse orange and white Beggiatoa mats in Guaymas
850 Basin hydrothermal sediments. *Deep Sea Res. I* 67**,** 21-31. doi:
851 10.1016/j.dsr.2012.04.011.
852 McMurdie, P.J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive
853 analysis and graphics of microbiome census data. *PLoS One* 8**,** e61217. doi:
854 10.1371/journal.pone.0061217.
855 McMurdie, P.J., and Holmes, S. (2014). Waste not, want not: why rarefying microbiome data
856 is inadmissible. *PLoS Comput. Biol.* 10. doi: 10.1371/.
857 Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'hara, R.B., Simpson,
858 G.L., Solymos, P., Stevens, M.H.H., and Wagner, H. (2015). vegan: Community
859 Ecology Package. R Package Version 2.2-1. *Available online at:* http://cran.r-
860 project.org/package=vegan.
861 Paull, C.K., Ussler, W., Peltzer, E.T., Brewer, P.G., Keaten, R., Mitts, P.J., Nealon, J.W.,
862 Greinert, J., Herguera, J.-C., and Elena Perez, M. (2007). Authigenic carbon entombed
863 in methane-soaked sediments from the northeastern transform margin of the Guaymas
864 Basin, Gulf of California. *Deep Sea Res. II* 54**,** 1240-1267. doi:
865 10.1016/j.dsr2.2007.04.009.
866 Peter, J.M., and Scott, S.D. (1988). Mineralogy, composition, and fluid-inclusion
867 microthermometry of seafloor hydrothermal deposits in the southern trough of
868 Guaymas Basin, Gulf of California. *Canadian Mineralogist* 26**,** 567-587.
869 Portail, M., Olu, K., Escobar-Briones, E., Caprais, J.C., Menot, L., Waeles, M., Cruaud, P.,
870 Sarradin, P.M., Godfroy, A., and Sarrazin, J. (2015). Comparative study of vent and
871 seep macrofaunal communities in the Guaymas Basin. *Biogeosciences* 12**,** 5455-5479.
872 doi: 10.5194/bg-12-5455-2015.

Racine, J.S. (2012). RStudio: A Platform-Independent IDE for R and Sweave. *Journal of Applied Econometrics* 27**,** 167-172. doi: 10.1002/jae.1278.

Rossel, P.E., Elvert, M., Ramette, A., Boetius, A., and Hinrichs, K.-U. (2011). Factors controlling the distribution of anaerobic methanotrophic communities in marine environments: Evidence from intact polar membrane lipids. *Geochim. Cosmochim. Acta* 75**,** 164-184. doi: 10.1016/j.gca.2010.09.031.

Ruff, S.E., Biddle, J.F., Teske, A.P., Knittel, K., Boetius, A., and Ramette, A. (2015). Global dispersion and local diversification of the methane seep microbiome. *Proc. Natl. Acad. Sci. U. S. A.* 112**,** 4015-4020. doi: 10.1073/pnas.1421865112.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., and Weber, C.F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75**,** 7537-7541. doi: 10.1128/AEM.01541-09.

Seewald, J.S., Seyfried, W.E., and Thornton, E.C. (1990). Organic-rich sediment alteration: an experimental and theoretical study at elevated temperatures and pressures. *Appl. Geochem.* 5**,** 193-209.

Simoneit, B.R.T., Lonsdale, P., Edmond, J.M., and Shank Iii, W.C. (1990). Deep-water hydrocarbon seeps in Guaymas Basin, Gulf of California *Appl. Geochem.* 5**,** 41-49.

Spang, A., Caceres, E.F., and Ettema, T.J.G. (2017). Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* 357. doi: 10.1126/science.aaf3883.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30**,** 1312-1313. doi: 10.1093/bioinformatics/btu033.

Starnawski, P., Bataillon, T., Ettema, T.J., Jochum, L.M., Schreiber, L., Chen, X., Lever, M.A., Polz, M.F., Jorgensen, B.B., Schramm, A., and Kjeldsen, K.U. (2017). Microbial community assembly and evolution in subseafloor sediment. *Proc. Natl. Acad. Sci. U. S. A.* 114**,** 2940-2945. doi: 10.1073/pnas.1614190114.

Teske, A., De Beer, D., Mckay, L.J., Tivey, M.K., Biddle, J.F., Hoer, D., Lloyd, K.G., Lever, M.A., Roy, H., Albert, D.B., Mendlovitz, H.P., and Macgregor, B.J. (2016). The Guaymas Basin Hiking Guide to Hydrothermal Mounds, Chimneys, and Microbial Mats: Complex Seafloor Expressions of Subsurface Hydrothermal Circulation. *Front. Microbiol.* 7**,** 75. doi: 10.3389/fmicb.2016.00075.

Teske, A., Hinrichs, K.U., Edgcomb, V., De Vera Gomez, A., Kysela, D., Sylva, S.P., Sogin, M.L., and Jannasch, H.W. (2002). Microbial Diversity of Hydrothermal Sediments in the Guaymas Basin: Evidence for Anaerobic Methanotrophic Communities. *Appl. Environ. Microbiol.* 68**,** 1994-2007. doi: 10.1128/aem.68.4.1994-2007.2002.

Teske, A., McKay, L., Ravelo, A.C., Aiello, I., Mortera, C., Núñez-Useche, F., Canet, C., Chanton, J.P., Brunner, B., Hensen, C., Ramírez, G.A., Sibert, R.J., Turner, T., White, D., Chambers, C.R., Buckley, A., Joye, S.B., Soule, S.A., and Lizarralde, D. (2019). Characteristics and evolution of sill-driven off-axis hydrothermalism in Guaymas Basin- the Ringvent site. *Scientific Reports* 9. doi: 10.1038/s41598-019-50200-5.

Vigneron, A., Cruaud, P., Pignet, P., Caprais, J.C., Cambon-Bonavita, M.A., Godfroy, A., and Toffin, L. (2013). Archaeal and anaerobic methane oxidizer communities in the Sonora Margin cold seeps, Guaymas Basin (Gulf of California). *ISME J.* 7**,** 1595-1608. doi: 10.1038/ismej.2013.18.

Vigneron, A., Cruaud, P., Roussel, E.G., Pignet, P., Caprais, J.C., Callac, N., Ciobanu, M.C., Godfroy, A., Cragg, B.A., Parkes, J.R., Van Nostrand, J.D., He, Z., Zhou, J., and Toffin, L. (2014). Phylogenetic and functional diversity of microbial communities associated

with subsurface sediments of the Sonora Margin, Guaymas Basin. *PLoS One* 9**,** e104427. doi: 10.1371/journal.pone.0104427.

Von Damm, K.L., Edmond, J.M., Measures, C.I., and Grant, B. (1985). Chemistry of submarine hydrothermal solutions at Guaymas Basin, Gulf of California. *Geochim. Cosmochim. Acta* 49**,** 2221-2237.

Walsh, E.A., Kirkpatrick, J.B., Pockalny, R., Sauvage, J., Spivack, A.J., Murray, R.W., Sogin, M.L., and D'Hondt, S. (2016). Relationship of bacterial richness to organic degradation rate and sediment age in subseafloor sediment. *Appl. Environ. Microbiol.* 82**,** 4994-4999. doi: 10.1128/AEM.00809-16.

Welhan, J.A., and Lupton, J.E. (1987). Light Hydrocarbon Gases in Guaymas Basin Hydrothermal Fluids: Thermogenic Versus Abiogenic Origin. *AAPG Bulletin* 71. doi: 10.1306/94886d76-1704-11d7-8645000102c1865d.

Willis, A., and Bunge, J. (2015). Estimating diversity via frequency ratios. *Biometrics* 71**,** 1042-1049. doi: 10.1111/biom.12332.

Xia, X., Guo, W., and Liu, H. (2017). Basin Scale Variation on the Composition and Diversity of Archaea in the Pacific Ocean. *Front. Microbiol.* 8**,** 2057. doi: 10.3389/fmicb.2017.02057.

Yarza, P., Ludwig, W., Euzeby, J., Amann, R., Schleifer, K.H., Glöckner, F.O., and Rossello-Mora, R. (2010). Update of the All-Species Living Tree Project based on 16S and 23S rRNA sequence analyses. *Syst. Appl. Microbiol.* 33**,** 291-299. doi: 10.1016/j.syapm.2010.08.001.

Yu, T., Wu, W., Liang, W., Lever, M.A., Hinrichs, K.U., and Wang, F. (2018). Growth of sedimentary Bathyarchaeota on lignin as an energy source. *Proc Natl Acad Sci U S A* 115**,** 6022-6027. doi: 10.1073/pnas.1718854115.

Zhou, Z., Liu, Y., Lloyd, K.G., Pan, J., Yang, Y., Gu, J.D., and Li, M. (2019). Genomic and transcriptomic insights into the ecology and metabolism of benthic archaeal cosmopolitan, Thermoprofundales (MBG-D archaea). *ISME J.* 13**,** 885-901. doi: 10.1038/s41396-018-0321-8.

Ziervogel, K., and Arnosti, C. (2020). Substantial Carbohydrate Hydrolase Activities in the Water Column of the Guaymas Basin (Gulf of California). *Front. Mar. Sci.* 6. doi: 10.3389/fmars.2019.00815.