Microbial community diversity, structure, and assembly across oxygen gradients in meromictic marine lakes, Palau

Matthew S. Meyerhof, Jesse M. Wilson, Michael N. Dawson, and J. Michael Beman*

Life and Environmental Sciences, University of California, Merced, California, USA

*Corresponding author: Prof. J. Michael Beman, Life and Environmental Sciences, University of California, Merced, 5200 North Lake Road, Merced, CA 95343, USA.
E-mail: mbeman@ucmerced.edu
Phone: 650.823.7925

Running title: Microbial communities in meromictic marine lakes

Originality–Significance Statement: This study (1) is one of the few studies of ‘marine lake’ microbial communities, (2) represents one of the first direct comparisons of microbial communities inhabiting low-oxygen aquatic ecosystems, and (3) joins a small number of studies examining the relative influence of deterministic versus stochastic processes on microbial community assembly.
Summary

Microbial communities consume oxygen, alter biogeochemistry, and compress habitat in aquatic ecosystems, yet our understanding of these microbial-biogeochemical-ecological interactions is limited by a lack of systematic analyses of low-oxygen ecosystems. Marine lakes provide an ideal comparative system, as they range from well-mixed holomictic lakes to stratified, anoxic, meromictic lakes that vary in their vertical extent of anoxia. We examined microbial communities inhabiting six marine lakes and one ocean site using pyrosequencing of 16S rRNA genes. Microbial richness and evenness was typically highest in the anoxic monimolimnion of meromictic lakes, with common marine bacteria present in mixolimnion communities replaced by anoxygenic phototrophs, sulfate-reducing bacteria, and SAR406 in the monimolimnion. These sharp changes in community structure were linked to environmental gradients (constrained variation in redundancy analysis = 68-76%)—particularly oxygen and pH. However, in those lakes with the steepest oxygen gradients, salinity and dissolved nutrients were important secondary constraining variables, indicating that subtle but substantive differences in microbial communities occur within similar low-oxygen habitats. Deterministic processes were a dominant influence on whole community assembly (all nearest taxon index values >4), demonstrating that the strong environmental gradients present in meromictic marine lakes drive microbial community assembly.
Introduction

Anoxia and hypoxia are widespread in aquatic ecosystems due to the relative insolubility of oxygen in water, as well as its consumption via microbial-mediated oxidative processes. Low oxygen (<20 µM) concentrations occur in stratified freshwater and saline lakes (Hamner and Hamner 1998, Bosshard et al. 2000, Hollibaugh et al. 2001, Koizumi et al. 2004, Crowe et al. 2008, Sarmento et al. 2008), coastal bays (Ferdelman et al. 2006, Zaikova et al. 2010), and marine basins with restricted circulation (Madrid et al. 2001, Vetriani et al. 2003). In the open ocean, oxygen minimum zones (OMZs) are found beneath productive upwelling regions with slow ventilation rates (Keeling et al. 2010). In all of these ecosystems, microbial communities both affect and are affected by the availability of dissolved oxygen (DO). Microbes deplete DO via respiration (Breitburg et al. 2010), and, as DO is consumed, alternative electron acceptors such as nitrate and sulfate are used by microorganisms (Lam and Kuypers 2011). Anaerobic microbial activity subsequently alters local and global biogeochemistry through production and consumption of key chemical compounds. Conversion of dissolved nitrogen to gaseous forms by anaerobic N cycling removes fixed forms of this important nutrient from ecosystems, for instance, and can produce the ozone-depleting greenhouse gas nitrous oxide (Ward et al. 2009, Gilly et al. 2013). Microbial sulfur (S) reduction can also produce hydrogen sulfide, which is toxic to many large organisms (Lavik et al. 2009, Bakun et al. 2010). Wherever they occur, low DO concentrations can reduce habitat area for large aerobic organisms and alter food web structure (Levin 2003, Stramma et al. 2012, Gilly et al. 2013).

Developing a predictive understanding of the interplay between microbial communities, DO concentrations in the water column, and other environmental variables is therefore essential. However, most studies of microbial communities inhabiting low-oxygen systems are confined to a single lake, bay, sea, or OMZ, and comparative studies are rare. For example, microbial communities in the oceans’ major OMZs have only been compared via meta-analysis; this identified ‘typical’ OMZ bacteria (e.g., SAR324, ARCTIC96BD-19/SUP05, Nitrospina; Wright et al. 2012), yet the contributions of these groups to C, N, and S cycling are still ill-defined (Ward et al. 2009, Canfield et al. 2010). Specific operational taxonomic units (OTUs) are also confined to particular OMZs (Wright et al. 2012), leading to high diversity despite low DO concentrations (Stevens and Ulloa 2008, Beman and Carolan 2013). In freshwater lakes, high beta diversity in sulfidic bottom waters is attributed to low dispersal and connectivity between
and among habitats, which may promote microbial ‘endemism’ (Barberan and Casamayor 2011). Collectively these studies suggest that geochemical conditions select for particular functional groups (e.g., denitrifiers, sulfate reducers), but that bacterial community composition as a whole exhibits substantial variation across oxygen-deficient aquatic ecosystems.

The presence of functionally-conserved but taxonomically-diverse communities is consistent with empirical and theoretical studies focusing on the relative roles of deterministic versus stochastic—or niche versus neutral—processes in structuring ecological communities (Chase 2007, Chase 2010, Vellend 2010, Hanson et al. 2012, Stegen et al. 2012, Stegen et al. 2013). Through meta-analysis, Hanson et al. (2012) demonstrated that measured environmental variables (including DO in some cases) typically have stronger effects than neutral processes, explaining 26.9% versus 10.3% of variation in microbial community composition. As a preferred electron acceptor, oxygen is a strong selective force for microbes, but the degree to which it shapes microbial community structure is rarely quantitatively investigated in comparison to other environmental factors. The subsequent, deterministic effects of ‘environmental selection’ on microbial community assembly are also rarely compared with the influence of stochastic processes on community assembly (Hanson et al. 2012, Stegen et al. 2012, Stegen et al. 2013).

We examined community assembly along DO gradients in meromictic marine lakes in the Republic of Palau. DO decreases strongly with depth in these meromictic lakes, creating compressed natural gradients ideal for examining relationships between environmental variation and microbial community diversity, function, structure, and assembly (Landing et al. 1991, Venkateswaran et al. 1993, Hamner and Hamner 1998). In Palau, dozens of marine lakes—bodies of seawater surrounded by land—formed as melting ice sheets raised global sea level after the last glacial maximum, flooding inland basins (Hamner and Hamner 1998, Dawson 2006). Marine lakes have limited connections to the ocean via tunnels and/or fissures in the surrounding karst, precipitation rates are high, and winds are restricted due to steep basin walls (Hamner et al. 1982, Hamner and Hamner 1998, Dawson and Hamner 2005). The degree of connectivity to the ocean, quantity of freshwater input, and extent of wind-driven mixing lead to a range of lake types: from well-mixed (holomictic) bodies of water, to stratified (meromictic) lakes—where a distinct chemocline separates the oxygenated mixolimnion from the anoxic, hydrogen sulfide-containing monimolimnion (Hamner et al. 1982, Hamner & Hamner 1998). We sampled and
compared multiple depths at 7 sites: an ocean site, one holomictic lake, and five meromictic lakes—including the well-known ‘Jellyfish Lake,’ which typically harbors millions of medusae and is meromictic.

Results and Discussion

Water column biogeochemistry in meromictic marine lakes

Along DO, pH, temperature, salinity, and nutrient gradients, we found highly structured communities exhibiting abrupt changes in composition and structure that closely track biogeochemical variation. The depth of anoxia varied across meromictic lakes, with the chemocline—defined as undetectable [DO]—ranging from 2 m depth in Spooky Lake (SLM) to 18 m depth in Ngermeuangel Lake (NLK; Figure 1 and Table 1). Although all meromictic lakes are therefore characterized by anoxia at depth, the steepness of oxygen gradients varies from lake to lake.

We observed sharp increases in NH$_4^+$ and PO$_4^{3-}$ at the chemocline in all meromictic lakes sampled (Figure 2; see also Hamner et al. 1982, Landing et al. 1991), including extreme variations in Goby Lake (GLK), where NH$_4^+$ reached 1090 µM and PO$_4^{3-}$ reached 52.8 µM. However, oxidized N was infrequently detected in our samples: in contrast to typical depth profiles in the ocean, nitrate (NO$_3^-$) ranged from 0-1.02 µM and was maximal above the chemocline in most of the lakes. NO$_3^-$ was rarely detected below the chemocline and only in small amounts, which we attribute to use of any available NO$_3^-$ as an electron acceptor under anoxic conditions. In contrast to micromolar-level accumulations of nitrite (NO$_2^-$) typically observed within oceanic OMZs, NO$_2^-$ was detected in 18 of 52 meromictic lake samples. This included samples near the chemocline in Jellyfish Lake (OTM) (12-13 m) and NLK (19-20 m), as well as elevated NO$_2^-$ in anoxic waters in SLM (>2 m depth) and GLK (>10 m). Like NO$_3^-$, some mixomolimnion samples also exhibited elevated NO$_2^-$ concentrations, but the general patterns for these two N species were opposed.

Microbial diversity

We examined alpha diversity of communities collected from the mixomolimnion, mid-depth, and the monimolimnion in each lake based on the number of operational taxonomic units (OTUs) observed, the ACE richness estimator (Chao et al. 1993), the Shannon index (Shannon 1948),
and Pielou’s evenness metric (Pielou 1966). Across samples, values ranged from 416–1,162 observed OTUs, 2,029–9,651 estimated OTUs, Shannon index values of 3.06 ± 0.08 to 5.74 ± 0.08, and Pielou’s evenness values of 0.43–0.79 (Figure 3). In all meromictic lakes, observed richness, ACE-estimated richness, Shannon, and Pielou’s evenness were uniformly highest under anoxic and sulfidic conditions in the monimolimnion (Fig. 3). SLM and NLK generally showed monotonic increases in diversity with depth (with the exception of ACE-estimated richness in SLM), but the other 3 lakes showed minima in richness, the Shannon Index, and Pielou’s evenness at intermediate depths. Intermediate minima can be attributed to the dominance of anoxygenic photosynthetic bacteria in the chemocline. This is particularly evident at 5 m depth in GLK and 15 m depth in CLM, where evenness values were low (0.49 and 0.59, respectively; Fig. 3) and Chlorobi were abundant (64% and 24% of communities; Figure 5). For all but ACE in SLM, trends in richness, Shannon, and Pielou’s evenness were identical in the lakes, indicating that both richness and evenness respond to similar factors. High diversity under anoxic and sulfidic conditions is consistent with work by Barberan and Casamayor (2011), who found higher alpha diversity at depth within sulfidic freshwater lakes. However, their dataset was based on fingerprinting and cloning approaches that could resolve at most 10s of OTUs, whereas our data extend this pattern to 1000s of OTUs.

Absolute values for richness, the Shannon index, and Pielou’s evenness varied from depth to depth and from lake to lake. Mixolimnia of SLM, GLK, CLM, and NLK were generally less diverse than other depths and lakes; GLK and CLM showed wide ranges in diversity; and OTM displayed relatively high diversity throughout the water column (ACE values in OTM ranged from 4,390 to 8,837 OTUs). By comparison, the holomictic lake (Mekeald Lake; MLN) and ocean (OS-GC) site (Fig. 1 and Table 1) were also consistently diverse across depths, with observed richness ranging from 927–1331 OTUs, and estimated richness ranging from 5,503-9,651 OTUs. In comparison to surrounding ocean waters, lower richness was observed in the lakes, and lake Shannon Index values bracketed those observed in the ocean (Fig. 3). Altogether, our data support the idea that low DO aquatic habitats can be microbially diverse (both rich and even)—specifically within the monimolimnion—but with significant variability with depth and from lake to lake.

Microbial community structure

Wiley-Blackwell and Society for Applied Microbiology
This article is protected by copyright. All rights reserved.
Variations in alpha diversity within and across lakes may reflect underlying differences in community structure and assembly. Marine lakes are ‘seeded’ with marine, freshwater, and terrestrial microorganisms through exchange with the ocean, freshwater runoff/groundwater, and erosion; lake communities are then assembled via deterministic or stochastic processes, and could closely resemble a given end-member, reflect a mixture of sources, or be entirely novel. These potential sources of microbes are increasingly well-studied and show consistent compositions with relatively little overlap (Zwart et al. 2002, Lozupone and Knight 2007, Wright et al. 2012). We found that marine lakes were dominated by characteristic marine bacteria (Figures 4 and 5). Cyanobacteria were detected in every sample, generally decreased with depth in the meromictic lakes, and were abundant throughout the open ocean samples (Figure 4). Lake cyanobacterial communities were dominated by Synechococcus; Prochlorococcus was not detected in any lake samples despite being present in ocean water at similar abundances to Synechococcus. This is consistent with the known distribution of these two common marine phototrophs, where Prochlorococcus is typically confined to oligotrophic gyres, while Synechococcus is more widely distributed (Partensky and Garczarek 2010).

α-proteobacteria were also prevalent throughout the meromictic lakes, and the vast majority of α-proteobacteria sequences were from SAR11, with minor contributions from other α-proteobacteria such as SAR116 (up to 6.58% of all 16S rRNA sequences) and the Roseobacter clade (up to 1.80%). This reflects the ubiquity and abundance of the SAR11 clade in the ocean (Morris et al. 2002), and SAR11 in fact constituted up to 44.4% of sequences collected at 30 m at the ocean site. SAR11 bacteria were also found at high levels at intermediate depths in the meromictic lakes OTM and NLK (>16% at 10 m in both lakes) but formed <2% of the community in anoxic waters (Fig. 4). These observations are consistent with the presence of SAR11 in OMZs (e.g., Stevens and Ulloa 2008, Stewart et al. 2011, Beman and Carolan 2013), as well as with the absence of SAR11 in sulfidic basins (Vetriani et al. 2003, Zaikova et al. 2010). These observations also reflect the significant diversity found within the SAR11 clade, and the majority of marine lake sequences were associated with SAR11 surface cluster 1 (83.8%) followed by surface cluster 2 (10.4%), with minor contributions from other SAR11 clusters.

Other marine bacteria that are common in surface waters of the ocean were present in the mixolimnia of marine lakes, including SAR86, marine Actinobacteria, and Bacteriodetes (specifically Flavobacteria and Sphingobacteria). For the most part, these mixolimnia
communities therefore resemble typical epipelagic ocean microbial communities. However, there were differences in community structure across different lakes, and substantial differences with depth in individual lakes. For example, the β-proteobacteria were present in all samples, and were comprised primarily of methylotrophs that were most abundant in mixolimnion samples. They formed 1.11% of the 5 m CLM library, 2.02% of the 0 m SLM library, and 2.12% of the 5 m OTM library. Although Archaea were found in all samples, no known methanogens were detected and nearly all archaeal sequences were unclassified Crenarchaeota or Euryarchaeota (Table 2). Archaea also showed relatively little variation with depth and between meromictic lakes, typically ranging from 4-10% of communities. NLK was a key exception where Archaea formed 15.8% of the community in the monimolimnion and 9.74% of the community at mid-depth.

δ-proteobacteria—specifically the sulfate-reducing bacteria (SRB) Desulfobacterales—were abundant in the anoxic monimolimnion of all the meromictic lakes (Figure 5). SRBs alone comprised 7.85-18.7% of 16S rRNA gene sequences in these samples. Anoxygenic photosynthetic bacteria were also present in monimolimnion samples from all meromicite lakes; these bacteria are known to form dense ‘plates,’ present in high abundance at discrete depths in some of the meromictic lakes (Hamner et al. 1982, Hamner & Hamner 1998, Venkateswaran et al. 1993). Chlorobi (green sulfur bacteria) constituted 24.4-64.0% of the intermediate depth microbial community in CLM and GLK, where light levels are higher and the availability of sulfide diffusing upwards provides ideal habitat (Fig. 5). γ-proteobacterial Chromatiales (purple sulfur bacteria) were variable among the meromictic lakes, and were most abundant at 20 m in NLK (19.7%) and 14 m in OTM (5.06%).

16S rRNA sequences from γ-proteobacteria were detected in all samples and were dominant in many samples, contributing 20% of all sequences recovered from the meromictic lakes. The γ-proteobacteria varied in abundance from lake to lake and from depth to depth within individual lakes, but in contrast to some other bacterial groups, this was not driven by the prevalence of any single sequence type or ecotype. In addition to SAR86 and Chromatiales, multiple other γ-proteobacteria groups were commonly detected, including various groups within the Alteromonas, Pseudoalteromonas, Thiotrichales, and Oceanosprilla. In some samples, many γ-proteobacterial sequences were also unclassified—e.g., 17.6% of the 1 m SLM library and 15.5% of the 15 m CLM library. These findings closely resemble those of Stevens and Ulloa
(2008) in the OMZ of the ETSP: they found high abundances of γ-proteobacteria (39-69% of bacterial clone libraries), representing many different known and unclassified groups.

However, a key difference between meromictic marine lake and OMZ communities is that the putative sulfur oxidizers SUP05/ARCTIC96BD-19 were not as abundant in lakes as they are in open ocean OMZs (Walsh et al. 2009, Lavik et al. 2009, Wright et al. 2012). SUP05 was detected in low abundance (<1%) at intermediate depth in CLM, and at intermediate depths and in the monimolimnion in OTM and NLK, and only a single ARCTIC96BD-19 sequence was detected. The S-oxidizing δ-proteobacterial SAR324 group (Swan et al. 2011, Sheik et al. 2013) was more abundant and peaked at intermediate depths in the meromictic lakes, exceeding 2% of sequences at intermediate depths in OTM, Clear Lake, and NLK. Because our knowledge of these groups is drawn from relatively few (meta)genomes (Walsh et al. 2009, Swan et al. 2011, Sheik et al. 2013), it is unclear what may drive differences in the presence and abundance of SUP05/ARCTIC96BD-19 and SAR324 in different lakes and in comparison to OMZs. Wright et al. (2012) argue for succession from ARCTIC96BD-19 to SUP05 to ε-proteobacteria (Sulfurimonas and Arcobacter) as DO decreases and H₂S increases. Consistent with this, Arcobacter were abundant in the monimolimnion of SLM (2.5%) and OTM (10% of the community), but not at intermediate depths.

SAR406 (or Marine Group A) bacteria were also most abundant at intermediate depths, and have been identified as S-cyclers based on genome content (Wright et al. 2014). Collectively our data therefore indicate that intermediate depths in some meromictic lakes resemble open ocean OMZs in that (1) DO is present at low levels with oxidized N (albeit at low levels), (2) marine bacteria prevalent in OMZs (SAR324 and SAR406) are abundant, and (3) sulfate reducers were not abundant. This niche appears to expand from lakes with shallower chemoclines and broad sulfidic zones (SLM and GLK) to those with deeper chemoclines (OTM, CLM, and NLK), as SAR324 and SAR406 are more abundant at intermediate depths in the latter lakes.

Factors influencing community structure and assembly

Community similarity (Sørensen abundance-based similarity) declined across these depth gradients, with top to bottom similarities ranging from 10% (CLM) to 51% (GLK) in meromictic lakes. Similarities between top and middle samples were 45–86%, and middle to anoxic
similarities were 22–79%. In contrast, communities were 75–83% similar between different depths in holomictic MLN and 86-87% similar at different depths in the surrounding ocean. In all lakes, communities become less similar with increasing depth and changing environmental conditions—a pattern that has been observed in other stratified, oxygen-deficient aquatic ecosystems. In the Black Sea, for instance, DO concentrations decrease with depth to the chemocline, at which point hydrogen sulfide begins to increase (Murray et al. 1989). Vetriani et al. (2003) demonstrated distinct changes in microbial community composition throughout the oxic and anoxic regions of the Black Sea: although many OTUs were found at more than one sampled depth, only 2 OTUs were detected in every sample collected, and most were confined to anoxic, suboxic, or oxic layers (Vetriani et al. 2003). The Cariaco Basin also shows microbial communities separated into three distinct layers, and microbial communities present in anoxic waters resemble those found in other anaerobic habitats (Madrid et al. 2001, Lin et al. 2006).

Detail can be added to these comparisons by examining variations across different lakes in addition to within lakes in Palau. Depth gradients reflect the interplay between microbial activity (e.g., consumption of DO, nutrient uptake and cycling) and physical limnology and oceanography, such that many factors co-vary and their relative influences cannot be easily isolated within a single lake. Across multiple marine lakes, differences in the steepness and strength of environmental gradients allow these factors and processes to be disentangled. We used redundancy analysis (RDA) to determine the degree to which shifts in community structure are constrained by different environmental variables. RDA demonstrated that 68% of community variation was explained by variations in DO concentrations, pH, conductivity, salinity, chlorophyll concentrations, temperature, and dissolved nutrient concentrations (ANOVA P=0.002; Figure 6). 24% of the variation in community structure was explained by the first RDA axis, and 12%, 9%, 7%, and 5% by the second, third, fourth, and fifth axes.

As expected, DO gradients were associated with shifts in community structure, and DO was a substantial contributor to RDA axis 1—along with a nearly equal contribution from pH (Fig. 6). This axis explained much of the variation in CLM, OTM, and NLK communities. GLK and SLM samples were less distinguishable along RDA axis 1, with more of the variation in community structure captured by axis 2. Salinity and dissolved nutrients—especially nitrate concentrations—contributed most strongly to RDA axis 2. This suggests that salinity and nutrients are important secondary factors affecting community structure across meromictic
marine lakes, and are primary factors affecting community structure in GLK and SLM. Holomictic lake and ocean samples mostly clustered together; excluding them from RDA produced similar patterns of influence for the different environmental variables, and increased the constrained proportion of community variation to 76% (Figure S1). This is likely due to subtle differences in community structure (Figs. 4 and 5) that occur without corresponding differences in environmental conditions.

Results from RDA are therefore consistent with the idea that strong biogeochemical gradients present in meromictic lakes drive deterministic—rather than stochastic—microbial community assembly. We used the framework used by Stegen et al. (2012; 2013) and adapted from the ecology literature (Webb et al. 2002), to evaluate whether or not this is the case based on the mean-nearest-taxon-distance (MNTD) and the nearest-taxon-index (NTI). MNTD is a measure of phylogenetic relatedness within a community, and NTI quantifies the number of standard deviations that the observed MNTD is from the mean MNTD values generated by null models. We found that all meromictic lakes sampled had NTI values significantly (all \( P<0.005 \)) and substantially (4.49-9.00) greater than 2 (Figure 7). This is indicative of deterministic processes affecting community assembly, leading to higher than expected phylogenetic clustering (Webb et al. 2002, Stegen et al. 2012, Stegen et al. 2013). In general, NTI values were highest in the monimolimnion and reached values of 7.64-8.38, meaning that the observed MNTD in these samples exceeds its null distribution by over 7 standard deviations.

Collective variations in oxygen, pH, and dissolved nutrients therefore appear to influence deterministic assembly of microbial communities within and across meromictic marine lakes. We expect that without coincident gradients in multiple variables, stochastic processes may have a greater influence on community assembly. For example, Chase (2007) demonstrated that the ‘harshness’ of environmental filters determines the relative dominance of deterministic and stochastic processes. Stochasticity may therefore be more important in holomictic lakes, or in the open ocean, where environmental gradients are less pronounced. Variations in community structure within and across lakes are explained primarily by DO and pH; these two variables track each other because decreases in pH with depth are driven primarily by carbon dioxide production via respiration of organic matter, which consumes DO. The OTUs that track variations in DO concentrations include both those that deplete DO via respiration, as well as those that respond favorably to its depletion. Based on results from RDA, individual
Synechococcus and SAR11 OTUs were frequently positively correlated with DO, while SAR406 and SRB OTUs were frequently inversely correlated with DO—however, production and respiration rate measurements, experimental manipulations, and additional sampling will provide more detailed insight into the specific OTUs that are drivers of, and responders to, decreasing DO concentrations. In those lakes with the steepest DO gradients (SLM and GLK), salinity and dissolved nutrients are important additional constraining variables. This indicates that subtle but substantive differences in microbial communities occur within similar low-oxygen habitats. In other words, DO exerts a strong influence on microbial communities but acts in concert with other variables to affect assembly. The continuum of marine lakes available in Palau provides an ideal comparative framework to quantitatively partition these effects and their relative influences as deterministic factors affecting microbial community assembly.

Experimental Procedures

Study Sites and Sample Collection. We studied five meromictic lakes in Palau: Spooky Lake (SLM), Goby Lake (GLK), Ongeim'l Tketau Lake (OTM, known colloquially as Jellyfish Lake), Clear Lake (CLM), and Ngermeuangel Lake (NLK). One holomictic lake (Mekeald Lake; MLN) and an ocean site at the German Channel on the southwestern side of the islands (OS-GC) were also sampled for comparison (Figure 1 and Table 1). We vertically profiled dissolved oxygen (DO), temperature, pH, chlorophyll fluorescence, and salinity/conductivity using a HydroLab DS5 (Hach Company, Loveland, CO, USA) and sampled three depth layers within each meromictic lake: the mixolimnion (0–5 m depth), the monimolimnion (5–20 m depth), and intermediate depths near the chemocline; these intermediate depths ranged from 1 to 15 m, depending on the depth of the chemocline within the individual lakes. We sampled comparable depths at MLN (5–20 m) and OS-GC (5–30 m). For QPCR analysis of functional genes (dsrA, amoA, nirS) and specific functional groups, as well as analysis dissolved nutrient concentrations, we collected additional samples above and below the chemocline to capture abrupt transitions in biogeochemical conditions across this interface.

Samples were collected from small boats using a horizontal, 2.5 L GoFlo bottle (General Oceanics, Miami, FL, USA), transferred to 1 L polycarbonate bottles, and stored in the dark during transit to the Coral Reef Research Foundation laboratory in Koror, Palau. Water samples were filtered using a peristaltic pump and 0.22 µm Durapore PVDF hydrophilic filters.
Filters were immediately frozen in 800 µL Sucrose-Tris-EDTA (STE) lysis buffer (750 mM sucrose, 20 mM EDTA, 400 mM NaCl, and 50 mM Tris) in 2 mL Lysing Matrix E tubes (MP Biomedicals, Solon, OH, USA), and stored at -20°C until transport to the United States, where they were stored at -80°C until extraction. (Dry ice and liquid nitrogen are not readily available in Palau.)

**Nutrient Measurements.** During sample filtration, 50 mL of filtrate was collected in high-density polyethylene bottles for subsequent nutrient analysis at the University of California, Santa Barbara (UCSB) Marine Analytical Laboratory. Samples were analyzed for ammonium (UCSB MAL analytical method for ammonium, see below; Diamond and Huberty 1996), nitrite (Environmental Protection Agency (EPA) Method 353.2, Schroeder 1997), nitrite+nitrate (EPA Method 353.2, Diamond 1997), and phosphate (EPA Method 365.1, Huberty and Diamond 1998), on a Lachat QuikChem 8000 Flow Injection Analyzer (Hach Company, Loveland, CO, USA). A handful of samples containing large concentrations of sulfide were not analyzed for nitrate, as sulfide damages the cadmium reduction column. For ammonium analysis, each sample was injected into a flowing carrier stream through an injection valve, and then merged with an alkaline solution stream; the produced ammonia was diffused through a hydrophobic, gas-permeable membrane into a recipient stream containing a pH indicator. Color change occurs in the indicator solution due to an increase in pH, and the concentration of ammonia was determined spectrophotometrically based on absorption at 570 nm. For all analyses, a mid-range check standard bracketed every 20 samples to verify the accuracy of the measurements, and samples that were detected outside of the standards’ range were diluted 1:10 and reanalyzed. Detection limits were 0.10 µM for phosphate, 0.10 µM for nitrite, 0.20 µM for nitrite+nitrate, and 0.10 µM for ammonium.

**DNA Extraction and quantification.** DNA was extracted following Beman et al. (2008): 100µL 10% sodium dodecyl sulfate (SDS) was added to tubes containing STE buffer and filters, samples were bead-beat for 2 minutes (BioSpec Products, Inc., Bartlesville, OK, USA), and incubated for 3 minutes on a dry heat block at 99°C. Following transfer of sample solutions to 1.5 mL LoBind Microcentrifuge tubes (Eppendorf, Hauppauge, NY, USA), 50µL proteinase K (20mg mL⁻¹; Qiagen, Inc., Valencia, CA, USA) was added, and tubes were incubated at 55°C for
3 hours. Lysates were purified using a DNeasy Blood and Tissue Kit according to the manufacturer’s protocol (Qiagen, Inc., Valencia, CA, USA). DNA concentrations were measured using the Quant-iT PicoGreen dsDNA Assay Kit and the manufacturer’s protocol (Invitrogen Corporation, Carlsbad, CA, USA) on a Stratagene MX 3005P (Agilent Technologies, Inc., Santa Clara, CA, USA).

**Pyrosequencing.** Sequencing and microbial community analysis followed Beman and Carolan (2013). In brief, we used the primers 926F (5\'AACCTGAGGTGACGGCAGC-3\') and 1392R (5\'ACGGGCGGTGTGTRC-3\')—which are effective universal primers (for bacteria and archaea) that do not inflate richness estimates (Engelbrekstrom et al. 2010)—to amplify a portion of the 16S rRNA gene. Linkers and 8-base ‘barcodes’ (Hamady et al. 2008) were attached to primers, and barcodes were used to sort individual samples. DNA samples were sequenced using Titanium chemistry on the Roche 454 FLX platform at Research and Testing Laboratories (Lubbock, TX, USA). We recovered 285,815 sequences (from 21 samples) with a median read length of 460 bp. Sequence data have been deposited in the Sequence Read Archive (SAMN04274137).

We used the program mothur (version 1.31.2; [http://www.mothur.org; Schloss et al. 2009]) for all sequence analyses, and employed the approach of Huse et al. (2007) for quality control. Sequences that met one or more of the following criteria were discarded: the sequence length was more than \( \pm 100 \) bp from the median, contained ambiguous bases, contained homopolymers \( >8 \) bp, had \( <25 \) average quality score, or did not exactly match the forward primer and barcode sequence. Of the sequences that were removed, most were removed because they did not exactly match the forward primer and barcode. (We did not screen sequences based on the reverse primer, as some of the reads are high quality sequences that did not extend to the reverse primer.) 13.9% of sequences did not meet quality control criteria and were excluded from subsequent analyses. The remaining 246,087 bacterial and archaeal 16S rRNA sequences were aligned to the greengenes universal alignment (2011 version) that includes both archaea and bacteria (DeSantis et al. 2006). Kunin et al. (2010) raised the issue of pyrosequencing errors inflating diversity estimates, and Huse et al. (2010) proposed a pseudo-single linkage ‘preclustering’ algorithm as a means of reducing these errors; we used this approach as implemented in mothur to remove sequences that may be affected by pyrosequencing errors. In addition, we analyzed
much longer sequences (460 bp vs. 60-110 bp for ‘pyrotags’), and calculated diversity indices at
97% identity, to avoid error-inflated diversity estimates.

For subsequent analysis, libraries were normalized by randomly removing sequences to
reduce the total number of sequences in each library to a common size of 7000 (individual
sample library size ranged from 7,063–23,983 sequences with a mean of 11,718 ± 4,223 sd).
Alpha diversity indices (observed richness, ACE, Shannon, Pielou’s evenness) were calculated
for libraries at 97% sequence identity using average neighbor clustering. Beta diversity
calculations were performed across these samples and changes in community
composition/structure were explored using multiple indices (abundance-based Sørensen, Bray-
Curtis, 0); we focus primarily on the abundance-based Sørensen index. Taxonomic composition
of individual samples was determined by classification of 16S rRNA sequences based on the
ARB SILVA dataset (release 102; Pruesse et al. 2007) in mothur, and data are reported for a
consensus confidence threshold of 65%.

**Statistical analyses** were conducted in the R statistical environment (v. 3.2.0). We used the
vegan package for redundancy analysis (RDA). Mean-nearest-taxon-distance (MNTD) and the
net-taxon-index (NTI) were analyzed in the picante package; for these calculations, we used the
ses.mntd command with null model=‘taxa.names,’ abundance.weighted=TRUE, and 999
randomizations. NTI is the inverse of ses.mntd and quantifies the number of standard deviations
that the observed MNTD is from the mean of its null distribution. NTI values greater than 2 or
less than -2 are interpreted as evidence of deterministic processes such as environmental
selection.

**Acknowledgments**

We thank Pat Colin and Lori Colin of the Coral Reef Research Foundation (CRRF) for providing
laboratory space, logistical support, and access to field sites. Emilio Basilius, Gerda Ucharm,
and Matt Mesubed (CRRF), Sharon Patris (UC Merced) and Herwig Stibor (Ludwig-
Maximilians University, Munich) provided invaluable help in the field, and we thank Emilio,
Gerda, and Sharon for collecting Hydrolab profiles. We also wish to thank Susan Alford, Lolo
Cardenas, Curtis Hayden, and Rodney Severson for assistance with laboratory analyses, and
thank the United States National Science Foundation Dimensions of Biodiversity Program
(OCE-1241255) and UC Merced Graduate and Research Council for providing funds to support this research.

**Figure Legends**

**Figure 1.** Map of sampling locations in Palau (A), and profiles of dissolved oxygen, temperature, and salinity in meromictic lakes (B-F). Sites are designated as follows: Spooky Lake (SLM), Goby Lake (GLK), Ongeim’l Tketau Lake (OTM), Clear Lake (CLM), Ngermeuangel Lake (NLK), and Mekeald Lake (MLN), as well as an open ocean site (OS-GC). (B-F) Depth profiles show oxygen, temperature, and salinity at the time of sampling. Grey shading indicates the depth of the chemocline in each lake, with the lakes ordered from left to right based on the depth of the chemocline. Data from all lakes are plotted on the same vertical axis for purposes of comparison; lake depths are reported in Table 1.

**Figure 2.** Depth profiles of (A) phosphate, (B) ammonium, (C) nitrate, and (D) nitrite concentrations in meromictic lakes show sharp increases in phosphate and ammonium below the chemocline, while nitrate and nitrite are more variable. Data (µmol L⁻¹) are plotted relative to the depth of the chemocline in each lake, such that all profiles are collapsed onto a common axis. Positive values indicate the relative depth above the chemocline, and negative values indicate depths below the chemocline. The color shading denotes the different lakes.

**Figure 3.** Alpha diversity of microbial communities displays higher richness and evenness in the monimolimnion of meromictic marine lakes. Site names are shown along the left side of the figure with the depths of sampling. For each meromictic lake, these depths represent, from top to bottom, the mixolimnion, intermediate depth, and the monimolimnion. The plots show (A) observed richness, (B) ACE-estimated richness, (C) the Shannon Index, and (D) Pielou’s evenness. Samples were normalized to a common library size of 7000 sequences, and all calculations were performed at 97% sequence identity.

**Figure 4.** The relative abundance of microbial groups prevalent in the mixolimnion of meromictic lakes, a holomictic lake (MLN), and an ocean site (OS-GC) demonstrates that common marine bacteria inhabit marine lakes. Site names are shown along the left side of the
Figure 5. The relative abundance of microbial groups prevalent in the monimolimnion of meromictic lakes includes known anaerobic groups. Site names are shown along the left side of the figure with the depths of sampling. Colors depict bacterial groups that comprise more than 1% of any library. For the Chromatiales, empty bars denote the abundance of all gammaproteobacteria, and for sulfate-reducing bacteria and SAR324, empty bars denote the abundance of all deltaproteobacteria. Note differences in scales between axes.

Figure 6. Biplot of RDA results demonstrates that variations in oxygen and pH are strongly related to changes in microbial community structure. Different samples and environmental variables are plotted on RDA axes 1 and 2, which respectively capture 24% and 12% of the variation in community structure. Overall, 68% of variation in community structure could be constrained. Different lakes are indicated by the color-coded circles, with lake and depth of sampling listed next to each data point. Arrows denote biplot scores for the different constraining environmental variables.

Figure 7. NTI values for marine lake microbial communities show the influence of environmental selection (NTI>2) at the whole community level. Site names are shown along the left side of the figure with the depths of sampling.
### Table 1. General properties of sampled marine lakes in Palau (from Colin 2009).

<table>
<thead>
<tr>
<th>Lake</th>
<th>Type</th>
<th>Chemocline Depth (m)</th>
<th>Maximum Depth (m)</th>
<th>Area (m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spooky (SLM)</td>
<td>Meromictic</td>
<td>2</td>
<td>14</td>
<td>11000</td>
</tr>
<tr>
<td>Goby (GLK)</td>
<td>Meromictic</td>
<td>5</td>
<td>15</td>
<td>21000</td>
</tr>
<tr>
<td>Jellyfish (OTM)</td>
<td>Meromictic</td>
<td>12</td>
<td>30</td>
<td>50000</td>
</tr>
<tr>
<td>Clear (CLM)</td>
<td>Meromictic</td>
<td>16</td>
<td>30</td>
<td>39000</td>
</tr>
<tr>
<td>Ngermeuangel (NLK)</td>
<td>Meromictic</td>
<td>18</td>
<td>38</td>
<td>43000</td>
</tr>
<tr>
<td>Mekeald (MLN)</td>
<td>Holomictic</td>
<td>n/a</td>
<td>27</td>
<td>20300</td>
</tr>
</tbody>
</table>
Table 2. Archaeal 16S rRNA sequences as a percentage of sequence libraries

<table>
<thead>
<tr>
<th>Lake</th>
<th>Depth</th>
<th>Archaea (% of 16S rRNA sequences)</th>
<th>Crenarchaeota (% of 16S rRNA sequences)</th>
<th>Euryarchaeota (% of 16S rRNA sequences)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLM</td>
<td>0</td>
<td>4.21</td>
<td>2.24</td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>5.06</td>
<td>3.77</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5.33</td>
<td>4.06</td>
<td>1.28</td>
</tr>
<tr>
<td>GLK</td>
<td>1</td>
<td>3.45</td>
<td>2.70</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>1.57</td>
<td>0.89</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6.60</td>
<td>3.64</td>
<td>2.96</td>
</tr>
<tr>
<td>OTM</td>
<td>5</td>
<td>6.67</td>
<td>5.86</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>8.45</td>
<td>5.81</td>
<td>2.64</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>3.59</td>
<td>2.69</td>
<td>0.89</td>
</tr>
<tr>
<td>CLM</td>
<td>5</td>
<td>8.60</td>
<td>7.08</td>
<td>1.52</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>4.74</td>
<td>3.92</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>6.84</td>
<td>4.20</td>
<td>2.64</td>
</tr>
<tr>
<td>NLK</td>
<td>5</td>
<td>6.48</td>
<td>4.84</td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>9.47</td>
<td>5.97</td>
<td>3.50</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>15.8</td>
<td>14.3</td>
<td>1.43</td>
</tr>
<tr>
<td>MLN</td>
<td>5</td>
<td>13.5</td>
<td>12.4</td>
<td>1.07</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6.41</td>
<td>5.05</td>
<td>1.36</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>5.51</td>
<td>3.74</td>
<td>1.77</td>
</tr>
<tr>
<td>Ocean</td>
<td>5</td>
<td>4.43</td>
<td>2.56</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>5.25</td>
<td>3.62</td>
<td>1.63</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>5.59</td>
<td>2.91</td>
<td>2.68</td>
</tr>
</tbody>
</table>
References


Figure 2
72x31mm (300 x 300 DPI)
Figure 7
90x111mm (300 x 300 DPI)