





## SYMPOSIUM

# Phenotypic Comparability from Genotypic Variability among Physically Structured Microbial Consortia

Stephanie K. Hoffman <sup>1,\*,\dagger</sup> Kiley W. Seitz,<sup>\*,\ddagger</sup> Justin C. Havird <sup>\*,§</sup> David A. Weese,<sup>\*,¶</sup> and Scott R. Santos<sup>\*</sup>

<sup>\*</sup>Department of Biological Sciences and Molette Laboratory for Climate Change and Environmental Studies, Auburn University, Auburn, AL 36849, USA; <sup>\dagger</sup>Department of Biological Sciences, Green River College, Auburn, WA 98092, USA; <sup>\ddagger</sup>Strutural and Computational Biology Unit, European Molecular Biological Laboratory, 69117 Heidelberg, Germany; <sup>\§</sup>Department of Integrative Biology, University of Texas at Austin, Austin, TX 78712, USA; <sup>\¶</sup>Department of Biological and Environmental Sciences, Georgia College & State University, Milledgeville, GA 31061, USA

From the symposium “SICB Wide Symposium: Building Bridges from Genome to Phenome: Molecules, Methods and Models” presented at the annual meeting of the Society for Integrative and Comparative Biology January 3–7, 2020 at Austin, Texas.

<sup>1</sup>E-mail: shoffman@greenriver.edu

**Synopsis** Microbiomes represent the collective bacteria, archaea, protist, fungi, and virus communities living in or on individual organisms that are typically multicellular eukaryotes. Such consortia have become recognized as having significant impacts on the development, health, and disease status of their hosts. Since understanding the mechanistic connections between an individual’s genetic makeup and their complete set of traits (i.e., genome to phenome) requires consideration at different levels of biological organization, this should include interactions with, and the organization of, microbial consortia. To understand microbial consortia organization, we elucidated the genetic constituents among phenotypically similar (and hypothesized functionally-analogous) layers (i.e., top orange, second orange, pink, and green layers) in the unique laminated orange cyanobacterial–bacterial crusts endemic to Hawaii’s anchialine ecosystem. High-throughput amplicon sequencing of ribosomal RNA hypervariable regions (i.e., *Bacteria*-specific V6 and *Eukarya*-biased V9) revealed microbial richness increasing by crust layer depth, with samples of a given layer more similar to different layers from the same geographic site than to their phenotypically-analogous layer from different sites. Furthermore, samples from sites on the same island were more similar to each other, regardless of which layer they originated from, than to analogous layers from another island. However, cyanobacterial and algal taxa were abundant in all surface and bottom layers, with anaerobic and chemoautotrophic taxa concentrated in the middle two layers, suggesting crust oxygenation from both above and below. Thus, the arrangement of oxygenated vs. anoxygenated niches in these orange crusts is functionally distinct relative to other laminated cyanobacterial–bacterial communities examined to date, with convergent evolution due to similar environmental conditions a likely driver for these phenotypically comparable but genetically distinct microbial consortia.

## Introduction

Elucidating relationships between the genetic makeup (i.e., genome) of an individual and its entirety of potential physical and/or physiological characteristics or traits (i.e., phenome) are among the Grand Challenges currently confronting biologists. As evident from the “Building Bridges from Genome to Phenome” symposium at the 2020 Annual Meeting of the Society for Integrative and Comparative

Biology (SICB), much work remains toward understanding the mechanistic basis of the “genome to phenome” in spite of numerous technological advances in this area. Among these challenges is the need to encompass different levels of biological organization, and recognize that important phenotypes such as development, health, and disease status can be attributed to the identity and genetics of microorganisms intimately associated with the specific

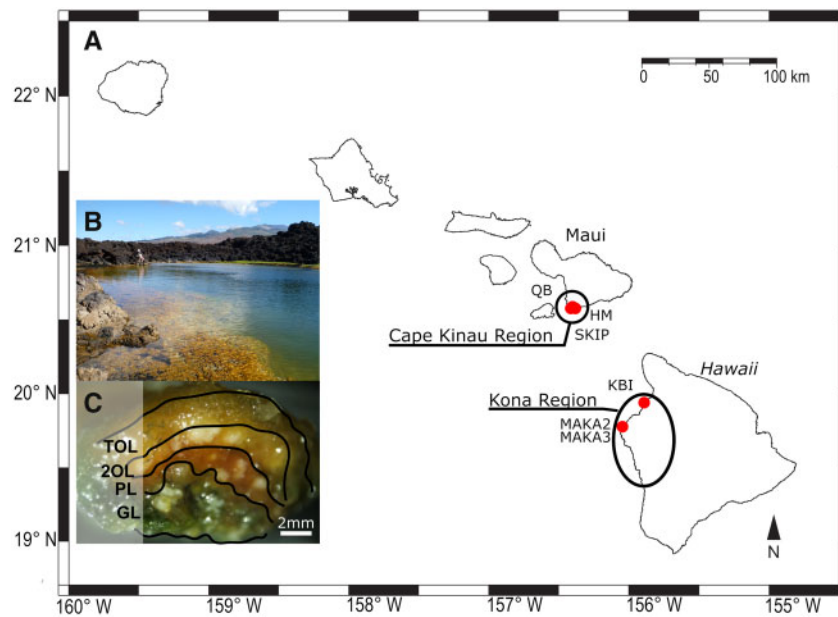
individual in both plants (e.g., Beilsmith et al. 2019) and animals (e.g., Dubé et al. 2019), including humans (e.g., Valles-Colomer et al. 2019; Zeevi et al. 2019). Such consortia, representing the collective bacteria, archaea, protist, fungi, and virus communities living in or on typically multicellular eukaryotes (Whipps et al. 1988; Prescott 2017), have become a target of intensive research over the last 15–20 years both for various organisms (reviewed by Cani 2018) and natural and man-made environments (reviewed by Gilbert and Stephens 2018; Jansson and Hofmockel 2020). However, while discussions of “genome to phenome” should be considered incomplete without accounting for microbial consortia interactions and organization, explicit studies in this context have been largely lacking.

Given current interests in the impacts of microbial consortia on the phenotypes of organisms and the environment, it should not be surprising that intense attention is being paid to further identify and develop models for such research. Efforts focused on plants and animals have gravitated to taxa that are time-efficient, cost-effective, and rich in tools and resources (reviewed in Compant et al. 2019; Douglas 2019). In contrast, the great heterogeneity among and within aquatic and terrestrial environments makes identifying a specific, and widely applicable, model more difficult. However, particular microbial consortia have garnered interest given their unusual or unique phenotype or properties. Laminated microbial mats, in which taxa exhibit vertically stratified distributions in response to environmental and/or chemical gradients, have proven to be a source of valuable biological and technological discoveries. For example, fossilized mats are central for studying the evolution of life (Walter et al. 1972; Konhauser et al. 2003; Nutman et al. 2016) and predicting how life may evolve on other planets (Walker et al. 2005) while extant ones have yielded organisms with high biotechnological value like *Thermus aquaticus* (Chien et al. 1976; Ishino and Ishino 2014). While some laminated microbial mats from particular ecosystems have received special interest or extensive study like those in the Yellowstone (USA) hot springs (Ferris and Ward 1997; Ward et al. 2006; Becraft et al. 2011) and the Hamelin Pool (AU) stromatolites (Awramik and Riding 1988; Papineau et al. 2005; Suosaari et al. 2016), many others have largely remained underexplored, including the unique, laminated orange cyanobacterial–bacterial crust endemic to particular habitats of the Hawaiian anchialine ecosystem.

The anchialine ecosystem encompasses near-shore bodies of water with fluctuating volumes and

salinities originating from subsurface connections (or influences) from fresh groundwater and seawater (Holthuis 1973; Maciolek 1983; Stock et al. 1986; Sket 1996). Habitats fitting this description are located primarily in the tropics (Holthuis 1973; Bishop et al. 2004; Krstulović et al. 2013; Menning et al. 2014) and occur within a variety of basin substrates (Holthuis 1973; Maciolek 1983; Sket 1996) where complex physical and chemical clines can occur within the water column (Stock et al. 1986; Humphreys 1999; Bishop et al. 2004). Although the anchialine ecosystem has been found to host high levels of species richness and endemism when it comes to macroorganisms (Yager 1981; Brock and Bailey-Brock 1998; Iliffe 2002; Mejía-Ortíz et al. 2007), less work has been done on the microbial communities within them (Sarbu et al. 1996; Seymour et al. 2007; Gonzalez et al. 2011; Humphreys et al. 2012).

Confined to the Cape Kinau and Kona regions of Maui and Hawaii, respectively (Wong 1975; Bailey-Brock and Brock 1993; Hoffman et al. 2018a, 2018b), the endemic laminated orange microbial crust communities from the Hawaiian anchialine ecosystem (Fig. 1) have structures comparable to stromatolites, laminated mats, or microbialites in being comprised of multiple discernable (i.e., top orange, second orange, pink, and green) layers. Each layer is primarily composed of filamentous cyanobacteria and multicellular algae colonized by diatoms and other microorganisms based on microscopy studies (Wong 1975; Bailey-Brock and Brock 1993). Structurally, stromatolites and other laminated microbial mats exhibit a common phenotype that can be subdivided from the surface to the base in a manner apparently dependent on physical and environmental properties such as the presence or absence of oxygen and/or light penetration. These stratifications generally include a surface oxic photic zone overlaying anoxic photic and anoxic aphotic zones (Bolhuis et al. 2014) with distinct taxa occupying each niche, specifically oxygenic phototrophs and aerobic heterotrophs most often near the surface while anaerobic phototrophs such as Chromatiales, Rhodobacterales, and Rhodospirillales and sulfate-reducing bacteria like Syntrophobacterales occurring in deeper zones (Paerl et al. 2000; Fourçans et al. 2004; Armitage et al. 2012; Bolhuis et al. 2014). Although representing just 10–20% of the total microbial population (Ley et al. 2006; Bolhuis et al. 2014), Cyanobacteria are major photoautotrophic contributors to primary production and vital to structural integrity of such mats as contributors of filamentous taxa as well as secretors of extracellular polymeric substances (EPS)



**Fig. 1** (A) Map depicting sampling sites of anchialine habitats on the islands of Maui and Hawaii with the regions of Cape Kinau, Maui, and Kona, Hawaii indicated (open circles). (B) Example of a Hawaiian anchialine open pool habitat (i.e., site SKIP) with the laminated orange cyanobacterial–bacterial crust communities endemic to Cape Kinau, Maui, and Kona, Hawaii. (C) Close-up of laminated orange cyanobacterial–bacterial crust with four phenotypically discernible layers. Total mat depth in the center is  $\sim 9.5$  mm. In order from crust surface to bottom: TOL, 2OL, PL, and GL.

(Bolhuis et al. 2014). This physical complexity provided by Cyanobacteria, along with the resulting oxygen and light gradients, ultimately allows for greater niche diversity (MacArthur and Levins 1967) that in turn fosters greater species richness further into the structure (Dillon et al. 2009; Armitage et al. 2012; Lindemann et al. 2013; Schneider et al. 2013).

This study drives further understanding into microbial consortia interactions and organization by reporting on the diversity, composition, and functional group partitioning among the four discernible layers within the distinct laminated, orange cyanobacterial–bacterial crust communities endemic to particular anchialine habitats within the Cape Kinau and Kona regions of Maui and Hawaii, respectively, in the Hawaiian Islands (Fig. 1). Notably, this work is the first to examine the microbial consortia composition and “genome to phenome” relationship between the different layers present in these laminated orange cyanobacterial–bacterial crusts from select habitats of the Hawaiian anchialine ecosystem. Given robust site- and island-based distinctions in the whole crust community previously reported (Hoffman et al. 2018a, 2018b), it was hypothesized such distinctions would carry through to analogous layers from different sites and islands. Furthermore, as these orange crusts exhibit a consistent physical appearance and lamination (i.e., phenotypes) between sites and islands,

the latter of which have never shared geological connections, it was hypothesized that similar stratifications of taxa (i.e., genotypes) and functional groups would be recovered between sites and islands that are reminiscent of other laminated microbial mats, with oxygenic phototrophs and aerobic heterotrophs concentrated near the surface and anaerobic and sulfur-cycling organisms toward the bottom.

## Methods

### Sites and sampling

Sampling of laminated orange cyanobacterial–bacterial crust communities occurred within an 8-day span during the spring of 2011 from six anchialine habitats on the islands of Maui and Hawaii (Fig. 1A). On Maui, sites were located along the southwest coast at Cape Hanamanioa and within the Ahihi-Kinau Natural Area Reserve at Skippy’s Pond and Queen’s Bath. Sites in Hawaii were located at Makalawena Beach ponds #2 and #3 (MAKA2 and MAK3, respectively) and Keawaiki Bay (KBI) on the west side of the island (Fig. 1A). Although all six habitats were open ponds occurring in basalt basins (e.g., Fig. 1B), they differed in ecological attributes such as degree of impact by invasive fishes, feral goats, and human visitation (pers. obs., Table 1). Whole crust samples ( $\sim 100$  g) were collected using disposable sterile spoons from three

**Table 1** Anchialine habitats from the islands of Maui and Hawaii possessing laminated orange cyanobacterial–bacterial crust communities and their corresponding categorical and continuous environmental factors

Environmental factors	Sites					
	HM	QB	SKIP	KBI	MAKA2	MAKA3
Island	Maui	Maui	Maui	Hawaii	Hawaii	Hawaii
Invasive fishes	No fish	No fish	No fish	Tilapia	Poecilids	Poecilids and Marine
Feral goats	Yes	Yes	Yes	No	Yes	Yes
Human visitation	No	No	No	Yes	Yes	Yes
Latitude (N)	20.58	20.6	20.6	19.89	19.79	19.79
Longitude (W)	156.41	156.43	156.42	155.9	156.03	156.03

HM, Hanamanioa, Maui; QB, Queen's Bath, Maui; SKIP, Skippy's Pond, Maui; KBI, Keawaiki Bay, Hawaii; MAKA2, Makalawena Beach, Pond 2, Hawaii; MAKA3, Makalawena Beach, Pond 3.

locations within each habitat and preserved in RNALater (Thermo Fisher Scientific, MA, USA) for later processing. Archival samples were also collected, preserved, and submitted to the Hawaiian Anchialine Microbial Repository with The Ocean Genome Legacy (<https://www.northeastern.edu/ogl/>) under accession numbers S23033–S23083 as described in Hoffman et al. (2018a).

### Sequence data generation

At all six sites, the laminated orange cyanobacterial–bacterial crust was composed of four distinctive layers clearly identifiable by their color and texture: a top orange layer (TOL), a second orange layer (2OL), a pink layer (PL), and a bottom green layer (GL; Fig. 1C). RNALater-preserved whole crust samples were dissected into these four phenotypic layers by carefully separating them at the boundaries between each layer's distinct color and texture differences (Fig. 1C) using sterile razors under 2–3× magnification prior to deoxyribonucleic acid (DNA) extraction utilizing MoBio PowerSoil DNA Isolation Kits (MOBIO, CA, USA) as specified in Hoffman et al. (2018a). Extracted DNA samples were shipped to the HudsonAlpha Institute for Biotechnology, Inc. Genomic Services Laboratory (Huntsville, AL, USA) where they were amplified in duplicate via the polymerase chain reaction (PCR) and sequenced on two independent runs of an Illumina HiSeq 2500 platform to minimize potential sample handling errors. Each PCR reaction utilized 20 ng of template DNA except in cases of low concentration, upon which the template volume was divided equally between the two reactions. Sequencing was performed on dual-barcoded amplicons to obtain 100 bp paired-end (PE) reads of the V6 and V9 hypervariable regions of small subunit ribosomal RNA (rRNA) genes from the microbial community under examination. Specifically, the V6

hypervariable region of the 16S rRNA gene was amplified with the *Bacteria*-specific primers 967–985 F and 1078–1061R (Gloor et al. 2010) and the V9 hypervariable region of the 18S-rRNA gene using the *Eukarya*-biased primers 1389 F and 1510 R (Amaral-Zettler et al. 2009). All raw high-throughput sequencing reads used in this study are available through the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA) database (Accession Number SAMN05833379–SAMN05833402, BioProject ID Number PRJNA325159).

### Operational taxonomic unit clustering

Alignment of forward and reverse sequencing reads and removal of primers, as well as sequences with uncalled bases, were done in PandaSeq v.2.5 (Masella et al. 2012). Reads were further filtered using the FASTQ Quality Filter in the FASTX-Toolkit v.13.2 (A. Gordon and G. J. Hannon, [http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)), with a conservative quality score cut-off >30 over at least 75% of the nucleotides in a read. Within the Quantitative Insights Into Microbial Ecology (QIIME) v.1.8 pipeline (Caporaso et al. 2010b), USEARCH61 (Edgar 2010) was employed to remove potentially chimeric sequences prior to clustering into operational taxonomic units (OTUs) via UCLUST (Edgar 2010) in the `pick_open_reference_otus.py` script. Sequences were clustered at 95% sequence similarity and 0.005% abundance using the 99% clustered GreenGenes 13.8 (DeSantis et al. 2006) and Silva 111 (Quast et al. 2012) databases as initial references for the bacterial V6 and *Eukarya*-biased V9 hypervariable regions, respectively. The 0.005% OTU abundance filter was utilized to improve clustering results (Bokulich et al. 2013) while OTUs were lumped, rather than split, using a conservative 95% sequence similarity parameter as Hoffman et al.

(2018a) previously demonstrated limited additional OTU recovery using the more stringent threshold of 97% sequence similarity. The most abundant sequence from each cluster was designated as the reference sequence and sequence clusters were assigned taxonomic identities with megaBLAST v.2.2.26 (Altschul et al. 1990) using the appropriate above-mentioned curated database at a sequence identity of  $\geq 90\%$  and *e*-value of  $1 \times 10^{-6}$ . While identifiable prokaryotic OTUs in the V9 dataset were not removed to maximize recovered taxonomic diversity (Hadziavdic et al. 2014), any OTUs failing to align with default parameters against the appropriate database by Python Nearest Alignment Space Termination (PYNAST) v.1.2.2 (Caporaso et al. 2010a) were excluded from the final OTU tables to minimize biases in community richness estimates (see below) due to unassigned identities (i.e., 2 [ $\sim 0.08\%$  sequences] V6 OTUs and 10 [0.30%] V9 OTUs).

### Analyses of layer consortia composition

Estimates of community richness for each layer were calculated in several ways, including as the number of observed OTUs as well as the Chao1 richness estimator (Chao 1984) and Shannon (Shannon 2001) and inverse Simpson (Simpson 1949) diversity indices using the package PhyloSeq v.1.10.0 (McMurdie and Holmes 2013) in the R v.3.1.13 statistical environment (R Development Core Team 2008). These four metrics (i.e., richness, Chao1 richness estimator, Shannon index, and inverse Simpson index) were used to produce rarefaction curves in R with 10 replicates at sequencing depths of 1, 10, 1000, 10,000, 20,000, and 30,000 sequences/sample to assess the thoroughness of sampling. Additionally, the four metrics were tested with one-way analysis of variances (ANOVAs) and Tukey's Honest Significant Difference (HSD) post-hoc tests to determine the influence of several moderators (i.e., layer of origin, island, feral goat visitation, invasive fish presence, and human visitation) on these metrics. ANOVAs and Tukey's HSD post hoc tests were performed using the R package agricolae v.1.2.3 (R Development Core Team 2008; De Mendiburu 2015), with  $P < 0.05$  considered significant.

The final OTU tables produced by the `pick_open_reference_otus.py` script in the QIIME pipeline were transformed to even sampling depths before the Bray–Curtis dissimilarity metric and Jaccard dissimilarity coefficient were applied in the R package vegan v.2.3.1 (R Development Core Team 2008; Oksanen et al. 2015). As common ecological

metrics, the Bray–Curtis dissimilarity metric assesses the abundance of OTUs shared between samples (Bray and Curtis 1957), while the Jaccard dissimilarity coefficient solely considers the presence or absence of OTUs when estimating the proportion of unshared taxa between samples (Jaccard 1908). These metrics were chosen due to their common application in assessing microbial communities (e.g., Yildirim et al. 2010; Sinha et al. 2017; Glasl et al. 2019) and utilization in previous studies of these habitats (Hoffman et al. 2018a, 2018b). Given this, Weighted UniFrac, which takes into consideration phylogenetic relatedness among OTUs (Lozupone et al. 2011), was deemed beyond the scope of the current analyses. Ordinations for the layer consortia were created using non-metric multidimensional scaling (NMDS) plots with 95% confidence ellipses in the R package PhyloSeq v.1.10.0 (R Development Core Team 2008; McMurdie and Holmes 2013) from the Bray–Curtis dissimilarity and Jaccard dissimilarity data matrices. Jaccard and Bray–Curtis NMDS ordinations were compared using a Procrustes test with 999 permutations in the R package vegan.

Tables of second-level (i.e., phylum and approximately phylum) and third-level (i.e., class and approximately class, both in GreenGenes and Silva, respectively) clades were created using the `summarize_taxa_through_plots.py` script in the QIIME v.1.8 pipeline (Caporaso et al. 2010b). Second-level clades and proteobacterial classes with relative abundances that varied between layers were identified with one-way ANOVAs with Tukey–Kramer post-hoc tests and Benjamini–Hochberg FDR corrections for multiple comparisons using STAMP v.2.1.3 (Parks et al. 2014). Additionally, correlation networks between second-level clades and proteobacterial classes with strong correlations in co-occurrence (Pearson's  $P > 0.7$  or  $< -0.7$ ,  $P < 0.001$ ) were produced using the R packages `igraph` v.1.2.4 (Csardi and Nepusz 2006) and `Hmisc` 4.2.0 (Harrell and Dupont 2019) and visualized with Cytoscape utilizing the degree sorted circle layout following the methods described by Ju et al. (2014) and Fisher et al. (2019). Identified clades with varying relative abundances and/or identified as part of a co-occurrence network were then classified by performing a literature search for the closest characterized relatives into one of the following metabolic groups: aerobic chemoautotroph, aerobic heterotroph, anaerobic chemoautotroph, anaerobic heterotroph, anaerobic photoautotroph, anaerobic photoheterotroph, fermenter, oxygenic photoautotroph, and parasite. All R code, QIIME scripts, and

other commands utilized here can be downloaded from <http://webhome.auburn.edu/~santosr/sequencedatasets.htm>.

## Results

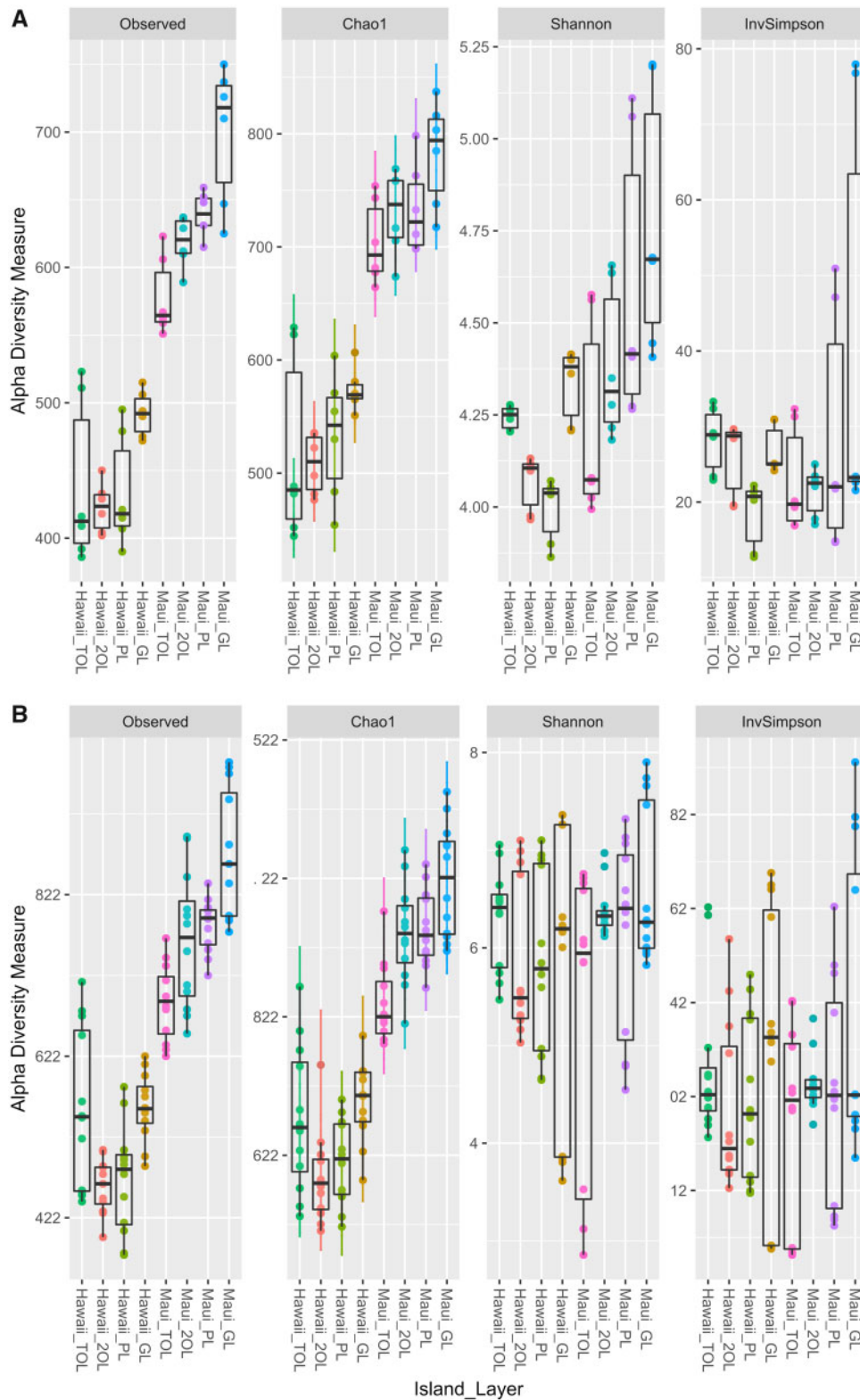
### Data generation and OTU clustering

Each of the four phenotypically distinct layers present in the laminated orange cyanobacterial–bacterial crusts from each site was successfully amplified in duplicate for the *Bacteria*-specific V6 and *Eukarya*-biased V9 hypervariable regions via PCR and sequenced with two independent sequencing runs, resulting in 96 samples originating from the six habitats. A total of 15,869,028 demultiplexed V6 Illumina reads were obtained in each PE direction, averaging ( $\bar{x}$ ) 82,651 reads per sample. For V9, a total of 9,578,684 demultiplexed V9 reads were returned ( $\bar{x}$  = 49,888 reads per sample). Following alignment, quality-filtering, removal of potential chimeric sequences, and abundance filtering, 1,674,332 bacterial V6 reads (an 89.4% reduction overall,  $\bar{x}$  = 17,440 reads per sample) belonging to 25 bacterial second-level clades and 1,281,037 eukaryote-biased V9 reads (an 86.6% reduction overall,  $\bar{x}$  = 13,344 reads per sample) from 14 bacterial and 12 eukaryotic second-level clades remained. Such stringent filtering parameters were utilized to reduce the noise-to-signal ratio as well as mitigate potential taxonomic assignment issues due to the short-read lengths (see below) produced by Illumina instruments. Reads from the V6 and V9 regions ranged in length from 69 to 80 bp ( $\bar{x}$  = 75 bp) and 65 to 162 bp ( $\bar{x}$  = 120 bp), respectively. PYNAST identified 2 (~0.08% sequences) V6 and 10 (~0.30% sequences) V9 OTUs as failing to align, and these were subsequently removed from the final datasets. Of the 1222 V6 OTUs in the final dataset, 132 (181,034 reads, 11.5% of total) were not assigned taxonomic identities using the GreenGenes 13.8 database (DeSantis et al. 2006) while 12 of the final 1083 V9 OTUs (10,283 reads, 0.85% of total) were not assigned identities using the Silva 111 database (Quast et al. 2012). When compared to NCBI's GenBank (Benson et al. 2009) using BLASTN v.2.3.0 (Altschul et al. 1990), the 132 unassigned V6 OTUs revealed affiliations primarily with uncultured members of the Acidobacteria, Actinobacteria, Alphaproteobacteria, Bacteroidetes, candidate division NC10, Chlamydiae, Chlorobi, Chloroflexi, Cyanobacteria, Deinococcus–Thermus, Deltaproteobacteria, Firmicutes, Mollicutes, Planctomycetes, and Verrucomicrobia at  $e$ -values  $\leq 1 \times 10^{-10}$  (Supplementary Table S1). For the 12

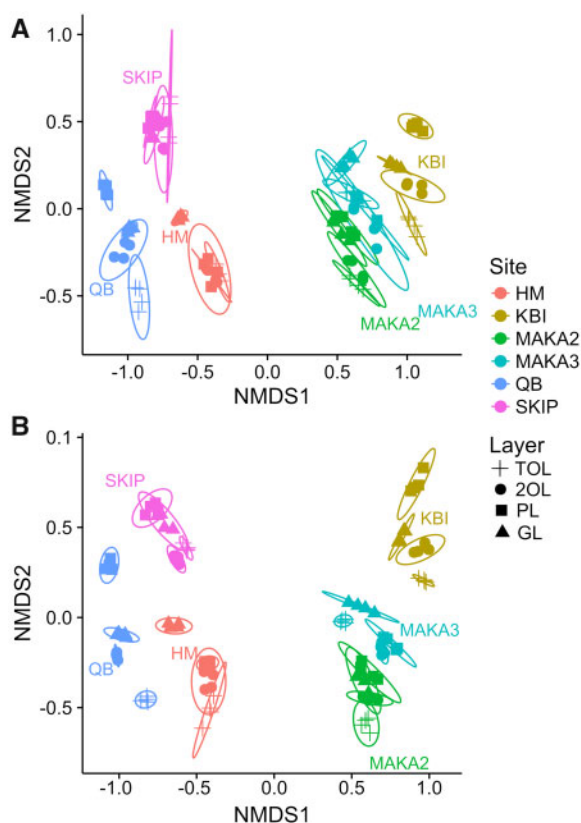
V9 OTUs which were not taxonomically assigned, these were most closely matched to members of Ciliophora, Amoebozoa, Bacillariophyceae, Alphaproteobacteria, and uncultured Archaea in the GenBank repository, also at  $e$ -values  $\leq 2 \times 10^{-16}$  (Supplementary Table S1). Seven V6 OTUs (0.34% of total) were identified as most likely originating from chloroplasts of eukaryotic haptophytes and/or stramenopiles.

### Analyses of consortia composition from phenotypically analogous layers

Sequence reads from phenotypically analogous (i.e., same position and color) layers of the laminated orange cyanobacterial–bacterial crust from the same site were combined for most downstream analyses because they were most similar to each other regardless of which PCR or sequencing run they were generated from (not shown). At 30,000 reads per sample, OTU richness and the Chao1 richness estimator did not saturate for either the V6 or V9 datasets while both the Shannon and inverse Simpson diversity indices appeared saturated at sampling depths  $\leq 10,000$  reads per sample (Supplementary Fig. S1). Based on ANOVAs, comparison of alpha diversity metrics between V6 samples grouped by specific crust layer identified the GL as having significantly greater OTU richness and Chao1 estimated richness compared to the other three layers ( $F_{3,88} = 7.353$ ,  $P < 0.001$  and  $F_{3,88} = 5.282$ ,  $P = 0.002$ , respectively; Fig. 2A). Furthermore, Shannon diversity was also significantly greater in the GL than the 2OL and PL ( $F_{3,88} = 3.584$ ,  $P = 0.017$ ). However, no significant differences in inverse Simpson diversities were found among these layers. For the *Eukarya*-biased V9 data, significantly greater OTU richness and Chao1 estimated richness were also recovered from the GL vs. other layers ( $F_{3,88} = 12.92$ ,  $P \ll 0.001$  and  $F_{3,88} = 10.487$ ,  $P \ll 0.001$ , respectively), but no effect of layer was evident from the V9 Shannon or inverse Simpson diversity indices (Fig. 2B). Additionally, significantly greater V6 and V9 OTU richness was observed in samples originating on Maui than in those from Hawaii ( $F_{1,88} = 124.904$ ,  $P \ll 0.001$  and  $F_{1,88} = 401.45$ ,  $P \ll 0.001$ , respectively) (Fig. 2A). The Chao1 richness estimator also indicated significantly greater V6 and V9 richness in samples from Maui than Hawaii ( $F_{1,88} = 160.909$ ,  $P \ll 0.001$  and  $F_{1,88} = 411.890$ ,  $P \ll 0.001$ ). While the V6 samples from Maui also had a significantly greater Shannon diversity than those from Hawaii ( $F_{1,88} = 33.759$ ,  $P \ll 0.001$ ), Shannon diversities from V9 samples were not correlated by island. Island of origin also



**Fig. 2** Diversity estimates, specifically number of observed OTUs, Chao1 richness estimator, Shannon diversity, and inverse Simpson diversity, of the *Bacteria*-specific V6 hypervariable region of the *16S-rRNA* gene (**A**), and the *Eukarya*-biased V9 hypervariable region of the *18S-rRNA* gene (**B**). Samples are colored by island of origin and layer within the laminated orange cyanobacterial–bacterial crust. In order from crust surface to bottom: TOL, 2OL, PL, and GL.



**Fig. 3** NMDS ordination using the Bray–Curtis dissimilarity index of samples grouped by laminated orange cyanobacterial–bacterial crust layer from their anchialine habitat of origin. In order from crust surface to bottom: TOL, 2OL, PL, and GL. (A) Samples generated using the *Bacteria*-specific V6 hypervariable region of the 16S-rRNA gene (stress = 0.0963). (B) Samples generated using the *Eukarya*-biased V9 hypervariable region of the 18S-rRNA gene (stress = 0.0811).

did not have a significant impact on inverse Simpson diversity for either the bacterial V6 or *Eukarya*-biased V9 datasets (Fig. 2). Finally, significantly greater V6 OTU richness and Chao1 estimated richness ( $F_{1,94} = 17.7$ ,  $P < 0.01$ ; V9:  $F_{1,94} = 21.83$ ,  $P < 0.01$  and  $F_{1,94} = 22.71$ ,  $P < 0.01$ ; V9:  $F_{1,94} = 24.05$ ,  $P < 0.01$ , respectively) and lower V9 diversity (Shannon  $F_{1,94} = 18.09$ ,  $P < 0.01$  and inverse Simpson  $F_{1,94} = 24.62$ ,  $P < 0.01$ ) were correlated with feral goats visitation to some of the sampled habitats. However, visitation by feral goats was not correlated with Shannon or inverse Simpson diversity metrics for the V6 data.

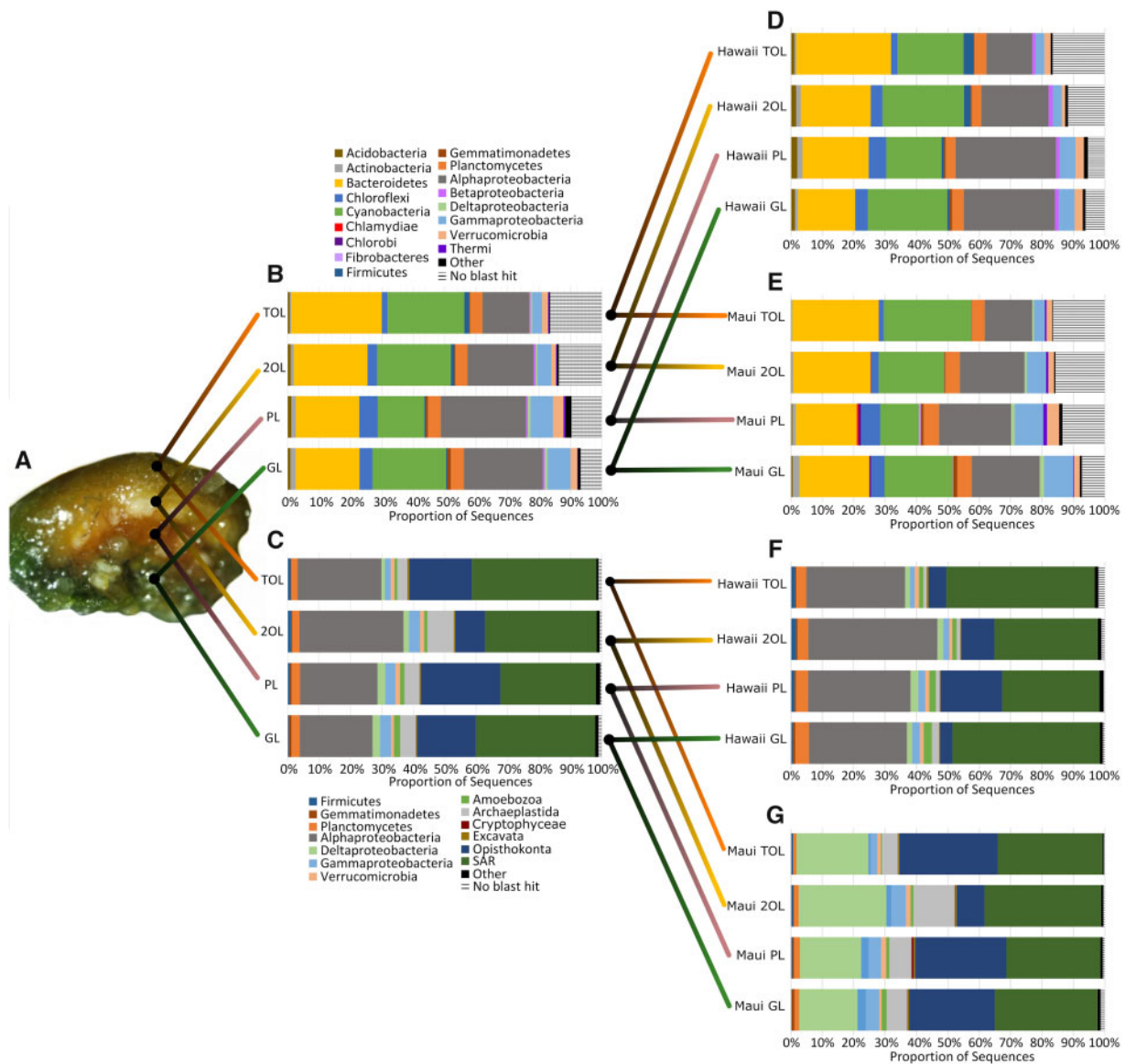
Minimal differences were observed between NMDS ordinations generated from the abundance-based Bray–Curtis dissimilarity metric or the binary Jaccard dissimilarity coefficient (V6: Procrustes  $m^2 = 0.04364$ ,  $P = 0.001$ ; V9:  $m^2 = 2.626 \times 10^{-9}$ ,  $P = 0.001$ ). Thus, only the Bray–Curtis ordinations for the V6 (Fig. 3A) and the V9 (Fig. 3B) data are

presented and discussed. Samples from crust layers primarily grouped by island, such that those from all Maui sites clustered separately from all Hawaii sites (Fig. 3). Within these island-specific clusters, samples were further grouped into site-specific clusters that included all four layers from their respective laminated orange cyanobacterial–bacterial crust. Notably, samples from phenotypically analogous layers across different islands or site did not cluster in any context (Fig. 3A and B).

Overall, genetic identification of bacterial phyla, proteobacterial classes (due to the great metabolic and physiological diversity within Proteobacteria), and eukaryotic second-level (approximately phyla) clades between the four layers revealed all possessed the same clades, but in different relative abundances (Fig. 4B and C, Supplementary Table S2). For example, while Cyanobacteria were present in all four layers, cyanobacterial OTUs were most abundant in the two TOLs of the crust as well as in the bottom-most GL (Fig. 4B). However, when the two islands were considered separately, Cyanobacteria were most abundant in the TOL and GL of Maui samples but most abundant in the GL and 2OL of Hawaii samples (Fig. 4D and E). Additionally, though Archaeplastida were ubiquitous in all four layers, OTUs belonging to Archaeplastida were most abundant in the 2OL and least abundant in the TOL of all samples and those from Maui (Fig. 4C and G) but increased in abundance with increasing depth in Hawaii samples (Fig. 4F). Similarly, Bacteroidetes OTUs were most abundant in the TOL and decreased in abundance with increasing depth into the crust (Fig. 4B), a pattern that was observed in Hawaii samples and generally in the Maui samples, though Bacteroidetes OTUs were least abundant in PL samples from Maui (Fig. 4D and E). Whereas gammaproteobacterial OTUs also increased in abundance with increasing depth in the crust in both V6 and V9 datasets and for both islands (Fig. 4B–G), differing patterns were observed for OTUs belonging to Alphaproteobacteria. Specifically, alphaproteobacterial OTUs detected by the *Bacteria*-specific V6 primers increased in abundance with increasing crust depth (Fig. 4B, D, and E) while those detected with the *Eukarya*-biased V9 primers were primarily identified in samples from Hawaii and were most abundant in the 2OL (Fig. 4C, F, and G).

Networks of bacterial phyla, eukaryotic phyla, and proteobacterial classes failed to identify any strong negative correlations in relative abundance between identified taxa, while positive correlations ranged from Pearson's  $P = 0.703$  (Deltaproteobacteria and Armatimonadetes, Supplementary Table S3) to





**Fig. 4** Relative abundance of taxa identified in samples grouped by layer of origin within the laminated orange cyanobacterial–bacterial crust. Taxa present at <0.5% relative abundance in every layer were summarized in the artificial group “Other.” (A) In order from crust surface to bottom: TOL, 2OL, PL, and GL. (B) Bacterial phyla and proteobacterial classes identified in the final OTU table by the *Bacteria*-specific V6 hypervariable region of the *16S-rRNA* gene using the GreenGenes 13.8 database, and for individual crust layers from the islands of Hawaii (D) and Maui (E). The artificial “Other” category included Armatimonadetes, BRC1, Chlamydiae, Chlorobi, FCPU426, Fibrobacteres, GOUTA4, Lentisphaerae, NKB19, Nitrospirae, OP3, OP8, Other Proteobacteria, Zetaproteobacteria, Spirochaetes, and WS6. (C) Approximately phyla-level bacterial clades, proteobacterial classes, and approximately phyla-level eukaryotic clades identified in the final OTU table by the *Eukarya*-biased V9 hypervariable region of the *18S-rRNA* gene using the Silva 111 database, and for individual crust layers from the islands of Hawaii (F) and Maui (G). The artificial “Other” category included Acidobacteria, Actinobacteria, Bacteroidetes, Candidate division OD1, Chloroflexi, Fibrobacteres, Fusobacteria, Lentisphaerae, Nitrospirae, Betaproteobacteria, Other Proteobacteria, Cryptophyceae, Haptophyta, Incertae\_Sedis (Protista), and Kathablepharidae.

Pearson’s  $P=0.966$  (Cryptophyceae and Nitrospirae, [Supplementary Table S3](#)). Samples collected on Maui generated two large networks with 17 or 10 clades and many links in each along with two networks of two clades ([Supplementary Table S3](#)), while samples from Hawaii generated seven networks consisting of two to eight clades in each ([Supplementary Table](#)

[S3](#)). Furthermore, the networks from Hawaii primarily included clades with greatest abundances in the bottom-most pink and GLs of the crust, with the V6 networks generally separated into heterotrophic and photoautotrophic OTUs ([Supplementary Table S3](#)) while the V9 networks generally separated into a network of anaerobic OTUs, one of heterotrophic

OTUs, and others with pairs of autotrophic OTUs (Supplementary Table S3). Taxa in networks from the *Bacteria*-specific V6 dataset were generally found in greatest relative abundance in the same layer as compared to networks from the *Eukarya*-biased V9 dataset (Figs. 4 and 5).

As previously noted, OTUs belonging to oxygenic photoautotrophs like the Cyanobacteria exhibited greater relative abundance in the bacterial V6 dataset within the two orange layers along with the bottom GL (Fig. 5A). An increased abundance in putative aerobic heterotrophic and anaerobic photoautotrophic OTUs in the V9 *Eukarya*-biased dataset (Fig. 5B) also further characterized the second, lower orange layer. Here, these potential aerobic heterotrophic OTUs were primarily members of the Alpha- and Gammaproteobacteria, Gemmatimonadetes, and Verrucomicrobia, and possible anaerobic photoautotrophs belonging to the gammaproteobacterial Chromatiales. An increased abundance of potential fermentative OTUs characterized the PL in both datasets (Fig. 5A and B), as well as an increased abundance of putative aerobic heterotrophic and anaerobic photoautotrophic OTUs in the bacterial V6 data (Fig. 5A). Putative fermentative OTUs in both datasets belonged to the Acidobacteria, Chloroflexi, Fibrobacteres, Firmicutes, Fusobacteria, Lentisphaerae, Planctomycetes, Beta-, Delta-, and Gammaproteobacteria, Spirochaetes, and Verrucomicrobia. OTUs assumed to be aerobic heterotrophs in the *Bacteria* V6 belonged to the Acidobacteria, Actinobacteria, Armatimonadetes, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Planctomycetes, Alpha- and Gammaproteobacteria, Deinococcus–Thermus, and Verrucomicrobia. Possibly anaerobic photoautotrophic OTUs in the V6 data belonged to Chloroflexi and the gammaproteobacterial Chromatiales. Finally, the GL was taxonomically rich, with the greatest relative abundance of putative anaerobic photoheterotrophic OTUs, members of the alphaproteobacterial Rhodospirillales (Fig. 5B), together with many OTUs potentially utilizing fermentative and oxygenic photoautotrophic metabolisms (Fig. 5A and B).

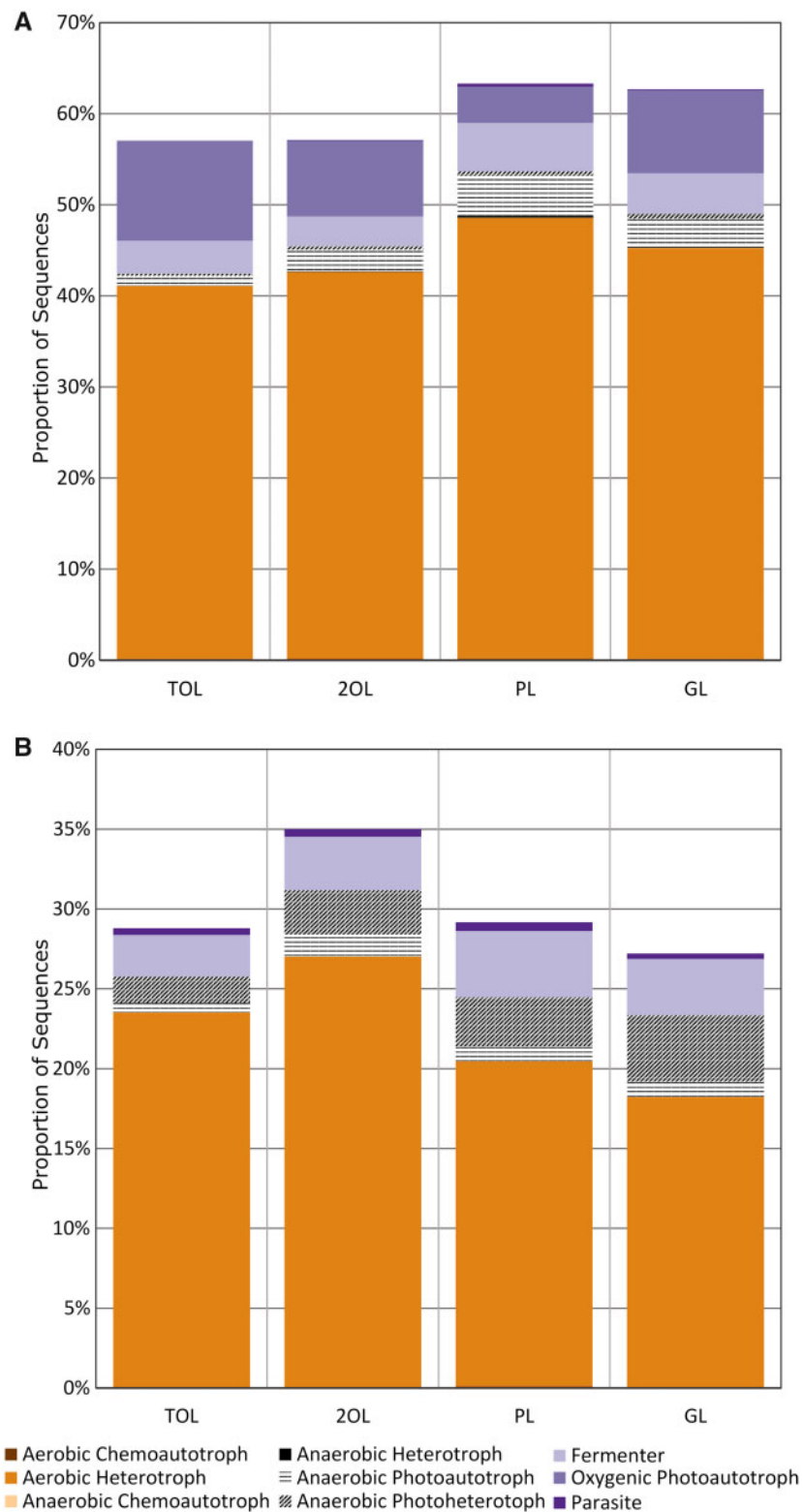
## Discussion

### Comparison of layer consortia diversity

This is the first work to specifically examine the individually discernable layers comprising the laminated orange cyanobacterial–bacterial crusts endemic to particular anchialine habitats of the Cape Kinau and Kona coast regions of Maui and

Hawaii, respectively. Such work is important as it offers the opportunity to begin developing a foundation for testing specific hypotheses on the ecological, biogeochemical, and metabolic roles microbes play in such communities. Rarefaction analyses suggest that while sampling depth may have failed to reach saturation in estimating microbial richness in these consortia, it was sufficient for estimating diversity (Supplementary Fig. S1). The recovery of greater OTU richness, Chao1 estimated richness, and V6 Shannon diversity at the bottom of these crusts (Fig. 2) is consistent with previous work from other laminated microbial mats, including those from hypersaline environments (Villanueva et al. 2007; Dillon et al. 2009; Lindemann et al. 2013; Schneider et al. 2013) and salt marshes (Armitage et al. 2012). Due to their seawater and freshwater influences, increasing depth results in more dynamic environmental conditions in anchialine habitats than at the surface (Holthuis 1973; Sket 1996; Humphreys 1999; Calderón-Gutiérrez et al. 2018), thus allowing different taxa to occupy niches meeting their metabolic needs. Additionally, microbial mats and crusts can create and maintain innumerable microniches that can compress high levels of microbial diversity into a relatively small spatial area (Macarthur and Levins 1967; Bolhuis et al. 2014).

Greater OTU richness, Chao1 estimated richness, and V6 Shannon diversity were also recovered from samples collected on Maui compared to Hawaii (Fig. 2). While notable, the apparent impacts of previously-quantified differences in water chemistry as well as invasive fishes, feral goats and human visitation on microbial richness may partially explain these island-based differences (Hoffman et al. 2018a). Specifically, it was previously demonstrated that salinity correlated with island and microbial diversity indices, such that anchialine habitats on Maui have greater salinity and diversity estimates than those on Hawaii (Hoffman et al. 2018a). Furthermore, all the sampled habitats on Hawaii were both human-visited and invaded by some combination of tilapia, poeciliid guppies, and/or marine fish species. Such invasive species can reduce microbial richness (and diversity in the case of poeciliid guppies) of Hawaiian anchialine habitats. Additionally, greater OTU richness, as well as decreased V9 diversity, was correlated with feral goats watering from the sampled habitats on Maui and the two MAKAs sites in Hawaii. In this case, decreased V9 diversity may be ultimately linked to shifts and declines in native vegetation impacting the overall watershed, effects documented on many Pacific



**Fig. 5** Relative abundance of functional groups identified in samples grouped by layer of origin within the laminated orange cyanobacterial–bacterial crust. In order from crust surface to bottom: TOL, 2OL, PL, and GL. **(A)** Bacterial functional groups identified by the *Bacteria*-specific V6 hypervariable region of the *16S-rRNA* gene using the GreenGenes 13.8 database. **(B)** Bacterial functional groups identified by the *Eukarya*-biased V9 hypervariable region of the *18S-rRNA* gene using the Silva 111 database.

islands with feral goat populations (Chynoweth et al. 2013). It is also possible that goats introduced transient microbes into the anchialine habitats they visited that had little impact on bacterial diversity (i.e., rare taxa), but subsequently inflated the observed richness estimates.

Although OTU richness was influenced by crust layer, NMDS ordinations using either the binary Jaccard dissimilarity coefficient or abundance-based Bray–Curtis dissimilarity revealed samples grouping together by site, not layer (Fig. 3). Site-specific microbial communities were also described previously for whole samples from these same habitats. (Hoffman et al. 2018a, 2018b). Thus, rather than phenotypically analogous layers (i.e., all PLs) containing more similar microbial consortia across the examined anchialine habitats, each habitat apparently possesses genetically unique microbial consortia regardless of layer. Additionally, both datasets included ~9 ubiquitous OTUs which were present in all 96 samples and ~9 rare OTUs that were present in only a few samples. Given this, the lack of clustering difference between the binary Jaccard and abundance-based Bray–Curtis ordinations, the limited number of ubiquitous and rare OTUs, and the similarity in relative phyla compositions suggests differences between islands, sites, and samples were not due to abundance-based differences of shared OTUs but instead by both different OTU memberships and differing abundances of shared OTUs. Taken together, the distinct community phenotype observed among laminated orange microbial crust communities endemic to particular anchialine habitats on Maui and Hawaii may be considered a plastic trait, where assembly has occurred independently into layered crust structures whose appearances are phenotypically convergent to the distinct consortia present in other crust-containing anchialine habitats, with the likely driver being gross similarity in environmental conditions among them (Havird et al. 2013; Hoffman et al. 2018a, 2018b).

#### Comparison of layer consortia composition and their potential physiological and metabolic functions

Taxonomic identity can serve as a statistically significant surrogate for inferring the functional potential of individual members in microbial communities (Royalty and Steen 2019), ultimately contributing to better understanding relationships between biodiversity and ecosystem functioning (Krause et al. 2014). When applied to the laminated orange cyanobacterial–bacterial crusts examined here, oxygenic phototrophs in the form of Cyanobacteria were

present in all four layers but had their greatest relative abundances in those both at the top and bottom (Figs. 4B and 5A). Specifically, OTUs identified as *Halomicronema*, a nonheterocystous filamentous cyanobacterium previously observed in laminated mats (Fourçans et al. 2004; Allen et al. 2010), was ubiquitous to every sample. While the abundance of Cyanobacteria at, and near, the surface of these crusts is consistent with many laminated cyanobacterial–bacterial mats (Paerl et al. 2000; Fourçans et al. 2004; Armitage et al. 2012; Bühring et al. 2014), the finding of an almost equal abundance of Cyanobacteria at the bottom of crusts like these has not been previously documented. Some prior studies examining these laminated orange cyanobacterial–bacterial crusts compared them to stromatolites (Wong 1975; Bailey-Brock and Brock 1993); however, stromatolites along with most laminated bacterial mats exhibit strict oxygen gradients with anoxic conditions at their bottom (Ley et al. 2006; Lindemann et al. 2013; Schneider et al. 2013; Bolhuis et al. 2014), in clear contrast to the microbial community composition reported here.

The high abundance of Cyanobacterial OTUs at the bottom of these Hawaiian laminated orange cyanobacterial–bacterial crusts relative to previous descriptions of stromatolites and more typical laminated bacterial mats might be explained by the fact that these crusts are unusual in that they only loosely adhere to the benthos and often grow in plates and hollow protuberances unattached to the underlying substrate (pers. obs.; Bailey-Brock and Brock 1993). Furthermore, seawater typically enters an anchialine habitat through the porous basin, thus providing oxygenation to the bottom of the mat as well as potentially facilitating circulation of the pockets of water between the crust structure and the substrate and through breaks in the crust itself. Even with the potential of imperfect dissection of layers prior to DNA extraction, the abundance of Cyanobacteria and algae detected in the bottom GL is unlikely an artifact since previous light microscopy identified the source of the bottom layer's green color as coming from taxa belonging to these groups (Wong 1975). Thus, these crusts seem to be unique in that both their top and bottom are apparently well-oxygenated, in contradiction to one of our initially proposed hypotheses.

Although specific OTUs varied significantly when comparing analogous layers across sites, similar functional groups were enriched in particular layers from all sites. For example, oxic zones at both the crust surface and bottom of all sites likely resulted in increased abundances of putative anaerobic taxa in

the middle layers (the 2OL and the PL; Fig. 5). Heterotrophy also appeared concentrated in these middle two layers while chemoautotrophy to the lower three layers, with putative chemoautotrophs primarily concentrated to an apparently appropriate niche in the PL (Fig. 5, Supplementary Table S2). The strong positive correlations between taxa primarily enriched in the pink and GLs (Supplementary Table S3) suggest that such anoxygenic niches are facilitated and maintained by the metabolisms of surrounding taxa as previously hypothesized (Schneider et al. 2013). In contrast, a laminated mat community from Kiritimati Atoll exhibited increased abundance of aerobic heterotrophs in the oxic layers near the top that decreased as fermenters increased in the transitional center layers, with relatively low abundances of either at the bottom anoxic layers where chemoautotrophs were localized (Schneider et al. 2013). Thus, whereas more typical laminated mats exhibit vertical gradients that stratify anaerobic metabolic niches across the lower layers of the mat, anaerobic conditions in the laminated orange cyanobacterial–bacterial crusts examined here appear concentrated to the center of their physical structure.

Anoxic niches were not solely restricted to the interior two layers, however, as known obligate anaerobes, such as the sulfate-reducing Deltaproteobacteria (Desulfobacterales and Desulfovibrionales), were found throughout the crusts (Fig. 4). The existence of anoxic niches, even at the surface of microbial mats, has been previously observed (Wong et al. 2015, 2018) and may be influential in the observed precipitation of calcium carbonate by the cyanobacterial members of these Hawaiian anchialine laminated orange cyanobacterial–bacterial crust consortia (Wong 1975; Bailey-Brock and Brock 1993). Microbial mat precipitation of calcium carbonate may be favored by increased calcium and inorganic carbon concentrations, produced by the degradation of EPS by heterotrophs, along with increased alkalinity resulting from sulfate-reduction by sulfate-reducing bacteria (Dupraz and Visscher 2005). Significant relative abundances of heterotrophs capable of breaking down high molecular weight macromolecules, such as those composing the EPS secreted by Cyanobacteria, were identified in the crusts (Figs. 4 and 6) in combination with a strong positive association between Bacteroidetes and Cyanobacteria (Supplementary Table S3). A similarly strong positive association between these two bacterial groups was observed and hypothesized as indicating the link between production and degradation of EPS in the

lithifying hypersaline microbial mats of Shark Bay, Western Australia (Wong et al. 2015). Taken together, the presence of sulfate-reducing bacteria with EPS production and degradation by Cyanobacteria and heterotrophs like Bacteroidetes, respectively, may together be responsible for the calcium carbonate precipitation known from these particular microbial crusts.

## Conclusions

Our study is the first to examine the specific consortia present in the phenotypically discernable layers of the laminated, orange cyanobacterial–bacterial crust endemic to particular habitats of the Hawaiian anchialine ecosystem. Overall, distinct microbial OTUs (i.e., genotypes) appear to have independently assembled multiple times toward a common physical appearance and lamination (i.e., phenotype). This represents an apparent case of convergent evolution, with gross similarity in environmental conditions having resulted in some anchialine habitats on Maui and Hawaii commonly arriving at a laminated orange microbial crust community despite the fact that the specific OTU makeup of the crust layers varies significantly among these habitats. Importantly, this provides a microbiome perspective on a common observation in studies of both plants and animals, namely that environment plays a large and significant role in determining the relationship(s) between genome and phenome. While “phenotype” in this study was largely limited to gross visual appearance, the recovery of similar functional groups in analogous layers (e.g., oxygenic groups in top and bottom layers) from different habitats implies assembly of distinct OTUs to generate convergent metabolic phenotypes, which represents an arena ripe for future research. Given the independent origins of the laminated, orange cyanobacterial–bacterial crust consortia on Maui and Hawaii, the number of distinct anchialine habitats across the islands, as well as our results here, microbial communities of the anchialine ecosystem is an attractive model for extending genotype to phenotype thinking to microbial consortia.

## Acknowledgments

We thank K. L. Kim and R. A. Kinzie III for generous help and support associated with fieldwork. P. M. Brannock and D. S. Waits graciously assisted with processing of sequence data. We are indebted to P. M. Brannock and M. Newman for providing helpful comments and feedback during the writing process. M. Ramsey assisted and provided comments

and photos, regarding work at the WC site. Site access and collections were conducted under the following scientific permits: State of Hawaii Native Invertebrate Research Permit # FHM10-232 and MAKA: Kamehameha Schools Permit # 4803. The experiments conducted in this study comply with current laws of the United States and the State of Hawaii. This represents contributions No. 199 and No. 100 to the Auburn University (AU) Marine Biology Program and Molette Biology Laboratory for Environmental and Climate Change Studies, respectively.

## Funding

This work was supported by the National Science Foundation [DEB #0949855 to S.R.S].

## Supplementary data

[Supplementary data](#) available at *ICB* online.

## References

- Allen MA, Neilan BA, Burns BP, Jahnke LL, Summons RE. 2010. Lipid biomarkers in Hamelin Pool microbial mats and stromatolites. *Org Geochem* 41:1207–18.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol* 215:403–10.
- Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM. 2009. A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. *PLoS One* 4:e6372.
- Armitage DW, Gallagher KL, Youngblut ND, Buckley DH, Zinder SH. 2012. Millimeter-scale patterns of phylogenetic and trait diversity in a salt marsh microbial mat. *Front Microbiol* 3:293.
- Awramik SM, Riding R. 1988. Role of algal eukaryotes in subtidal columnar stromatolite formation. *Proc Natl Acad Sci U S A* 85:1327–9.
- Bailey-Brock JH, Brock RE. 1993. Feeding, reproduction, and sense organs of the Hawaiian anchialine shrimp *Halocaridina rubra* (Atyidae). *Pac Sci* 47:338–55.
- Becraft ED, Cohan FM, Kühl M, Jensen SI, Ward DM. 2011. Fine-scale distribution patterns of *Synechococcus* ecological diversity in microbial mats of Mushroom Spring, Yellowstone National Park. *Appl Environ Microbiol* 77:7689–97.
- Beilsmith K, Thoen MP, Brachi B, Gloss AD, Khan MH, Bergelson J. 2019. Genome-wide association studies on the phyllosphere microbiome: embracing complexity in host–microbe interactions. *Plant J* 97:164–81.
- Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2009. GenBank. *Nucleic Acids Res* 37:D26–31.
- Bishop RE, Kakuk B, Torres JJ. 2004. Life in the hypoxic and anoxic zones: metabolism and proximate composition of Caribbean troglobitic crustaceans with observations on the water chemistry of two anchialine caves. *J Crustac Biol* 24:379–92.
- Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI, Knight R, Mills DA, Caporaso JG. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* 10:57–9.
- Bolhuis H, Cretoiu MS, Stal LJ. 2014. Molecular ecology of microbial mats. *FEMS Microbiol Ecol* 90:335–50.
- Bray JR, Curtis JT. 1957. An ordination of the upland forest communities of southern Wisconsin. *Ecol Monogr* 27:325–49.
- Brock RE, Bailey-Brock JH. 1998. An unique anchialine pool in the Hawaiian Islands. *Int Rev Hydrobiol* 83:65–75.
- Bühning SI, Kamp A, Wörmer L, Ho S, Hinrichs K-U. 2014. Functional structure of laminated microbial sediments from a supratidal sandy beach of the German Wadden Sea (St. Peter-Ording). *J Sea Res* 85:463–73.
- Calderón-Gutiérrez F, Sánchez-Ortiz CA, Huato-Soberanis L. 2018. Ecological patterns in anchialine caves. *PLoS One* 13:e0202909.
- Cani PD. 2018. Human gut microbiome: hopes, threats and promises. *Gut* 67:1716–25.
- Caporaso JG, Bittinger K, Bushman FD, DeSantis TZ, Andersen GL, Knight R. 2010a. PyNAST: a flexible tool for aligning sequences to a template alignment. *Bioinformatics* 26:266–7.
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, et al. 2010b. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7:335–6.
- Chao A. 1984. Nonparametric estimation of the number of classes in a population. *Scand J Stat* 11:265–70.
- Chien A, Edgar DB, Trela JM. 1976. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *J Bacteriol* 127:1550–7.
- Chynoweth MW, Litton CM, Lepczyk CA, Hess SC, Cordell S. 2013. Biology and impacts of Pacific island invasive species. 9. *Capra hircus*, the feral goat (Mammalia: Bovidae). *Pac Sci* 67:141–57.
- Compant S, Samad A, Faist H, Sessitsch A. 2019. A review on the plant microbiome: ecology, functions and emerging trends in microbial application. *J Adv Res* 19:29–37.
- Csardi G, Nepusz T. 2006. The igraph software package for complex network research. *InterJournal Complex Syst* 1695:1–9.
- De Mendiburu F. 2015. *Agricolae*: statistical procedures for agricultural research. R Package Version 12-3 1–6 (<https://cran.r-project.org/web/packages/agricolae/index.html>).
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* 72:5069–72.
- Dillon JG, Miller S, Bebout B, Hullar M, Pinel N, Stahl DA. 2009. Spatial and temporal variability in a stratified hypersaline microbial mat community. *FEMS Microbiol Ecol* 68:46–58.
- Douglas AE. 2019. Simple animal models for microbiome research. *Nat Rev Microbiol* 17:764–12.
- Dubé CE, Ky CL, Planes S. 2019. Microbiome of the black-lipped pearl oyster *Pinctada margaritifera*, a multi-tissue

- description with functional profiling. *Front Microbiol* 10:1548.
- Dupraz C, Visscher PT. 2005. Microbial lithification in marine stromatolites and hypersaline mats. *Trends Microbiol* 13:429–38.
- Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–1.
- Ferris MJ, Ward DM. 1997. Seasonal distributions of dominant 16S rRNA-defined populations in a hot spring microbial mat examined by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* 63:1375–81.
- Fisher A, Wangpraseurt D, Larkum AWD, Johnson M, Kühl M, Chen M, Wong HL, Burns BP. 2019. Correlation of bio-optical properties with photosynthetic pigment and microorganism distribution in microbial mats from Hamelin Pool, Australia. *FEMS Microbiol Ecol* 95: 1–13.
- Fourçans A, de Oteyza TG, Wieland A, Solé A, Diestra E, Bleijswijk J, Grimalt JO, Kühl M, Esteve I, Muyzer G, et al. 2004. Characterization of functional bacterial groups in a hypersaline microbial mat community (Salins-de-Giraud, Camargue, France). *FEMS Microbiol Ecol* 51:55–70.
- Gilbert JA, Stephens B. 2018. Microbiology of the built environment. *Nat Rev Microbiol* 16:661–70.
- Glasl B, Bourne DG, Frade PR, Thomas T, Schaffelke B, Webster NS. 2019. Microbial indicators of environmental perturbations in coral reef ecosystems. *Microbiome* 7:94.
- Gloor GB, Hummelen R, Macklaim JM, Dickson RJ, Fernandes AD, MacPhee R, Reid G. 2010. Microbiome profiling by Illumina sequencing of combinatorial sequence-tagged PCR products. *PLoS One* 5:e15406.
- Gonzalez B, Iffliffe T, Macalady J, Schaperdorth I, Kakuk B. 2011. Microbial hotspots in anchialine blue holes: initial discoveries from the Bahamas. *Hydrobiologia* 677:149–56.
- Hadziavdic K, Lekang K, Lanzen A, Jonassen I, Thompson EM, Troedsson C. 2014. Characterization of the 18S rRNA gene for designing universal eukaryote specific primers. *PLoS One* 9:e87624.
- Harrell FE Jr, Dupont M. 2019. Package ‘Hmisc.’ R Package Version 4.2.0 (<https://cran.r-project.org/web/packages/Hmisc/Hmisc.pdf>).
- Havird J, Weeks J, Hau S, Santos S. 2013. Invasive fishes in the Hawaiian anchialine ecosystem: investigating potential predator avoidance by endemic organisms. *Hydrobiologia* 716:189–201.
- Hoffman SK, Seitz KW, Havird JC, Weese DA, Santos SR. 2018a. Diversity and the environmental drivers of spatial variation in *Bacteria* and micro-*Eukarya* communities from the Hawaiian anchialine ecosystem. *Hydrobiologia* 806:265–82.
- Hoffman SK, Seitz KW, Havird JC, Weese DA, Santos SR. 2018b. Comparing the community structure of *Bacteria* and micro-*Eukarya* from the Hawaiian anchialine ecosystem during wet and dry seasons. *Aquat Microb Ecol* 82:87–104.
- Holthuis LB. 1973. Caridean shrimps found in land-locked saltwater pools at four Indo-West Pacific localities (Sinai Peninsula, Funafuti Atoll, Maui and Hawaii Islands), with the description of one new genus and four new species. *Zool Verh* 128:1–48.
- Humphreys WF. 1999. Physico-chemical profile and energy fixation in Bundera Sinkhole, an anchialine remiped habitat in north-western Australia. *J R Soc West Aust* 82:89–98.
- Humphreys W, Tetu S, Elbourne L, Gillings M, Seymour J, Mitchell J, Paulsen I. 2012. Geochemical and microbial diversity of Bundera Sinkhole, an anchialine system in the eastern Indian ocean. *Nat Croat* 21:59–63.
- Iffliffe TM. 2002. Conservation of anchialine cave biodiversity. In: Martin JB, Wicks CM, Sasowsky ID, editors. *Hydrogeology and biology of post-paleozoic carbonate aquifers*. Charles Town (WV): Karst Waters Institute. p. 99–102.
- Ishino S, Ishino Y. 2014. DNA polymerases as useful reagents for biotechnology – the history of developmental research in the field. *Front Microbiol* 5:465.
- Jaccard P. 1908. Nouvelles recherches sur la distribution florale. *Bull Soc Vaud Sci Natl* 44:223–70.
- Jansson JK, Hofmockel KS. 2020. Soil microbiomes and climate change. *Nat Rev Microbiol* 18:35–12.
- Ju F, Xia Y, Guo F, Wang Z, Zhang T. 2014. Taxonomic relatedness shapes bacterial assembly in activated sludge of globally distributed wastewater treatment plants. *Environ Microbiol* 16:2421–32.
- Konhauser KO, Jones B, Reysenbach A-L, Renaut RW. 2003. Hot spring sinters: keys to understanding Earth’s earliest life forms. *Can J Earth Sci* 40:1713–24.
- Krause S, Le Roux X, Niklaus PA, Van Bodegom PM, Lennon JT, Bertilsson S, Grossart H-P, Philippot L, Bodelier PL. 2014. Trait-based approaches for understanding microbial biodiversity and ecosystem functioning. *Front Microbiol* 5:251.
- Krstulović N, Šolić M, Šantić D, Jasna M-L, Ordulj M, Šestanović S. 2013. Microbial community structure in two anchialine caves on Mljet Island (Adriatic Sea). *Acta Adriat* 54:183–98.
- Ley RE, Harris JK, Wilcox J, Spear JR, Miller SR, Bebout BM, Maresca JA, Bryant DA, Sogin ML, Pace NR. 2006. Unexpected diversity and complexity of the Guerrero Negro hypersaline microbial mat. *Appl Environ Microbiol* 72:3685–95.
- Lindemann SR, Moran JJ, Stegen JC, Renslow RS, Hutchison JR, Cole JK, Dohnalkova AC, Tremblay J, Singh K, Malfatti SA, et al. 2013. The epsomitic phototrophic microbial mat of Hot Lake, Washington: community structural responses to seasonal cycling. *Front Microbiol* 4:323.
- Lozupone C, Lladser ME, Knights D, Stombaugh J, Knight R. 2011. UniFrac: an effective distance metric for microbial community comparison. *ISME J* 5:169–72.
- Macarthur R, Levins R. 1967. The limiting similarity, convergence, and divergence of coexisting species. *Am Nat* 101:377–85.
- Maciolek JA. 1983. Distribution and biology of Indo-Pacific insular hypogean shrimps. *Bull Mar Sci* 33:606–18.
- Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD. 2012. PANDAseq: paired-end assembler for Illumina sequences. *BMC Bioinformatics* 13:31.
- McMurdie PJ, Holmes S. 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217.

- Mejía-Ortíz LM, Yáñez G, López-Mejía M. 2007. Echinoderms in an anchialine cave in Mexico. *Mar Ecol* 28:31–4.
- Menning D, Boop LM, Graham ED, Garey JR. 2014. Molecular analyses of microbial abundance and diversity in the water column of anchialine caves in Mallorca, Spain. *Int J Speleol* 43:9.
- Nutman AP, Bennett VC, Friend CRL, Van Kranendonk MJ, Chivas AR. 2016. Rapid emergence of life shown by discovery of 3,700-million-year-old microbial structures. *Nature* 537:535–8.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2015. *Vegan: community ecology package*. R package version 2.3-0 (<https://cran.r-project.org/web/packages/vegan/index.html>).
- Paerl HW, Pinckney JL, Steppe TF. 2000. Cyanobacterial-bacterial mat consortia: examining the functional unit of microbial survival and growth in extreme environments. *Environ Microbiol* 2:11–26.
- Papineau D, Walker JJ, Mojzsis SJ, Pace NR. 2005. Composition and structure of microbial communities from stromatolites of Hamelin Pool in Shark Bay, Western Australia. *Appl Environ Microbiol* 71:4822–32.
- Parks DH, Tyson GW, Hugenholtz P, Beiko RG. 2014. STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics* 30:3123–4.
- Prescott SL. 2017. History of medicine: origin of the term microbiome and why it matters. *Hum Microbiome J* 4:24–5.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:D590–6.
- R Development Core Team. 2008. *R: a language and environment for statistical computing*, Vienna, Austria.
- Royalty TM, Steen AD. 2019. Quantitatively partitioning microbial genomic traits among taxonomic ranks across the microbial tree of life. *mSphere* 4:e00446–19.
- Sarbu SM, Kane TC, Kinkle BK. 1996. A chemoautotrophically based cave ecosystem. *Science* 272:1953–5.
- Schneider D, Arp G, Reimer A, Reitner J, Daniel R. 2013. Phylogenetic analysis of a microbialite-forming microbial mat from a hypersaline lake of the Kiritimati Atoll, Central Pacific. *PLoS One* 8:e66662.
- Seymour JR, Humphreys WF, Mitchell JG. 2007. Stratification of the microbial community inhabiting an anchialine sinkhole. *Aquat Microb Ecol* 50:11–24.
- Shannon CE. 2001. A mathematical theory of communication. *SIGMOBILE Mob Comput Commun Rev* 5:3–55.
- Simpson EH. 1949. Measurement of diversity. *Nature* 163:688.
- Sinha R, Abu-Ali G, Vogtmann E, Fodor AA, Ren B, Amir A, Schwager E, Crabtree J, Ma S, Abnet CC, et al., The Microbiome Quality Control Project Consortium. 2017. Assessment of variation in microbial community amplicon sequencing by the Microbiome Quality Control (MBQC) project consortium. *Nat Biotechnol* 35:1077–86.
- Sket B. 1996. The ecology of anchialine caves. *Trends Ecol Evol* 11:221–25.
- Stock JH, Illiffe TM, Williams D. 1986. The concept 'anchialine' reconsidered. *Stygologia* 2:90–2.
- Suosaari EP, Reid RP, Araujo TAA, Playford PE, Holley DK, Mcnamara KJ, Eberli GP. 2016. Environmental pressures influencing living stromatolites in Hamelin Pool, Shark Bay, Western Australia. *Palaio* 31:483–96.
- Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, Schiweck C, Kurilshikov A, Joossens M, Wijmenga C, et al. 2019. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol* 4:623–632.
- Villanueva L, Navarrete A, Urmeneta J, White DC, Guerrero R. 2007. Analysis of diurnal and vertical microbial diversity of a hypersaline microbial mat. *Arch Microbiol* 188:137–46.
- Walker JJ, Spear JR, Pace NR. 2005. Geobiology of a microbial endolithic community in the Yellowstone geothermal environment. *Nature* 434:1011–4.
- Walter MR, Bauld J, Brock TD. 1972. Siliceous algal and bacterial stromatolites in hot spring and geyser effluents of Yellowstone National Park. *Science* 178:402–5.
- Ward DM, Bateson MM, Ferris MJ, Kühl M, Wieland A, Koeppel A, Cohan FM. 2006. Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Philos Trans R Soc Lond B Biol Sci* 361:1997–2008.
- Whipps JM, Lewis K, Cooke RC. 1988. Mycoparasitism and plant disease control. In: Burge MN, editor. *Fungi in Biological Control Systems*. Manchester: Manchester University Press. p. 161–87.
- Wong D. 1975. *Algae of the anchialine pools at Cape Kinau, Maui, and aspects of the trophic ecology of Halocaridina rubra* Holthuis (Decapoda, Atyidae) [Master of Science Thesis]. Honolulu: University of Hawai'i at Mānoa.
- Wong HL, Smith D-L, Visscher PT, Burns BP. 2015. Niche differentiation of bacterial communities at a millimeter scale in Shark Bay microbial mats. *Sci Rep* 5:15607.
- Wong HL, White RA, Visscher PT, Charlesworth JC, Vázquez-Campos X, Burns BP. 2018. Disentangling the drivers of functional complexity at the metagenomic level in Shark Bay microbial mat microbiomes. *ISME J* 12:2619–2639.
- Yager J. 1981. Remipedia, a new class of Crustacea from a marine cave in the Bahamas. *J Crustac Biol* 1:328–33.
- Yildirim S, Yeoman CJ, Sipos M, Torralba M, Wilson BA, Goldberg TL, Stumpf RM, Leigh SR, White BA, Nelson KE. 2010. Characterization of the fecal microbiome from non-human wild primates reveals species specific microbial communities. *PLoS One* 5:e13963.
- Zeevi D, Korem T, Godneva A, Bar N, Kurilshikov A, Lotan-Pompan M, Weinberger A, Fu J, Wijmenga C, Zhernakova A, et al. 2019. Structural variation in the gut microbiome associates with host health. *Nature* 568:43–48.