
# Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors

Kelly, L.W.; Williams, G.J.; Barott, K.L.; Carlson, C.A.; Dinsdale, E.A.; Edwards, R.A.; Haas, A.F.; Haynes, M.; Lim, Y.W.; McDole, T.; Nelson, C.E.; Sala, E.; Sandin, S.A.; Smith, J.E.; Vermeij, M.J.; Youle, M.; Rohwer, F.

**Proceedings of the National Academy of Sciences of the United States of America**

*DOI:*
10.1073/pnas.1403319111

Published: 30/06/2014

# Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors

Linda Wegley Kelly[1*], Gareth J. Williams[2], Katie L. Barott[1,2], Craig A. Carlson[3], Elizabeth A. Dinsdale[1], Robert A. Edwards[4], Andreas F. Haas[2], Matthew Haynes[1], Yan Wei Lim[1], Tracey McDole[1], Craig E. Nelson[5], Enric Sala[6], Stuart A. Sandin[2] Jennifer E. Smith[2], Mark J. A. Vermeij[7,8], Merry Youle[9], and Forest Rohwer[1]

[1]Department of Biology
San Diego State University
San Diego, CA USA

[2]Marine Biology Research Division
Scripps Institution of Oceanography
University of California, San Diego
La Jolla, CA USA

[3]Marine Science Institute
Department of Ecology, Evolution and Marine Biology
University of California, Santa Barbara
Santa Barbara, CA USA

[4]Department of Computer Sciences
San Diego State University
San Diego, CA USA

[5]Center for Microbial Oceanography: Research and Education
Department of Oceanography
University of Hawai`i
Honolulu, HI USA

[6]National Geographic Society
Washington, DC USA

[7]Caribbean Research and Management of Biodiversity (CARMABI)
Willemstad, Curacao

[8]Aquatic Microbiology, Institute for Biodiversity and Ecosystem Dynamics
University of Amsterdam
Amsterdam, The Netherlands

[9]Rainbow Rock,
Ocean View, HI USA

Running title: Selection and adaptation of coral reef-associated microbiomes

*corresponding author
Email: lwegley@gmail.com, Phone: 619-594-1336, Fax: 619-594-5676

**Author Contributions**

The manuscript was written by LWK, MY, and FR. Metagenomic analyses were completed by LWK. Multivariate statistical analyses were completed by GJW. Water samples for metagenomes and nutrient analysis were collected by KLB, CAC, EAD, AH, CEN, TM, SAS, ES and FR. The benthic characterizations were completed by JES, GJW, and KLB. YWL and MH completed all of the library prep and sequencing reactions. RAE provided valuable computational support. All of the authors offered helpful comments and edits to the manuscript.

**Abstract**

Holobionts are species-specific associations between macro- and micro-organisms. On coral reefs the benthic coverage of coral and algal holobionts varies due to natural and anthropogenic forcings. Different benthic macroorganisms are predicted to have specific microbiomes. In contrast, local environmental factors are predicted to select for specific metabolic pathways in microbes. To reconcile these two predictions, we hypothesized that adaptation of microbiomes to local conditions is facilitated by horizontal transfer of genes responsible for specific metabolic capabilities. To test this hypothesis, microbial metagenomes were sequenced from 22 coral reefs at eleven Line Islands in the central Pacific that together span a wide range of biogeochemical and anthropogenic influences. Consistent with our hypothesis, the percent cover of major benthic functional groups significantly correlated with particular microbial taxa. Reefs with higher coral cover had a "coral microbiome" with higher abundances of *Alphaproteobacteria* such as *Rhodobacterales* and *Sphingomonadales*, whereas microbiomes of algae-dominated reefs had higher abundances of *Gammaproteobacteria* , such as *Alteromonadales*, *Pseudomonadales*, and *Vibrionales*, *Betaproteobacteria* and *Bacteriodetes*. In contrast to taxa, geography was the strongest predictor of microbial community metabolism. Microbial communities on reefs with higher nutrient availability (e.g., equatorial upwelling zones) were enriched for genes involved in nutrient-related metabolisms (e.g., nitrate and nitrite ammonification, Ton/Tol transport, etc.). On reefs further from the equator, microbes had more genes encoding chlorophyll biosynthesis and photosystems I/II. These results support the hypothesis that core microbiomes are determined by the holobiont macroorganisms, and that those core taxa adapt to local conditions by selecting for advantageous metabolic genes.

## Statement of significance

Microbial communities associated with coral reefs influence the health and sustenance of the keystone benthic organisms (e.g., coral holobionts). The present study investigated the community structure and metabolic potential of microbes inhabiting coral reefs located across an extensive area in the central Pacific. We found that the taxa present correlated strongly with the percent coverage of corals and algae, while community metabolic potential correlated best with geographic location. These finding are inconsistent with prevailing biogeographic models of microbial diversity (e.g., distance decay) and metabolic potential (i.e., similar functional profiles regardless of phylogenetic variability). Based on these finding we propose that the primary carbon sources determine community structure and that local biogeochemistry determines finer scale metabolic function.

\body

## Introduction

Coral reefs are complex ecosystems that provide habitats for diverse, interdependent macro- and microorganisms. A coral colony itself is a complex holobiont, each made up of a coral polyp and a suite of prokaryotic microbes, viruses, protists, endolithic fungi and algae, and other invertebrates (1-4). Some coral-associated microbes confer benefits by, for example, remineralizing nutrients that are essential for the coral holobiont (5-9). Others contribute to coral demise by causing a number of specific diseases as well as non-specific detrimental effects (e.g., hypoxia) (10-12). On degraded reefs, where coral cover is reduced and the benthic surface is dominated by fleshy algae, the microbial community includes higher abundances of copiotrophic microbes, many of which are known pathogens (13). Higher abundances of potential pathogens on reefs are also known to correlate with higher prevalence of coral disease (14), indicating a link between the community structure of reef-associated microbes and coral health.

Previous studies have described the biogeographic distribution of pelagic microbial communities by investigating statistical relationships between pelagic microbes and environmental parameters (15-18). However, application of this approach to coral reef-associated microbes is complicated by a number of factors. First, for microbial members of specific coral holobionts, microbial biogeography is directly linked to the distribution of the coral species. Second, reef-associated microbial communities are influenced by the other benthic macroorganisms present, such as macroalgae – both calcifying and fleshy, which may vary markedly between locations. Third, these microbial communities are subject to abiotic factors, such as variable nutrient, temperature and hydrodynamic regimes associated with a particular geographic location. Given this

complexity, understanding the drivers that influence the community structure of reef-associated microbes requires unraveling numerous interdependent factors.

The relationships between microbial community structure, the metabolic capacity of the assemblage, and their habitat are complex. Numerous taxa share 'core' genes required for survival in the marine habitat. Supplementing these core housekeeping genes in each strain are a varied combination of metabolic genes (the pan-genome) associated with specialized pathways that contribute to fitness under particular local conditions, e.g., limited phosphate availability. These specialization genes do not respect species boundaries and may be found in multiple taxa adapted to similar environmental conditions (19, 20). Due to the mobility of these genes via horizontal gene transfer, the microbes can be considered to share a common gene pool, with specific genes being enriched within communities in particular niche habitats where they increase fitness. As a result, the similar community metabolism (i.e., functional redundancy) can be associated with high phylogenetic variability (21), and likewise communities comprised of similar taxa may differ in metabolic capabilities (22).

The mechanisms that govern community structure and gene flow in complex microbial communities, such as those associated with benthic marine habitats, remain largely unknown to the field of microbial ecology. Coral reefs are of particular interest because of their importance as centers of biodiversity, their contribution to global marine productivity, and their alarming decline. Coral reefs of the Line Islands (LIs) in the central Pacific offer a unique opportunity to investigate these questions as they span a latitudinal gradient from $6^{o}$ north to $11^{o}$ south. These islands and atolls (heron referred to as atolls) also span across the Equatorial Counter Current, Intertropical Convergence Zone and thus experience significant variability in nutrient concentrations, temperature and precipitation.

In addition to oceanographic variability, the northern LIs also span a gradient of human disturbance where Teraina, Tabuaeran, and Kiritimati support populations of approximately 1,000, 2,500, and 5,000 people, respectively. Reefs at these atolls are impacted by subsistence and commercial fishing, as well as some pollution (e.g., sewage, chemicals) and agricultural runoff. Some of the highest known biomass of the fishes for a coral reef ecosystem were observed on the unpopulated atolls (14, 23), where reefs were characterized by high cover of reef-building corals and crustose coralline algae, abundant coral recruits, and low levels of coral disease (14). In contrast, the populated atolls, most notably Kiritimati, had reefs with as low as 2% coral cover and were associated with higher abundances of super heterotrophs, many of which are known pathogens (13), and higher prevalence of coral diseases (14). Since the reefs at the uninhabited atolls have been largely spared from such anthropogenic disturbances, they provide a baseline for a comparative evaluation of the effects of human activity on coral reef-associated microbes. However, to definitively attribute any observed differences to anthropogenic activities, the role of other environmental drivers that differ between atolls must also be examined. For instance, the three inhabited atolls are clustered together in a region spanning <3° latitude, inciting a counterargument that local biogeochemical factors were responsible for reef degradation rather than fishing or other local activities as had been suggested by a prior study (14).

Here we used comparative metagenomics to tease out the key environmental factors driving the composition and metabolism of reef-associated microbial communities in the LIs. Although the eleven atolls are clustered in the same oceanic region, they differ in three key environmental variables that are predicted to influence their microbial communities: nutrient levels, latitudinal distance from the equator, and the percentage of the benthic surface occupied by various

functional groups of macroorganisms. In this study, we collected reef-associated microbes, then extracted and sequenced the community DNA. Taxonomic and functional annotations were assigned to the resultant reads by comparison to the SEED protein database. We then quantified variation in the structure and metabolic potential of the communities in relation to the three key variables. These comparisons show that (1) the microbial taxa present and their relative abundances reflect the benthic community whose carbon-containing exudates provide the primary local energy source, and (2) the presence of various specialized metabolic capabilities correlates with nutrient levels and other latitude-dependent factors.

## Results

Studies were conducted at 22 reef sites distributed across eleven LIs spanning 18° latitude (Table S1). At each site, seawater samples were collected at the surface of the benthos for microbial metagenome preparation and from the immediately overlying water for nutrient analysis. The macroorganisms comprising the benthic cover were surveyed. Subsequent analyses assessed the relationships between three predictor variables (benthic macroorganisms, nutrient levels, and latitude) and both the structure and the metabolic capabilities of the microbial communities at these atolls.

*Nutrient concentration.* Inorganic nitrogen (nitrate+nitrite) and phosphate concentrations were generally highest near the equator and declined with increasing latitude both north and south (Figure 1, Table S2). Nitrate+nitrite concentrations ranged from 0.52 to 4.83 μM, whereas phosphate concentrations varied less (0.15 to 0.44 μM). When compared to the northernmost

(Kingman) and southernmost (Flint) atolls, nitrate+nitrite and phosphate concentrations at equatorial Jarvis were approximately five- and two-fold higher, respectively.

*Benthic macroorganisms*. The benthic cover was quantified as the percentage covered by each of seven functional groups: hard coral, crustose coralline algae (CCA), calcified macroalgae, soft coral, fleshy macroalgae, fleshy turf algae, and 'other' (Table S2). A list of the genera within each category is also provided (Table S3). Coral cover varied markedly from 2.2% at one site on Kiritimati to 86.7% at one site on Malden (mean = 44.4%, Table S2). In general, the uninhabited atolls were dominated by reef building calcifiers including coral, CCA and calcified macroalgae (24) while fleshy algae such as turf and fleshy macroalgae dominate the inhabited atolls (14).

*Reef-associated microbes*. DNA isolated from microbes sampled at each site was sequenced to yield 22 metagenomic libraries totaling 2.25 million quality reads (average length 389 bp; Table S1). The sequenced reads were translated *in silico* into predicted protein sequences; subsequent comparison to the SEED database provided taxonomic annotations for 21% to 47% of the reads and assignments to functional subsystems for 27% to 62% of the reads from each site. These annotations were the basis for comparative analyses of the microbial community structure and metabolic capabilities across the LI archipelago.

The relative abundances of the major taxonomic groups were tabulated (Figure S1), plotted in 2D using non-metric multidimensional scaling (nMDS; Figure 2A), and analyzed for multivariate structure using SIMPROF (Figure S2). By all measures, geographic location of the atoll was a poor predictor of similarity for microbial community structure. For example, the two northernmost atolls, Kingman and Palmyra, clustered with the Southern LIs in Group 1 (Figure 2A) and were most similar to Millennium, one of the southernmost atolls. Likewise Malden and

Flint, separated by nearly 900 km, had similar taxonomic composition. In contrast, the metabolic capabilities (based on level 1 subsystem designations in the SEED; N=20) of microbial communities in geographic proximity were more similar, forming three groups corresponding to the northern, middle, and southern atolls (Figure 2B, Figure S3). SIMPROF analyses conducted at the site level resulted in a higher number of significant groupings, though each site generally remained located within its own atoll group (Figure S2, Figure S3) provided some exceptions, particularly in the metabolic groupings (e.g., Flint 2 clustered with Group 3 atolls, Figure S3). Further analyses were performed to quantify correlations between three key variables and both microbial community structure and metabolism across the LIs.

*Community structure*. The correlations visualized by the CCA (Figure 3A) illustrate that microbial community structure on LI reefs is closely associated with benthic community composition. Reefs at all of the uninhabited LIs (Group 1 in Figures 2A and S2) associated with higher percent cover of reef building calcifiers were characterized by higher abundances of Cyanobacteria, *Alphaproteobacteria* (i.e., orders *Rhodobacterales* and *Rickettsiales*), and Firmicutes. Reefs with the highest hard coral coverage, such as Malden and Flint, had higher abundances of *Sphingomonadales* and Cyanobacteria (Figure 3A). Though the abundance of the genus *Synechococcus* correlates positively with nutrient concentration in pelagic microbial communities, here it was positively correlated with the percentage of hard coral cover (Table 1, $r = 0.665$, $p = 0.026$). In contrast, hard coral cover showed a strong negative correlation with the abundance of *Alteromonadales* ($r = -0.819$, $p = 0.002$).

The inhabited Group 2 atolls associated with higher percent cover of fleshy macroalgae (Tabuaeran and Teraina; Figure 3A) had greater abundances of *Gammaproteobacteria* (e.g., orders *Enterobacteriales*, and *Pseudomonadales*) and *Betaproteobacteria*. In contrast, the reefs

252 at populated Kiritimati were dominated by fleshy turf algae (58.9-82.4%) and supported a
253 markedly increased abundance of *Bacteriodetes* (25.1% ±4.2%, N=2) compared to the other
254 atolls (7.2%±3.5%, N=20). Specifically, five genera within the class *Flavobacteria* (genera
255 *Croceibacter*, *Dokdonia*, *Gramella*, *Leeuwenhoekiella*, and *Polaribacter*) were consistently
256 overrepresented compared to sites on other atolls. Overall, the percent coverage of fleshy turf
257 algae on LI reefs was positively correlated with bacteria from the orders *Flavobacteriales* and
258 *Alteromonadales* (Table 1, $r = 0.815$, $p = 0.002$ and $r = 0.682$, $p = 0.021$, respectively). The CCA
259 also depicted a correlation between the percent cover of other benthic organisms and Kiritimati
260 reefs. Though other benthic organisms contributed to <1% of the benthic composition on most LI
261 reefs, the 2 sites on Kiritimati had a higher percentage of sand, which contributed to the higher
262 percent cover of this category (5.2% ± 0.5%).

263 A distance-based linear model (DistLM) was used to formally quantify which suite of predictor
264 variables formed the best-fit model (balancing performance with complexity) for explaining
265 variations in microbial communities across LI reefs. Hard coral alone had the largest impact on
266 microbial community structure explaining 15.2% of the variation between reefs (Table S4).

267 *Community metabolism.* Distance from the equator was the strongest predictor of community
268 metabolism, explaining 18.4% of the variation in microbial metabolic potential (Table S4). The
269 two northern atolls (Group 2 in Figure 2B; Kingman and Palmyra) were characterized by high
270 abundances of genes encoding cofactors, RNA metabolism, and protein metabolism. Moving
271 southward, the mid-latitude atolls (Group 3 in Figure 2B; Jarvis, Kiritimati, Teraina, and
272 Tabuaeran) were characterized by higher abundances of genes for aromatic compound
273 utilization, iron metabolism, membrane transport, nitrogen metabolism, potassium metabolism,

regulation, and virulence. All of the southern Line Islands were combined into one group and had similar community metabolism (Group 1, Figure 2B).

The question remained as to which environmental parameters associated with latitude were driving these variations. Nutrient levels varied across the LIs as expected due to the influence of equatorial upwelling (Figure 1). As such, a number of metabolic pathways (SEED level 3 subsystems) demonstrated significant correlations with local phosphate concentrations across all eleven atolls. These included six pathways positively correlated with phosphate concentration: conjugative transfer, chemotaxis, nitrate and nitrite ammonification, cobalt-zinc-cadmium resistance, multidrug resistance efflux pumps, and Ton and Tol transport (Figure 4A, Table S5). Phosphate concentration was negatively correlated with two metabolic pathways involved in photosynthesis: chlorophyll biosynthesis and photosystems I and II (Figure 4B and Table S5), and also with the abundance of *Prochlorococcus* (Table 1). Genes for ribosomal proteins were also overrepresented at oligotrophic sites (Figure 4B).

*Interisland Comparison.* Atolls in close proximity were observed to have similar metabolic capabilities despite differences in their taxonomic composition. For example, microbial communities from the geographically close Jarvis and Kiritimati had similar metabolic profiles (Figure 2), but the taxonomic profile of Jarvis was most similar to Vostok and Starbuck, while that for Kiritimati was the most dissimilar of all (Figure 2). Conversely, the distant atolls of Kingman and Malden supported taxonomically similar microbial communities that encoded divergent metabolic capabilities. Hence, microbial communities composed of different taxa can encode similar functions, and vice versa.

## Discussion

This study reports the first large-scale metagenomic survey of the microbial communities associated with coral reefs that simultaneously characterizes both taxonomic composition and metabolic capabilities. We have demonstrated that, at the ecosystem level, benthic macroorganisms most strongly influence the taxonomic composition of the microbial community, while metabolic specialization genes carried by these taxa vary between locations and reflect functional adaptations to local oceanographic conditions.

For this study, microbial communities were sampled from 22 coral reef sites at eleven atolls across the Line Island (LI) archipelago, atolls that differed with respect to their benthic community, nutrient levels, and latitude. The microbes collected by our procedure were closely associated with the surface of the benthic macroorganisms (corals and algae). As a result, they included species-specific bacterial components of the coral holobiont (1) as well as specific bacterial taxa associated with some algal functional groups (1, 25). In addition, the microbial communities sampled on these reefs reflected selection by the adjacent benthic macroorganisms, as evidenced by the differences between reef-associated bacterioplankton communities and open ocean communities (26). There is evidence that reef-associated communities undergo selection in shallow reef environments by the locally available labile organic matter exuded by the benthic organisms (27). For example, in an empirical study Nelson and colleagues demonstrated that exudates collected from coral and macroalgae selectively fostered growth of distinct bacterioplankton communities (27). Coral exudates promoted communities with higher diversity, including lineages of *Alphaproteobacteria* with relatively few virulence factors (e.g., *Erythrobacteraceae)*; whereas exudates from fleshy macroalgae selected for less diverse

communities with more copiotrophic *Gammaproteobacteria* lineages (e.g., the families *Alteromonadaceae, Pseudoalteromonadaceae*, and *Vibrionaceae*).

*Community structure*. The current study confirms and extends earlier findings (27) by demonstrating similar correlations between the benthic community composition and the enrichment of specific microbial taxa on coral reefs *in situ* (Table 1). Consistent with the effects of individual exudates, high coral cover was associated with higher abundances of *Alphaproteobacteria*, while the abundant fleshy macroalgae at Tabuaeran and Teraina were accompanied by more *Gammaproteobacteria* (e.g., *Enterobacteriales* and *Pseudomonadales*). In addition, the fleshy turf algae that dominated Kiritimati favored *Flavobacteria* (phylum Bacteriodetes) including genera increased by turf algal exudates (*Dokdonia*, *Gramella*, and *Leeuwenhoekiella*) (28). Together, these complementary research approaches indicate that coral- and algae-derived organic exudates enrich for specific types of bacteria living in close association with coral reefs.

Nutrient levels have also been postulated to influence microbial community composition. Here we tested this hypothesis using the natural nutrient gradient present across the LIs. Due to the equatorial Pacific upwelling in this region, phosphate and nitrate are elevated at the equator and decrease with latitude both north and south (Figure 1). In high-nitrate, low-chlorophyll ecosystems such as this, iron may be the nutrient limiting primary production (29). Other unspecified biogeographic factors also vary with latitude across the LIs. In this study, neither nutrients nor other latitude-dependent variables were included in the best fit model for determining microbial community structure. Therefore, we propose that on these geographically separate coral reefs, microbial community structure is determined by the available energy source,

i.e., the DOC provided in the form of benthic exudates, which provides a mechanism for the correlations observed between the macro- and microbial components of reef communities.

*Community metabolism*. In contrast to community structure, specialized and ecologically-relevant metabolic capabilities of these communities reflected local nutrient concentrations. For example, six level 3 metabolic subsystems (the SEED database) correlated positively with phosphate concentration across the LIs (Figure 4A). Some of these, such as the TonB system, contribute to nutrient acquisition. The TonB system transports large molecules in through the outer membrane of Gram-negative Bacteria, e.g., polysaccharides, proteins, and siderophores. Its importance in marine environments is evidenced by the presence of these genes in marine bacterial genomes and pelagic metagenomes (30-32), their high levels of expression in metatranscriptome data (33), and the proteomic identification of their products as the predominant membrane proteins in pelagic Bacteria (34). In this study, they accounted for nearly 1% of gene function annotations at some high nutrient sites (Figure 4A). Genes of the conjugative transfer subsystem, also overrepresented at high nutrient sites, may function in energy and nutrient acquisition via type IV secretion of ectoenzymes and siderophores, and may support active horizontal gene transfer via conjugation. Conversely, the more oligotrophic sites exhibited overrepresentation of two photosynthesis pathways (chlorophyll biosynthesis and photosystems I and II) (Figure 4B and Table S5), as well as greater abundance of *Prochlorococcus*, a key primary producer in the oligotrophic oceans (Table 1).

Previous studies have shown that the anaerobic ammonification of nitrate and nitrite (also referred to as dissimilatory nitrate reduction to ammonium, DNRA) is significant for nitrogen metabolism in the diffusive boundary layer, an environment with heterogeneous distribution of dissolved oxygen during the day (12) that then becomes anoxic at night (35). That anaerobes

dominate coral-associated microbial communities suggested that this anaerobic nitrogen metabolism may be important on coral surfaces (25). An interesting observation from the nutrient measurements is that atolls with higher nitrate+nitrite availability have lower ammonium concentrations whereas low nitrate+nitrite atolls have higher ammonium. Nitrate+nitrite to ammonium ratios were 0.26, 0.29, and 0.22 on Malden, Jarvis, and Kiritimati compared to 3.23 and 1.47 on Flint and Kingman, respectively (Table S2). Therefore, the overrepresentation of DNRA may reflect the lower abundances of ammonium at these high nutrient sites.

Reef-associated microbial communities in high nutrient environments encoded greater metabolic complexity, suggesting that they carry more specialization genes and thus generally possess larger genomes (Figure S4). Consistent with this hypothesis, single-copy genes encoding ribosomal proteins were overrepresented at oligotrophic sites (Figure 4B), indicating that the community overall possessed smaller genomes compared to those at high nutrient sites.

Although both phosphorus and nitrogen concentrations correlated with distance from the equator ($r = -0.74$ and $-0.64$, respectively, Table S6), neither was as strong a predictor of metabolism as was latitudinal distance from the equator (as assessed by DistLM analysis). Distance from the equator may serve as a proxy for other influential but unsampled variables such as seawater temperature, salinity, PAR, or micronutrient concentrations (e.g., iron). In addition, the limited sampling (1-4 sites at each atoll) may have obscured significant correlations to specific nutrients. Had the atoll averages been based on sampling of 20+ sites per atoll, significant correlation with specific nutrients might have been discernible. Nevertheless, the availability of the macronutrients nitrate+nitrite and phosphate are posited to be important factors influencing microbial community metabolism on LI reefs.

*Anthropogenic impacts on LI reefs*. The findings of this study indicate that local human populations influence the reef-associated microbial community indirectly by influencing composition of benthic macroorganisms. Typically activities such as fishing remove important grazing herbivore species resulting in increased cover of fleshy algae, and this in turn profoundly impacts microbial community structure at the populated atolls (Figure 2, Figure S1). Increased coverage by fleshy algae selects for specific microbes that may be detrimental to coral health (27, 36), thereby opening additional benthic space for further algal colonization (37).

*Discordance between taxa and metabolism.* Both the abundance of specific taxonomic groups and the community metabolic capabilities of the reef-associated microbial communities varied across the Line Islands. Both correlated with ecological factors, but did so independent of each other. As a result, atolls as far apart as Kingman and Malden (~1,400 km) hosted taxonomically similar communities, but these communities effectuated different metabolisms. Conversely, the different microbial communities at equatorial Jarvis and Kiritimati encoded similar metabolic specialization genes. This discordance between taxonomy and metabolism is novel and intriguing. We hypothesize that while community structure is attributable to the core genes that classify each taxon, community metabolism reflects the particular complement of specialization genes that comprise the dynamic genome of each strain present. Previously, strain-specific adaptation to different nutrient levels had been documented in marine cyanobacteria for genes involved in phosphate acquisition. The particular genes present and their genomic organization depended on phosphate availability in each isolate's source environment. Strains of *Prochlorococcus* that showed 99.9% similarity of their 16S rRNA genes nevertheless possessed different phosphate metabolism genes located in different genomic locations (19). Conversely, some more divergent strains that occupied environments with similar nutrient regimes shared

similar phosphate gene content and organization. Additionally, although *Prochlorococcus* typically assimilates only ammonium, in regions of nitrogen limitation strains have adapted to utilize nitrate and nitrite by using genes acquired horizontally from *Synechococcus* (20).

The observed adaptation of microbial community metabolism patterns could have resulted from either gene acquisition and loss or shifts in the relative abundances of strains adapted to different conditions. Traditionally, only changes in strain abundance (i.e., beta diversity) have been considered as possible drivers of rapid adaptation in ecological time. Increased genetic diversity, i.e., evolution, by mechanisms such as horizontal movement of genes between strains or species, has been expected to require evolutionary time. We posit that in these microbial communities, evolution is rapid, occurring in ecological time.

Attempts to identify the evolutionary mechanisms active in this situation have been hampered by the limited representation of marine microbes in databases (38), such as the SEED, due to our inability to culture most species (39). The availability of single-cell whole genome amplification methods (40) promises to enable genomic characterization of unculturable marine microbes, thereby substantially accelerating resolution of this question.

## Materials and Methods

Metagenomic sequence reads were compared to the SEED protein database using BLASTx. For taxonomic annotation, sequences with significant similarities (E-value $<10^{-5}$) were assigned to the closest identified microbial representative. For functional annotation, sequences were assigned the function of the closest identified protein and these functions were then grouped into metabolic pathways according to the subsystems in the SEED database. Community structure

was compared using the relative abundances of 19 higher rank microbial taxa (see SI Text and Table S7 for clarification of taxonomic groups). Similarly, community metabolism was determined by comparing the relative abundance of 20 Level 1 subsystem categories in the SEED database.

Non-metric multidimensional scaling (nMDS) analyses were used with the annotated metagenome data to visualize between-atoll similarity in terms of two discrete response variables: community structure and community metabolism. For an initial exploration of potential correlations between the three predictor variables and either microbial community structure or metabolism, a canonical correspondence analysis (CCA) was performed using the R package vegan. To formally quantify how much variation in the microbial communities or their metabolism could be explained by the predictors measured (continuous variables), a permutational distance-based multivariate linear model (DistLM) was used in PERMANOVA+. Full methods and any associated references are available in the SI.

**Acknowledgements**

The microbial samples were collected during two research expeditions to the LIs funded by the National Geographic Society, the Moore Family Foundation, the Hawaii Undersea Research

Lab of Coral Reef Conservation (HURL,) a program of NOAA, and several private donors and during one RAMP cruise to Jarvis supported and executed by NOAA-CRED. This work was carried out under research permits from the Palmyra Atoll National Wildlife Refuge, US Fish and Wildlife Service, and the Environment and Conservation Division of the Republic of Kiribati. Thank you Beltran Rodriguez-Mueller, Rob Schmieder, Bahador Nosrat, Federico

Lauro, Nao Hisakawa, Jeremy Frank, Bas Dutilh, Katrine Whiteson, Barbara Bailey, and Jim Nulton for mathematical and bioinformatics support and to the Palmyra Atoll Research Consortium and the Nature Conservancy for field support. We are grateful to Jennifer Martiny for valuable discussions and Heather Maughan for her editing expertise. This research was sponsored by NSF awards OCE-0927415 and DEB-1046413 to FR, OCE-0927411 to CAC, OCE-0417412 to the MCR-LTER, and the CIFAR IMB Fellowship 141679 to FR.

**References**

1. Rohwer F, Seguritan V, Azam F, & Knowlton N (2002) Diversity and distribution of coral-associated bacteria. *Mar Ecol-Prog Ser* 243:1-10.
2. Bourne DG & Munn CB (2005) Diversity of bacteria associated with the coral *Pocillopora damicornis* from the Great Barrier Reef. *Environ Microbiol* 7(8):1162-1174.
3. Koren O & Rosenberg E (2006) Bacteria associated with mucus and tissues of the coral Oculina patagonica in summer and winter. *Appl Environ Microb* 72(8):5254-5259.
4. Sunagawa S, Woodley CM, & Medina M (2010) Threatened Corals Provide Underexplored Microbial Habitats. *Plos One* 5(3):e9554.
5. Knowlton N & Rohwer F (2003) Multispecies microbial mutualisms on coral reefs: The host as a habitat. *Am Nat* 162(4):S51-S62.
6. Lesser MP, *et al.* (2007) Nitrogen fixation by symbiotic cyanobacteria provides a source of nitrogen for the scleractinian coral Montastraea cavernosa. *Marine Ecology Progress Series* 346:143-152.
7. Fiore CL, Jarett JK, Olson ND, & Lesser MP (2010) Nitrogen fixation and nitrogen transformations in marine symbioses. *Trends in Microbiology* 18(10):455-463.
8. Wegley L, Edwards R, Rodriguez-Brito B, Liu H, & Rohwer F (2007) Metagenomic analysis of the microbial community associated with the coral Porites astreoides. *Environ Microbiol* 9(11):2707-2719.
9. Raina JB, Tapiolas D, Willis BL, & Bourne DG (2009) Coral-associated bacteria and their role in the biogeochemical cycling of sulfur. *Appl Environ Microb* 75(11):3492-3501.
10. Smith JE, *et al.* (2006) Indirect effects of algae on coral: algae-mediated, microbe-induced coral mortality. *Ecol Lett* 9(7):835-845.
11. Barott K, *et al.* (2009) Hyperspectral and Physiological Analyses of Coral-Algal Interactions. *Plos One* 4(11):e8043.
12. Haas AF, *et al.* (2013) Visualization of oxygen distribution patterns caused by coral and algae. *PeerJ* 1:e106.
13. Dinsdale EA, *et al.* (2008) Microbial ecology of four coral atolls in the northern Line Islands. *Plos One* 3(2):e1584.
14. Sandin SA, *et al.* (2008) Baselines and degradation of coral reefs in the northern Line Islands. *Plos One* 3(2):e1548.

15. Hewson I, Paerl RW, Tripp HJ, Zehr JP, & Karl DM (2009) Metagenomic potential of microbial assemblages in the surface waters of the central Pacific Ocean tracks variability in oceanic habitat. *Limnol Oceanogr* 54(6):1981-1994.
16. Fuhrman JA, *et al.* (2008) A latitudinal diversity gradient in planktonic marine bacteria. *P Natl Acad Sci USA* 105(22):7774-7778.
17. Pommier T, *et al.* (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Mol Ecol* 16(4):867-880.
18. Martiny AC, Tai APK, Veneziano D, Primeau F, & Chisholm SW (2009) Taxonomic resolution, ecotypes and the biogeography of Prochlorococcus. *Environ Microbiol* 11(4):823-832.
19. Martiny AC, Coleman ML, & Chisholm SW (2006) Phosphate acquisition genes in Prochlorococcus ecotypes: Evidence for genome-wide adaptation. *P Natl Acad Sci USA* 103(33):12552-12557.
20. Martiny AC, Kathuria S, & Berube PM (2009) Widespread metabolic potential for nitrite and nitrate assimilation among Prochlorococcus ecotypes. *P Natl Acad Sci USA* 106(26):10787-10792.
21. Burke C, Steinberg P, Rusch D, Kjelleberg S, & Thomas T (2011) Bacterial community assembly based on functional genes rather than species. *P Natl Acad Sci USA* 108(34):14288-14293.
22. Moore LR, Rocap G, & Chisholm SW (1998) Physiology and molecular phylogeny of coexisting Prochlorococcus ecotypes. *Nature* 393(6684):464-467.
23. Williams ID, *et al.* (2011) Differences in Reef Fish Assemblages between Populated and Remote Reefs Spanning Multiple Archipelagos Across the Central and Western Pacific. *Journal of Marine Biology* 2011:14pp.
24. Williams GJ, *et al.* (2013) Benthic communities at two remote Pacific coral reefs: effects of reef habitat, depth, and wave energy gradients on spatial patterns. *PeerJ* 1:e81.
25. Barott KL, *et al.* (2011) Microbial diversity associated with four functional groups of benthic reef algae and the reef-building coral Montastraea annularis. *Environ Microbiol* 13(5):1192-1204.
26. Nelson CE, Alldredge AL, McCliment EA, Amaral-Zettler LA, & Carlson CA (2011) Depleted dissolved organic carbon and distinct bacterial communities in the water column of a rapid-flushing coral reef ecosystem. *The ISME journal* 5(8):1374-1387.
27. Nelson CE, *et al.* (2013) Coral and macroalgal exudates vary in neutral sugar composition and differentially enrich reef bacterioplankton lineages. *Isme Journal* 7(5):962-979.
28. Kelly LW, *et al.* (In Preparation) Taxonomic and functional gene analysis of the microbial communities stimulated by the dissolved organic matter released from three benthic coral reef primary producers.
29. Martin JH, *et al.* (1994) Testing the iron hypothesis in ecosystems of the equatorial Pacific-Ocean. *Nature* 371(6493):123-129.
30. Hopkinson BM & Barbeau KA (2012) Iron transporters in marine prokaryotic genomes and metagenomes. *Environ Microbiol* 14(1):114-128.
31. Tang K, Jiao NZ, Liu KS, Zhang Y, & Li SH (2012) Distribution and Functions of TonB-Dependent Transporters in Marine Bacteria and Environments: Implications for Dissolved Organic Matter Utilization. *Plos One* 7(7): e41204.

32. Dupont CL, *et al.* (2012) Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J* 6(6):1186-1199.
33. Ottesen EA, *et al.* (2011) Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *Isme Journal* 5(12):1881-1895.
34. Morris RM, *et al.* (2010) Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction. *ISME J* 4(5):673-685.
35. Shashar N, Cohen Y, & Loya Y (1993) Extreme Diel Fluctuations of Oxygen in Diffusive Boundary-Layers Surrounding Stony Corals. *Biol Bull* 185(3):455-461.
36. Kelly LW, *et al.* (2012) Black reefs: iron-induced phase shifts on coral reefs. *Isme Journal* 6(3):638-649.
37. Barott KL & Rohwer FL (2012) Unseen players shape benthic competition on coral reefs. *Trends in Microbiology* 20(12):621-628.
38. Yooseph S, *et al.* (2010) Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* 468(7320):60-66.
39. Rappe MS & Giovannoni SJ (2003) The uncultured microbial majority. *Annu Rev Microbiol* 57:369-394.
40. Stepanauskas R & Sieracki ME (2007) Matching phylogeny and metabolism in the uncultured marine bacteria, one cell at a time. *Proceedings of the National Academy of Sciences* 104(21):9052-9057.

**Figure Legends**

Figure 1. The Line Islands and their nutrient concentrations. (A) The eleven main atolls sampled in this study. Scale bar indicates latitude and distance between atolls. Atoll sizes are proportionate, but not to scale. (B) Average nutrient concentrations at the eleven atolls. Nutrient concentrations were measured in triplicate for each of the 22 study sites (N = 66) and averaged; sites were then averaged for each atoll. Solid and dashed error bars show the standard error for atoll and site replicates, respectively. Average values for each site are provided in Table S2.

Figure 2. Non-metric multidimensional scaling plots for the relative abundances of taxonomic similarities (*A*) and metabolic subsystem similarities (*B*). Sites were averaged for each atoll. The 2D stress values for are 0.05 and 0.03 for the taxonomic and metabolic similarities, respectively. Dark gray circles indicate significant groupings from the SIMPROF analysis (Figures S2 and S3;

570 Bray-Curtis similarity, p-value <0.01). Light gray circles cluster atolls with greatest similarity
571 within each statistically significant group.
572
573 Figure 3. Canonical correspondence analysis (CCA) depicting the correlations between predictor
574 variables (blue) and the relative abundance of taxonomic similarities (*A*) and metabolic
575 similarities (*B*) at each Line Island. Loading vectors for the taxa and subsystems are shown in
576 red. Altero, *Alteromonadales*; Betaproteo, *Betaproteobacteria*; Enterob, *Enterobacteriales*;
577 Oceano, *Oceanospirillales*; OtherAlphas, Other *Alphaproteobacteria*; Pseudomon,
578 *Pseudomonadales*; Rhodobact, *Rhodobacterales*; Sphing, *Sphingomonadales*; calc macro,
579 calcified macroalgae; cca, crustose coralline algae; macro, fleshy macroalgae; soft, soft coral;
580 dist, distance from the equator in degrees latitude.
581
582 Figure 4. Metabolic pathways that correlate positively (*A*) and negatively (*B*) with increasing
583 distance from the equator (decreasing nutrient concentrations) across the Line Islands. Pathways
584 are level 3 subsystem annotations from the SEED database.