MMBR
Journals.ASM.org

Evolutionary Ecology of the Marine *Roseobacter* Clade

Haiwei Luo, Mary Ann Moran

Department of Marine Sciences, University of Georgia, Athens, Georgia, USA

SUMMARY 573
INTRODUCTION 573
WHAT DEFINES A ROSEOBACTER? 573
ECOLOGICAL AND EVOLUTIONARY GENOMICS OF CULTIVATED ROSEOBACTERS 574
Genome Content 574
Genome Evolution 574
ECOLOGICAL ASSOCIATIONS BETWEEN ROSEOBACTERS AND EUKARYOTES 577
Roseobacter-Phytoplankton Interactions 577
Roseobacter-Coral Interactions 577
ECOLOGICAL AND EVOLUTIONARY GENOMICS OF UNCULTIVATED ROSEOBACTERS 578
Genome Analyses 578
Evolution of an Enigmatic *Roseobacter* Lineage 578
Inclusion of Uncultured Roseobacters in Phylogenomic Analyses 579
Genome Streamlining in the *Roseobacter* Clade 579
A New Look at Population Size 580
CONCLUDING REMARKS 580
ACKNOWLEDGMENTS 582
REFERENCES 582
AUTHOR BIOS 587

SUMMARY

Members of the *Roseobacter* clade are equipped with a tremendous diversity of metabolic capabilities, which in part explains their success in so many different marine habitats. Ideas on how this diversity evolved and is maintained are reviewed, focusing on recent evolutionary studies exploring the timing and mechanisms of *Roseobacter* ecological diversification.

INTRODUCTION

The major clades of heterotrophic bacteria inhabiting surface ocean waters were initially discovered with culture-independent 16S rRNA gene surveys (1, 2) and found to include the alphaproteobacterial SAR11, *Roseobacter*, and SAR116 clades and the gammaproteobacterial SAR86 clade, among others (1–3). Studies over the next 2 decades revealed enormous genomic and physiological diversity among these major clades, correlating with differences in genome size, gene repertoire, G+C content, ecological strategy (free living, patch adapted, or eukaryote associated), and trophic strategy (heterotrophy, photoheterotrophy, or autotrophy). Such diversity is relevant from an ecological perspective because it determines the roles of these bacterial lineages in oceanic elemental cycles and their interactions with marine eukaryotes. However, how this diversity evolved and is maintained in a well-mixed seawater matrix has only begun to be addressed (4–8). In efforts to understand the ecology of marine bacteria and predict their responses in a changing ocean, the history and mechanisms of genome change are crucial puzzle pieces. In this review, we focus on the evolutionary processes underlying the ecological roles of the marine *Roseobacter* clade.

The *Roseobacter* clade is found predominantly in marine environments, representing up to 20% of bacterial cells in some coastal ecosystems and 3 to 5% of bacterial cells in open ocean surface waters (9). In addition to living freely in bulk seawater (10, 11), roseobacters are dominant members of the bacterial communities associated with phytoplankton (12–17), macroalgae (18, 19), and various marine animals (20–23). Both mutualistic (24, 25) and pathogenic (26–29) life-styles have been suggested. Clade members are ubiquitous in temperate and polar oceans (30); the latter include sea ice habitats, in which *Roseobacter* is a major bacterial phylotype (31, 32). Roseobacters are also abundant in coastal sediments (33), deep pelagic ocean (34), and deep-sea sediments (35).

Since the publication of the first *Roseobacter* genome from *Ruegeria pomeroyi* DSS-3 (formerly *Silicibacter pomeroyi* DSS-3) a decade ago (36), numerous other *Roseobacter* genome sequencing projects have been launched. To date, genomes of >50 isolates and 4 uncultivated single cells have been sequenced. An important lesson learned through genome analyses is that members of the *Roseobacter* clade are equipped with a tremendous diversity of metabolic capabilities and regulatory circuits (9, 25, 36–39), which in part explains their success in a variety of marine habitats. Ideas on how this diversity evolved and is maintained are reviewed, focusing on recent evolutionary studies yielding information on the timing and mechanisms of ecological diversification.

WHAT DEFINES A ROSEOBACTER?

Members of the *Roseobacter* clade form a subgroup of the *Rhodobacterales* that shares ≥89% identity of 16S rRNA gene sequences (40). Most of the cultured members have large genomes, high

Address correspondence to Mary Ann Moran, mmoran@uga.edu.
Copyright © 2014, American Society for Microbiology. All Rights Reserved.
doi:10.1128/MMBR.00020-14

G+C content (60% ± 4%), and versatile metabolic capabilities. An interesting exception is a cultured member branching at the base of the *Roseobacter* phylogeny, strain HTCC2255, which has a streamlined genome, low G+C content (37%), and a paucity of genes for transcriptional regulation, motility, and cell-cell interactions (41). A few uncultured *Roseobacter* single-cell amplified genomes (SAGs), although evolutionarily divergent from the basal HTCC2255 strain, also have streamlined genomes and low G+C content (39% ± 1%) (42).

Despite this broad genomic diversity, roseobacters comprise a well-supported clade in the *Alphaproteobacteria* tree. The strong evolutionary coherence of roseobacters originally discovered by 16S rRNA gene analysis (40) has been validated recently by applying phylogenomic approaches on concatenated protein data sets consisting of conserved single-copy orthologous sequences (37, 43). Concern over potential systematic errors in the data set (for example, due to the extreme differences in G+C content among *Alphaproteobacteria* taxa) motivated the use of multiple models and computational approaches, including a model correcting for amino acid compositional bias (44), a model integrating heterogeneity in the amino acid replacement process across sites of a protein alignment (45), a model accounting for variation of the substitution rate at a site across time (46), and a gene partition model considering heterogeneity among genes based on a gene partition framework (47). These approaches all give strong support for a monophyletic clade of roseobacters, as does an analysis of nearly 200 gene trees for orthologous gene families conserved in *Alphaproteobacteria* (H. Luo, P. G. Foster, and M. A. Moran, unpublished data). Therefore, phylogenetic coherence, and not necessarily conserved metabolic capabilities or phenotypic traits, defines a roseobacter.

While the *Roseobacter* clade itself is strongly supported, the within-group phylogeny based on the most commonly used approach of 16S rRNA gene analysis is not (40). One consequence of this is that the current *Roseobacter* taxonomy is at odds with the phylogeny (37, 41–43). For instance, the genera *Ruegeria*, *Roseobacter*, *Phaeobacter*, and *Oceanicola* are clearly paraphyletic lineages. This noncoherent assignment of scientific names can cause confusion when interpreting the ecology and evolutionary biology of the group.

ECOLOGICAL AND EVOLUTIONARY GENOMICS OF CULTIVATED ROSEOBACTERS

Genome Content

Roseobacter genome content has been addressed in several reviews and research papers (9, 37, 48, 49), and here we update the current status by including ~15 genome sequences that have recently become publically available. Based on 105 single-copy genes that are present in all the *Roseobacter* closed genome sequences (37), the currently sequenced *Roseobacter* isolates have at least 91% coverage and a median of 100% coverage (42).

An unusual genetic capability characteristic of roseobacters is the alphaproteobacterial gene transfer agent (GTA), first discovered in *Rhodobacter capsulatus* (RcGTA) (50). The RcGTA consists of 15 to 17 cotranscribed genes, some of which show homology with phage genes (51). While phage particles package complete phage genomes with occasional inclusion of host bacterial DNA, GTA particles randomly package bacterial DNA only (52). They have been shown to mediate high rates of gene transfer, as much as 10^6-fold higher than estimated transformation and transduction rates (53). Among marine bacterial lineages, the *Roseobacter* clade is the only one to carry GTA genes, and our survey of 52 cultured *Roseobacter* genomes showed that 46 of them contain a complete or nearly complete set (Fig. 1). Three clade members have been experimentally demonstrated to produce RcGTA-like particles, including *Ruegeria pomeroyi*, *Roseovarius nubinhibens*, and *Ruegeria mobilis* (53, 54). Diverse *Roseobacter*-like GTA genes can be amplified from estuarine waters (161), yet they are rarely observed in ocean metagenomes where amplification is not employed, including the Global Ocean Sampling (GOS) data set (54). Thus, the abundance and distribution of GTA genes in marine ecosystems remain unclear, as does their role in mediating lateral gene transfer (LGT) during *Roseobacter* evolution.

Members of the *Roseobacter* clade have a variety of mechanisms for obtaining energy. In addition to heterotrophic energy acquisition through the oxidation of organic matter, genome analysis indicates that some members are also capable of phototrophy. Light utilization involving bacteriochlorophyll *a* synthesized by aerobic anoxygenic phototrophic (AAP) roseobacters is found in 14 phylogenetically diverse strains (Fig. 1). Light utilization based on proteorhodopsin is found only in the streamlined strain HTCC2255 (Fig. 1), and that based on xanthorhodopsin is present in three strains (Fig. 1). For all phototrophic roseobacters, genes for carbon fixation are absent, and thus, these processes may produce energy but do not directly provide fixed carbon; there has been speculation, however, of heightened anaplerotic pathways fueled by light (55). Nonobligate chemolithotrophy is also a common signature in *Roseobacter* genomes. Twenty-four of the 52 strains have the form I version of carbon monoxide dehydrogenase (Fig. 1), and CO oxidation has been demonstrated experimentally for several strains (36, 56). Thirty-six genomes carry the *sox* genes that mediate the oxidation of sulfide or thiosulfate (57) (Fig. 1) and may also provide a source of energy. Among the two types of dissimilatory nitrate reductases (periplasmic *napA* and cytoplasmic *narG*) and dissimilatory nitrite reductases (*nirK* and *nirS*), about half of the 52 strains contain at least one of these genes (Fig. 1). The taxonomic distribution of these chemolithotrophic gene sets does not map coherently to species phylogeny, suggesting that gain and loss of these traits have occurred multiple times during *Roseobacter* evolution, consistent with the scenario that roseobacters have continuously explored new ecological habitats. A few of these ecologically relevant genetic traits, including complete photosynthetic gene clusters (58), capabilities for production of the antibiotic tropodithietic acid (TDA) (59, 60), and type IV secretion systems (61), are found on plasmids, suggesting that extrachromosomal DNA could be an important mechanism by which genes useful for environmental adaptation are transferred among roseobacters.

Genome Evolution

By using a likelihood-based ancestral genome content reconstruction method (62), gene families summarizing all protein-coding genes in 39 *Roseobacter* genomes and 26 other genomes representing major lineages in the *Alphaproteobacteria* were mapped to their ancestral nodes in the *Alphaproteobacteria* tree (41). In general, early evolution in the *Roseobacter* clade was predicted to experience a net genome reduction from a large common ancestral genome, followed by two episodes of genome innovation and ex-

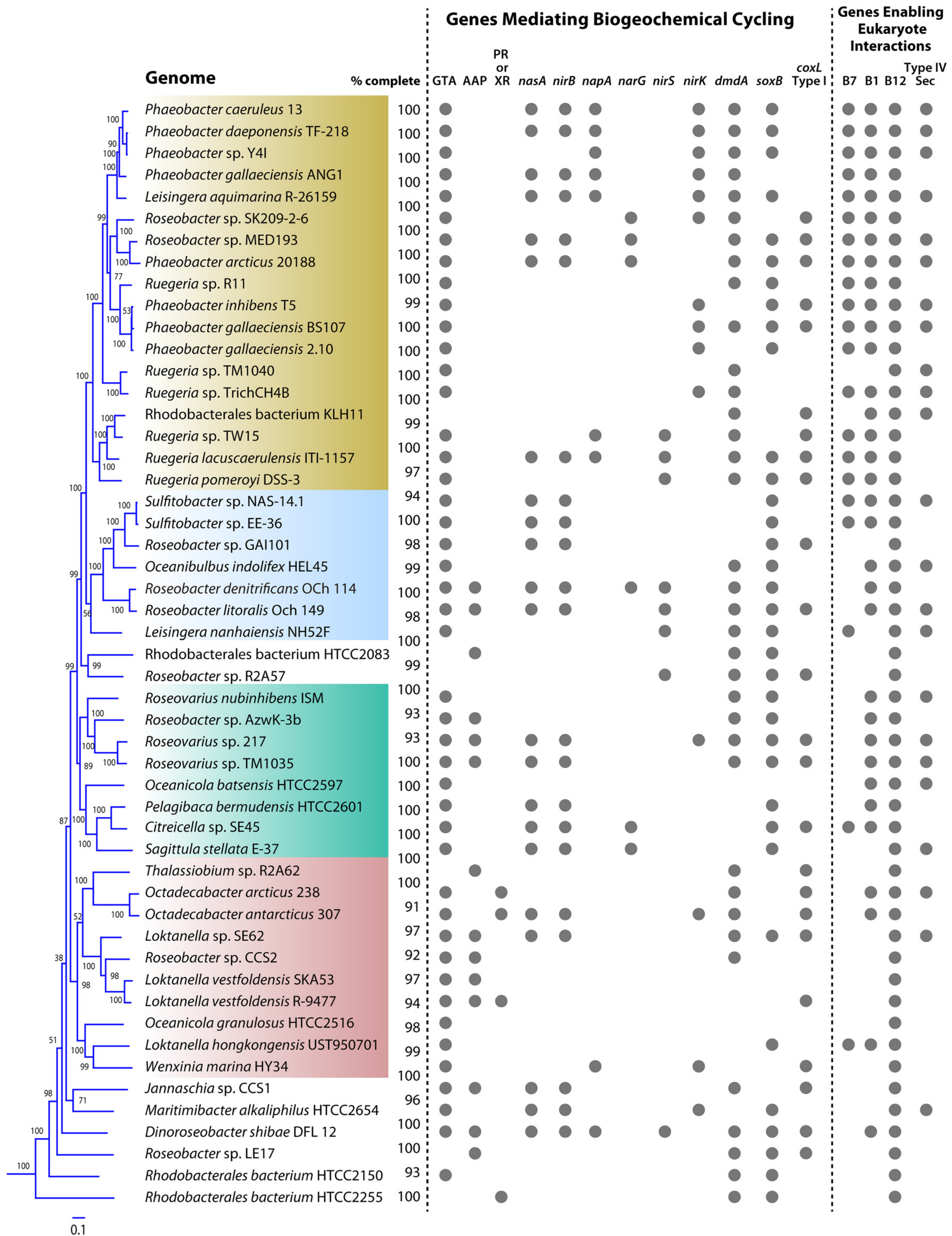

FIG 1 Survey of select genes and metabolic pathways in 52 *Roseobacter* isolate genomes. % complete, estimate of genome completeness; GTA, gene transfer agent; AAP, aerobic anoxygenic photoheterotrophy; PR, proteorhodopsin; XR, xanthorhodopsin; *nasA*, assimilatory nitrate reductase; *nirB*, assimilatory nitrite reductase; *napA*, periplasmic dissimilatory nitrate reductase; *narG*, dissimilatory nitrate reductase; *nirS*, dissimilatory nitrite reductase; *nirK*, dissimilatory nitrite reductase; *dmdA*, dimethylsulfoniopropionate demethylase; *soxB*, sulfur oxidation gene; *coxL* type I, carbon monoxide oxidation; B7, biotin synthase; B1, thiamine synthase; B12, cobalamin synthase; Type IV Sec, type IV secretion system. Colors indicate four major clades of isolate genomes. The phylogenetic tree was constructed based on a concatenation of ~50 single-copy conserved protein sequences using the RAxML software.

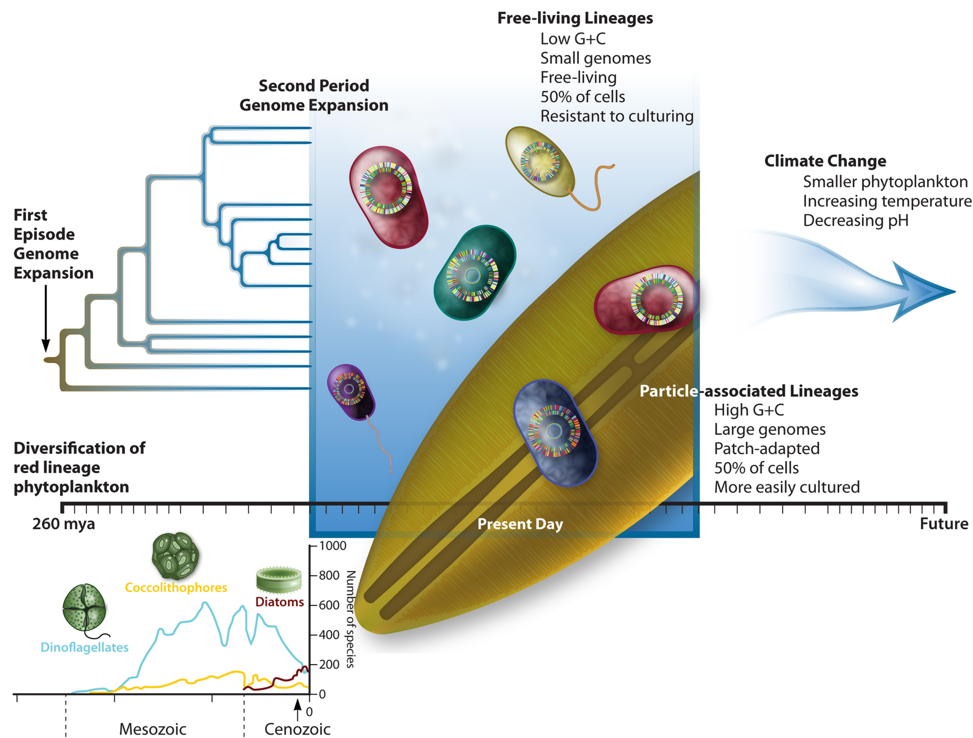

FIG 2 Evolutionary time line of the *Roseobacter* lineage. (Left) The first predicted episode of *Roseobacter* genome expansion coincided with the radiation of marine dinoflagellates and coccolithophorids ~250 mya. The diversification of diatoms occurred more recently. (Middle) Free-living and particle-associated roseobacters cooccur in the surface ocean, each with a distinct set of traits. (Right) In a future ocean, greater dominance by phytoplankton lineages with smaller cell sizes could lead to a decrease in particle-associated roseobacter populations.

pansion through lateral gene transfer (41) (Fig. 2). Although other authors analyzing single-gene trees suggested that roseobacters might have arisen more anciently (63), molecular dating analysis based on the currently available genomes and using a relaxed molecular clock model predicts that the *Roseobacter* ancestor existed around 260 million years ago (mya) (Fig. 2). A first predicted episode of genome expansion shortly after the emergence of this *Roseobacter* ancestor shows that it would have coincided with the rise of two red-plastid-lineage eukaryotic phytoplankton groups, the dinoflagellates and coccolithophorids, around 250 mya (41) (Fig. 2).

One explanation for the predicted coincidence of *Roseobacter* genome expansion and eukaryotic phytoplankton diversification is that the eukaryotes provided new ecological habitats for ancestral roseobacters (41). Indeed, modern lineages of roseobacters are abundant and consistent components of the phycosphere community of these phytoplankton groups (12, 15–17, 64–67). The predicted occurrence of genes involved in motility and chemotaxis in the ancestor of the *Roseobacter* clade would have potentially allowed cells to sense and swim toward phytoplankton (41), enabled by a cell size of red-plastid-lineage phytoplankton large enough to be detected by chemotaxis (68); earlier-evolving phytoplankton groups dominated by the cyanobacteria and green algal lineages were considerably smaller. Along with relationships with dinoflagellates and coccolithophorids, some lineages in the *Roseobacter* clade have been found to be consistently associated with marine diatoms (13, 14), another large-celled, red-plastid-lineage phytoplankton group that diversified somewhat later. This coincidence of *Roseobacter* genome innovation with the radiation of red plastid phytoplankton is consistent with adaptive evolution, but the underlying population genetic mechanism of genome change is not clear and will be difficult, if not impossible, to test for such an ancient event. Genome changes may have been dominated by exaptations (69, 70), in which case changes occurred by chance prior to roseobacters encountering the environment in which they proved useful. Alternatively, genome evolution may have been dominated by positive selection, as has been suggested for other marine bacterioplankton clades (71), in which case environmental change was followed by LGT events, which were then selectively favored.

A second episode of genome innovation during the evolution of the *Roseobacter* clade is predicted to have occurred more recently (Fig. 2), based on elevated LGT rates calculated for several leaf branches (41). The basal lineage consisting of reduced genomes related to *Roseobacter* member strain HTCC2255 appears to have escaped both of these expansion events and become streamlined directly from the common ancestor of the clade (41). Thus, not all extant roseobacters are predicted to be descendants of the ancestral lineages that experienced genome innovation.

The evolutionary mechanisms underlying the diversification of roseobacters can be compared to those driving the evolution of another dominant surface ocean lineage that shares a common ancestor in the *Alphaproteobacteria*, the SAR11 clade. In contrast to *Roseobacter* genomes, SAR11 genomes are uniformly small, have low G+C content (~30%), and harbor limited metabolic capabilities (72, 73). Ecologically, the roseobacters have been proposed to represent a patch-adapted ecological strategy that takes advantage of seawater microenvironments with elevated nutrient

concentrations, while members of the SAR11 clade are considered to use a free-living planktonic strategy adapted to nutrient-depleted bulk seawater (36, 72, 74). The SAR11 clade may also have a much longer evolutionary history, inhabiting the ocean for >800 million years (41). In contrast to roseobacters, there is currently no evidence for major genome innovations in SAR11, and genome size changes over the evolution of this group suggests that it evolved at a lower rate than did the roseobacters (41). The common ancestor of the SAR11 clade was predicted to have contained only ~2,000 genes (41). For the *Roseobacter* clade, gene gain appears biased toward transcriptional regulators, replication/recombination/repair genes, and defense mechanism genes; such traits might be important for roseobacters to compete well with cooccurring microbial populations on particles and living surfaces in the marine environment. For SAR11, gene gain is predicted to have been biased toward genes for cell wall biogenesis and pilus synthesis (41).

ECOLOGICAL ASSOCIATIONS BETWEEN ROSEOBACTERS AND EUKARYOTES

Roseobacter-Phytoplankton Interactions

Roseobacter-phytoplankton interactions are likely to be based on acquiring organic matter on the bacterial side and obtaining essential vitamins and regenerated inorganic nutrients on the phytoplankton side (24). Roseobacters may also function as probionts that deter algal pathogens (24). As these interactions are potentially relevant for explaining evolutionary relationships between roseobacters and phytoplankton, we summarize known interaction mechanisms and chemical exchanges. Both genomic and experimental evidence suggests that some interaction mechanisms are common to most members of the *Roseobacter* clade and most types of eukaryotic phytoplankton, while others are limited to specific phylotypes of each. General attributes that may be useful in interactions with eukaryotic organisms include motility and chemotaxis (49, 75), type IV secretion systems (9, 76), quorum-sensing systems (77), and probiotic biosynthesis (78).

Among the list of organic compounds released by phytoplankton that have been shown experimentally to be assimilated by roseobacters is the two-carbon compound glycolate, which is excreted during autotrophic photorespiration (79, 80) and can comprise 10 to 50% of the phytoplankton exudate in marine environments (80–82). Roseobacters are among a select group of marine bacterial lineages capable of metabolizing this compound (83, 84). Carbohydrates are also exuded by phytoplankton, with the composition possibly varying among lineages. *Roseobacter* clade members not only grow on these carbohydrates (e.g., *Planktotalea frisia*) but also may grow better on exudates produced by certain species (85), raising the possibility of an adaptive association of roseobacters with specific phytoplankton lineages. The organic sulfur compound dimethylsulfoniopropionate (DMSP) is produced in abundance by dinoflagellates and coccolithophorids (12, 17, 67), two phytoplankton groups often found associated with roseobacters in ocean waters (15, 16, 64–66). DMSP acts as a specific chemical cue that attracts motile and chemotactic bacteria, including roseobacters (86). The fact that the *Roseobacter* clade is only one of two marine bacterial lineages harboring both of the known pathways for DMSP degradation (the other being the SAR116 clade) (87) suggests that this compound may be a particularly important currency in *Roseobacter*-phytoplankton interactions (75). Finally, polyunsaturated aldehydes (PUAs) produced mainly by diatoms may selectively inhibit *Roseobacter* members, with evidence for poor growth in the presence of PUAs for strains related to *Roseobacter litoralis* and *Phaeobacter gallaeciensis* (88).

Interactions based on the release of metabolites are not a one-way street, as roseobacters may also release compounds that affect phytoplankton physiology. A recent study showed that exudates from the *Roseobacter* clade member *Dinoroseobacter shibae* stimulate the metabolism of the diatom *Thalassiosira pseudonana*, including enhanced production of picolinic acid, which can form metal complexes with limiting trace metals such as iron (89). The ability to synthesize the soluble vitamins B_1 (thiamine), B_7 (biotin), and B_{12} (cobalamin) is often lacking in phytoplankton, with ~50% of >300 eukaryotic phytoplankton species being auxotrophic for vitamin B_{12}, ~25% being auxotrophic for vitamin B_1, and 8% being auxotrophic for vitamin B_7 (90). Every one of the 52 *Roseobacter* genomes has evidence of a functioning vitamin B_{12} biosynthesis pathway (90) (Fig. 1), even though >60% of marine bacterial species appear to be unable to make this vitamin. Over half of the sequenced *Roseobacter* strains also have genes for biosynthesis of vitamins B_1 and B_7 (Fig. 1). Whether metabolite exchanges between marine bacteria and phytoplankton rise to the level of symbiosis is a matter of debate (91, 92), but species-specific mutualism is not required for such interactions to have evolutionary consequences.

On the pathogenic side of *Roseobacter*-phytoplankton interactions, *Phaeobacter gallaeciensis* has been shown to produce an algicidal compound against senescent *Emiliania huxleyi* populations and may help to terminate blooms (27, 28). A similar pathogenic function was also found for a *Roseobacter* clade-affiliated (RCA) cluster strain, LE17, which demonstrated algicidal activity against dinoflagellates (93). A set of genes potentially noteworthy for interactions with phytoplankton are those of the type IV secretion system, noticed in one of the first *Roseobacter* genome sequences (36) and now known to be present in nearly half of the strains surveyed (Fig. 1). The roles of these systems in roseobacters have yet to be deciphered, although studies of other model bacteria (e.g., *Agrobacterium*) showed that they mediate transfer of DNA and proteins to eukaryotic cells and thereby modulate physiological processes in the recipient organism (94).

Roseobacter-Coral Interactions

Roseobacters are also known to associate with a variety of marine macroalgae and animals, including green and red seaweed (18, 19), sponges (95), squid (22), fish larvae (96), sea urchins (21), and oysters (26, 97). A few of them have been demonstrated to induce physiological changes in the host, both positive and negative, including gall disease in the red alga *Prionitis lanceolata* (98), juvenile oyster disease (99), infection of the red alga *Delisea pulchra* (100), and inhibition of the settlement of fouling organisms on the green alga *Ulva australis* (19). The most extensive research, however, has focused on associations with corals (101–107), as roseobacters can account for up to 50% of the coral mucus microbiome (108–110). It has been hypothesized that DMSP and dimethylsulfide (DMS) are the two compounds that structure the community of coral-associated microbes, since isolates from these populations are enriched with the ability to degrade DMSP/DMS (111). As is the case for *Roseobacter* interactions with other eukaryotes, it is not yet clear whether a physical association signifies a specific mutualistic interaction. There is evidence that roseobac-

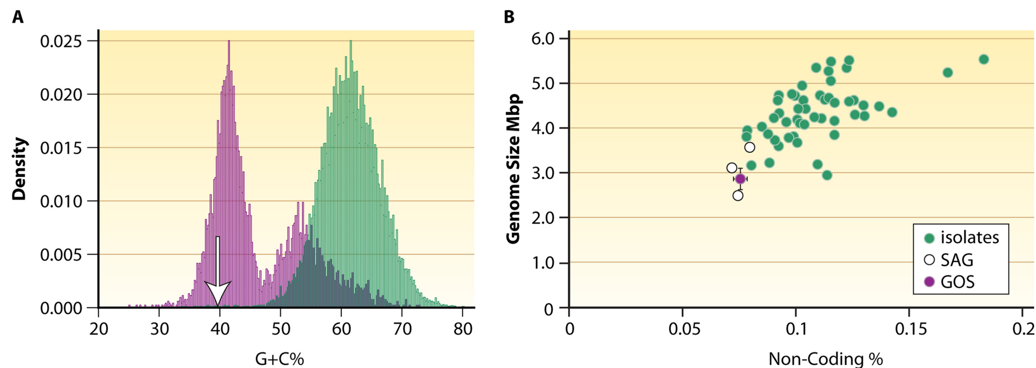

FIG 3 Comparison of characteristics of cultured *Roseobacter* genomes, single-cell *Roseobacter* genomes, and *Roseobacter* sequences in the GOS metagenome. (Left) Distribution of G+C content. The arrow represents the mean value for the SAG genomes. (Right) Estimated average genome size plotted against the fraction of noncoding DNA.

ters may play a role in coral reproduction, based on both consistent associations between the bacteria and early developmental stages of corals (112, 113) as well as increases in *Roseobacter* abundance after coral mass spawning (114). Roseobacters may also have a role in coral bleaching by acting as probiotics against pathogens (115, 116). Some *Roseobacter* lineages have been found to be significantly associated with lesioned corals (106, 117, 118). Whether interactions with corals are specific or general and whether the outcomes are beneficial or pathogenic, the associations are nonetheless likely to have shaped aspects of *Roseobacter* evolution.

ECOLOGICAL AND EVOLUTIONARY GENOMICS OF UNCULTIVATED ROSEOBACTERS

Genome Analyses

Many major *Roseobacter* lineages, including those that are most numerically dominant in the ocean, do not yet have cultivated members. Metagenomic data sets are valuable resources for studying these uncultivated roseobacters, as long as the taxonomic classification of metagenomic reads is unequivocal. Unfortunately, the widely applied approach to metagenomic analysis of using BLAST best hits can produce an unacceptable level of false-positive results due both to uncorrected genetic distances of similarity scores and the biased taxonomic composition of the database (119). *De novo* genome assembly from deeply sequenced metagenomes is one way around this problem, and indeed, a near-complete *Roseobacter* genome closely related to strain HTCC2255 was assembled from a sample from coastal seawater near Seattle, WA (120). When roseobacters are not dominant or when the sequencing depth of a metagenome is shallow, however, an alternate approach uses the computed patterns in d_N, the number of nonsynonymous (amino-acid-changing) substitutions per nonsynonymous site. A two-step bioinformatics procedure first assigns metagenomic gene fragments to a clade based on the values of d_N computed between the metagenomic gene fragment and its orthologs in cultured members and then uses the paired-end read of these conservatively assigned metagenomic sequences as a homology-independent survey of biological function (43). Applied to the ~7 million Sanger reads in the 2007 GOS metagenomic data set, the method identified ~3,000 unambiguous *Roseobacter* reads exhibiting several systematic differences in genome content compared to their cultured counterparts. The metagenomic *Roseobacter* data set had fewer genes for signal transduction and cell surface modifications but more genes for Sec-like protein secretion systems (43). Interestingly, these gene functions have been suggested to be signatures for distinguishing patch-adapted from free-living ecological strategies in marine bacteria (121, 122). One conclusion arising from the analysis of metagenomic data is that genomic analyses focusing on cultured roseobacters is biasing our view of the lineage's ecology (43).

Another approach to understanding *Roseobacter* biology without the bias of cultivation is the use of single-cell genome sequencing. Recently, four SAGs of uncultivated roseobacters were obtained from surface waters of the North Pacific, South Atlantic, and Gulf of Maine (123). Phylogenomic analyses showed that 3 of the SAGs comprise a novel clade (the SAG-O19 clade) in which no cultured representatives have been found. Moreover, this clade appears to represent up to 35% of the *Roseobacter* sequences in samples from surface ocean waters in the GOS (42). These SAGs have low G+C content (39% ± 1%) and a reduced percentage of noncoding DNA (7.5% ± 0.4%) and are predicted to have streamlined genomes (2.6 to 3.5 Mbp). These characteristics are consistent with those of the oceanic roseobacters identified by the *dN* pipeline, and they show the same systematic divergences from cultured strains (42) (Fig. 3).

Evolution of an Enigmatic *Roseobacter* Lineage

As the SAG-O19 clade represents an abundant extant lineage with low G+C content, it may be informative to consider how the evolutionary forces affecting this clade compare to mechanisms invoked for other G+C-poor bacterial genomes. In a situation where synonymous (silent) nucleotide sites are saturated with substitutions, as is the case for the three SAG-O19 clade members, an alternate population genetic approach based on the ratio of radical to conservative nonsynonymous nucleotide substitution rates (d_R/d_C ratio) can be used instead, based on classification of amino acids according to physicochemical properties such as charge (positive/negative/neutral) (124, 125). By using this approach, a higher mean d_R/d_C ratio was found for >500 orthologous genes in the uncultivated SAG-O19 clade than in other *Roseobacter* lineages. Theory predicts that such inflated d_R/d_C ratios occur when purifying selection is not efficient and thus argues that genetic drift may have played a prominent role during the evolution of the SAG-O19 clade (42). Such an interpretation might be

problematic if the selective pressure to decrease G+C content and, hence, nitrogen (N) usage in the proteome (126) interferes with the selective pressure for conservation of amino acid physicochemical properties. In other words, the fitness cost of radical changes could be balanced by the benefit of using less N in both genomes and proteomes (127). However, two observations that cast doubt on the idea that low G+C content is under particularly strong selection in marine bacteria are that freshwater relatives of the low-G+C-content SAR11 clade similarly have low-G+C-content genomes, even though N is not limiting in freshwater environments (128), and that both low-G+C- and high-G+C-content lineages coexist in many major marine clades, including not only the roseobacters but also the SAR116 lineage and several marine *Gammaproteobacteria* groups.

Based on recent studies showing that mutational bias from G/C to A/T is universal among bacterial genomes (129–132), one explanation for the observed paucity in G+C-rich codons in genomes of the SAG-O19 clade is that an inefficiency of purifying selection under such a mutational bias would accelerate the replacement of G+C-rich codons. Thus, although counter to the more established theory that selection has been the exclusive force driving the evolution of surface ocean bacterial populations (8, 133), there are also reasonable arguments in favor of the inefficiency of purifying selection playing a role, at least in the evolution of SAG-O19 genome composition. Regardless of the mechanism, it is clear that different evolutionary processes have been at work in the SAG-O19 clade compared to its cultured *Roseobacter* relatives.

Inclusion of Uncultured Roseobacters in Phylogenomic Analyses

With the availability of genomic sequences of single cells, uncultivated bacterial and archaeal lineages have been regularly placed in phylogenetic trees along with their cultured relatives (134, 135). However, when the three uncultivated genomes in the SAG-O19 clade (SCGC AAA015-O19, SCGC AAA298-K06, and SCGC AAA300-J04) are included in phylogenetic trees based on standard approaches, such as multiprotein concatenation using maximum likelihood (134, 135), the branching order of several major clades is no longer supported (42). The substantially different G+C contents of the 3 SAGs compared to those of most isolates (39% ± 1% versus 60% ± 4%) may be one cause of this problem, a phenomenon known as nonstationarity (136–140). Indeed, posterior predictive simulation assessing the fit of homogeneous versus heterogeneous models to each data set (i.e., models that do not account for compositional heterogeneity in G+C content versus those that do) showed that half of the orthologs used for tree building significantly violated the composition-homogeneous assumption. Since many standard phylogenetic programs cannot account for compositional heterogeneity, the node-discrete composition heterogeneity (NDCH) and node-discrete rate matrix heterogeneity (NDRH) models that allow for variation in G+C contents and rate matrices across branches were used instead (44). These models resolved the 3 SAGs into a well-supported clade while maintaining the evolutionary relationships of cultured lineages identified previously (42). An understanding of the evolutionary history of the marine *Roseobacter* clade requires the inclusion of the uncultured members, even if they pose computational challenges, particularly given the genomic evidence that they represent very different organisms and life history strategies. It can be envisioned that this nonstationary phylogenetic model will become increasingly useful as additional SAGs fill in our understanding of *Roseobacter* diversity and ecology.

Genome Streamlining in the *Roseobacter* Clade

Genome streamlining is a well-known process in obligate intracellular bacteria. For these taxa, the primary mechanism of gene loss is thought to be relaxation of purifying selection on genes for which host proteins perform the same biological function, rendering mutations selectively neutral (141). Additionally, obligate intracellular bacteria are assumed to have small effective population sizes as a result of frequent population bottlenecks, leading to fixation of slightly deleterious mutations through genetic drift (142, 143). Genome streamlining is also recognized as a prevalent feature among diverse lineages of free-living planktonic marine bacteria (123). A few prominent examples are the alphaproteobacterial SAR11 clade (72), the gammaproteobacterial SAR86 clade (144), and the cyanobacterial *Prochlorococcus* clade (145, 146). The primary evolutionary mechanisms of streamlining proposed for marine bacteria is selection for metabolic efficiency (72, 147). Secondarily, these bacteria also typically have reduced cell sizes and thus can maximize surface-to-volume ratios to facilitate nutrient acquisition (72, 121). However, the current theory cannot explain why bacteria with average-sized genomes cooccur with bacteria with streamlined genomes in oligotrophic seawater and even represent dominant groups in some ocean habitats. This "paradox of genome streamlining" has been perplexing (73).

Although the marine *Roseobacter* clade was initially thought to contain mainly members with large genomes, both metagenomic approaches and single-cell genome sequencing indicate that for many oceanic roseobacters, this is not the case (42). Genome streamlining is predicted to have occurred during the evolution of the basal lineage of the clade, represented by isolate HTCC2255, with a nearly complete genome of 2.54 Mb, and a highly related single-cell genome, with ~80% coverage and an estimated genome size of 2.64 Mb. Ancestral genome content reconstruction and COG functional category analysis showed that HTCC2255 shares some similarities in genetic architecture with the common ancestor of the SAR11 clade, both of which are more similar in gene content to the extant SAR11 genomes than to extant *Roseobacter* genomes (41). This convergent evolution of genetic makeup suggests that the basal HTCC2255 clade is consistent with a free-living planktonic strategy. In the predicted streamlining from the *Roseobacter* common ancestor that led directly to the HTCC2255 clade, genes that were lost were biased toward functions involving cell-cell interactions, transcriptional regulation, and motility (41). While these biological functions are likely essential for cells living on surfaces or switching between free-living and surface-associated strategies, they may become dispensable and expensive to maintain when cells grow singly in bulk seawater. One explanation for functionally biased gene loss could be inefficiency of purifying selection due to restriction of habitats (population bottleneck) or exploration of new habitats (founder effect), resulting in the preferential loss of nonessential genes. While comparisons of ancestor-descendant genome contents have yielded useful information regarding the evolutionary mechanism giving rise to the basal HTCC2255 clade, a fuller resolution of the evolutionary path of this basal lineage during the past 260 million years is lacking.

In the case of the SAG-O19 clade, all three genome sequences

were only partially recovered during single-cell sequencing, yet several independent lines of evidence support genome streamlining. First of all, genome size estimates based on single-copy genes suggest that SAG-O19 clade members are smaller than most cultured roseobacters (see above). Second, nonobligate endosymbiotic bacterial genome reduction often occurs with a reduction in the percentage of noncoding DNA (72), a feature which is also observed for these SAGs (Fig. 3). Finally, gene content characteristics attributed to streamlined marine bacterial genomes are found in the SAG-O19 clade, including an underrepresentation of genes involved in transcriptional regulation and replication/recombination/repair (41, 121, 122) relative to roseobacters with larger genomes. Since a substantial part of the SAG genomes was not sequenced (45 to 77%), it is more challenging to reconstruct the evolutionary path giving rise to this clade. If inefficiency of purifying selection drove the evolution of the HTCC2255 clade, as suggested above, it is plausible that it also played an important role in genome streamlining of this clade. Interestingly, the *Roseobacter* R11 lineage that has been observed to live endosymbiotically with the marine macroalga *Delisea pulchra* has not undergone genome reduction, having a genome size of 3.93 Mb (42). This is expected, however, since as the bacterium is facultative symbiont, it must maintain a genome repertoire sufficient for independent living.

A New Look at Population Size

There are two ways of conceptualizing the population size of organisms: census population size (N_c), which is a headcount of a population and has significance from the ecological standpoint, and effective population size (N_e), which determines the capability for environmental adaptation and has relevance from an evolutionary viewpoint. The widely used GOS metagenomic data from ocean surface waters, which provide taxonomic information free of the biases inherent in PCR-amplified 16S rRNA gene surveys, suggest a lower N_c for roseobacters (4% of surface ocean bacteria) than for the SAR11 clade (31%) and several other bacterial taxa (SAR86, 6%; *Bacteroidetes*, 6%; *Actinobacteria*, 8%) (148). However, the genomes sequenced in the GOS were primarily from small bacteria (i.e., those falling in the size range of 0.2 to 0.8 μm in diameter) at the expense of larger and particle- or host-associated cells, instituting a bias that may affect estimates of census population size. To better address the N_c of the *Roseobacter* clade, we analyzed new marine metagenomic data from the Southern Ocean (149) and Monterey Bay (A. Z. Worden, unpublished data), in which both free-living (0.1- to 0.8-μm) and particle-associated (0.8- to 3.0-μm) communities were sequenced. Based on queries of >2,000 orthologous gene families, the metagenomic data sets showed that roseobacters accounted for a higher percentage of the genomes in particle-associated communities than in free-living communities by >2-fold overall (Fig. 4), with 11 of 13 Southern Ocean samples and 3 of 4 Monterey Bay samples following this pattern ($P < 0.001$ by χ^2 test). Thus, calculation of the census population size of the *Roseobacter* clade may be a sizeable underestimate in cases where sampling protocols systematically exclude larger and particle-associated cells.

From an evolutionary perspective, it is the N_e that determines whether a deleterious mutation can be effectively purged from the population and whether a favorable mutation can be effectively spread, two major ways in which populations adapt to a changing environment (150). In a population with a low N_e value, there is a greater chance that deleterious mutations drift to fixation while favorable mutations drift to loss, both of which reduce fitness. Calculation of the absolute values of N_e requires knowledge of the mutation rate for a given population, but this has not yet been measured for any marine bacterial population (151). Streamlined marine bacteria that grow slowly and do not form colonies on solid media pose a particular challenge for approaches that rely on direct measurements of the mutation rate for the calculation of N_e (151).

Alternatively, population genetic theory provides a method for comparing the relative values of N_e based on the expectation that a population with a reduced N_e will have an increasing number of slightly deleterious mutations that drift to fixation (152). Since nonsynonymous mutations are more likely to be slightly deleterious, it is expected that the ratio (ω) of the number of nonsynonymous substitutions per nonsynonymous site (d_N) to the number of synonymous substitutions per synonymous site (d_S) is greater in populations with a lower N_e value (153). Applying this approach to surface ocean bacteria, comparisons of ω were made for >400 orthologous genes among cooccurring *Alphaproteobacteria* populations. *Roseobacter* lineages exhibited a range of relative N_e values, including some populations whose substitution patterns produced higher calculated N_e values than those for SAR11 populations (151). The apparent mismatch between N_e and N_c is accounted for to some extent by the undersampling of lineages that have larger cells and/or are associated with particulate material (Fig. 4). Other possible factors include variability in the fraction of cells that are inactive and the population history of the lineages (e.g., population bottlenecks) (154).

Overall, neither census population size nor life history strategy is a good predictor of effective population sizes, and direct measures of N_e are ultimately necessary to understand the ability of *Roseobacter* populations to adapt to a changing ocean. Under the neutral theory model, the nucleotide diversity at synonymous sites in a bacterial population equals twice the product of its effective population size and spontaneous mutation rate (per generation). Thus, one way to directly estimate the absolute value of N_e is to determine the mutation rate through a mutation accumulation experiment in which multiple parallel cell lines are bottlenecked to a single cell regularly over several thousand generations, as has been done for a few model bacteria (155) but not for any that are marine. Roseobacter lineages can serve as model organisms for this experiment, since many grow readily on solid media and are amenable to experimental evolution studies.

CONCLUDING REMARKS

Although the *Roseobacter* clade is among the most frequently sampled of surface ocean bacterial lineages, it is perhaps one of the least known due to the enormous diversity of clade members with regard to habitats and metabolic capabilities. Over 2 decades of research on this lineage have established its importance in global carbon and sulfur cycles. Many of the attributed biogeochemical functions appear to be linked to the ability of group members to physically associate with phytoplankton, corals, and other eukaryotes and affect eukaryote function as mutualists, probionts, and pathogens. How climate change may diminish or strengthen the abundance and activity of *Roseobacter* lineages is currently unknown.

Following the emergence of the *Roseobacter* ancestor, estimated to have appeared about 260 mya based on analysis of cur-

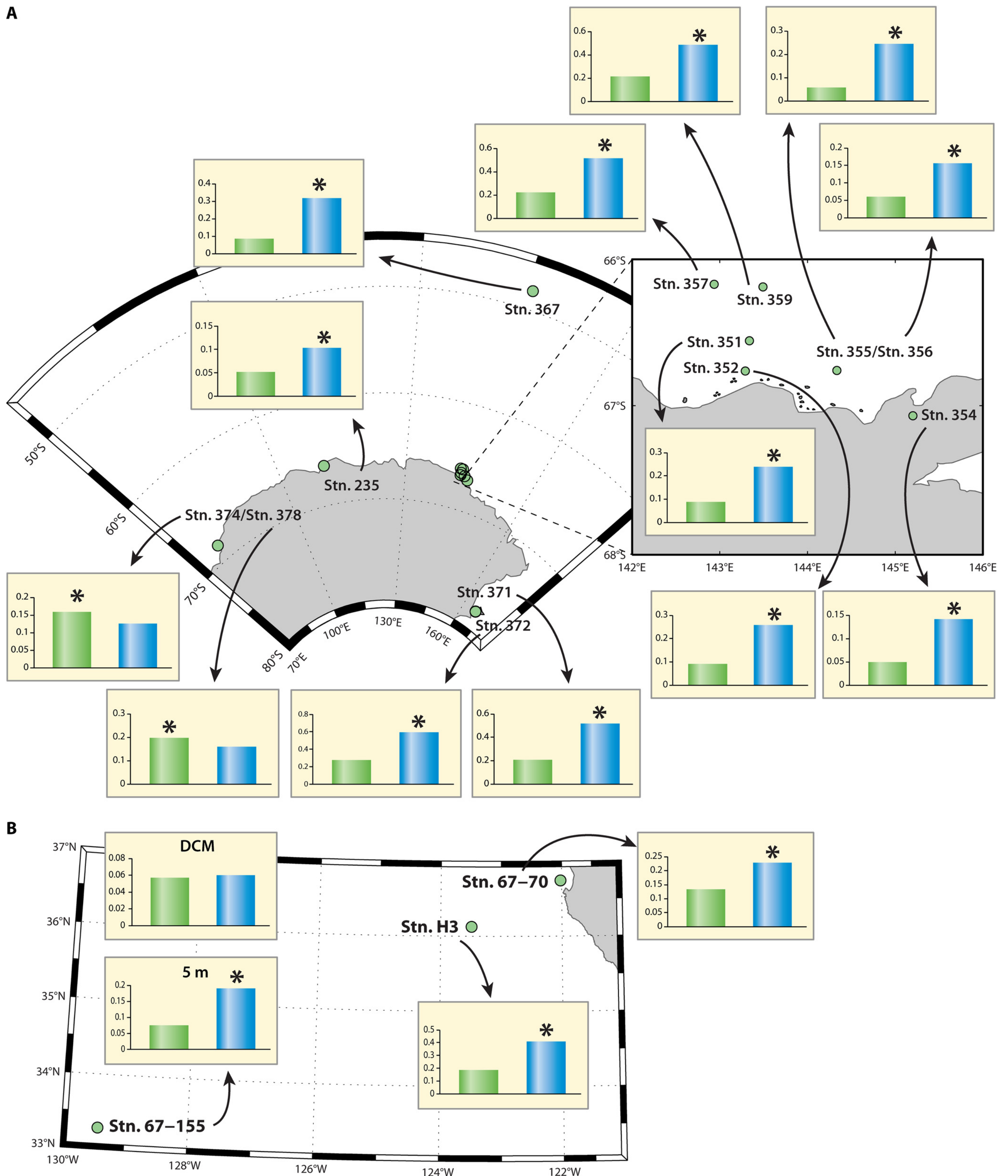

FIG 4 *Roseobacter* sequences in metagenomic data sets, calculated as a percentage of free-living (0.1 to 0.8 μm) (green bars) and particle-associated (0.8 to 3.0 μm) (blue bars) reads in DNA from the Southern Ocean (13 stations [Stn.]) and Monterey Bay, CA (4 stations). Asterisks indicate significant differences between size fractions ($P < 0.05$). Data were obtained by analyzing 2,235 orthologous gene families shared by at least 20 of the 40 *Roseobacter* isolate genomes.

rently available genomes, some lineages are surmised to have experienced adaptive genome expansion through LGT. The radiation is predicted to have coincided in time with the emergence of dinoflagellates and coccolithophorids, and these phytoplankton groups may have provided new habitats and new roles in oceanic elemental cycles. In addition to eukaryotic phytoplankton, a number of other marine eukaryotic organisms are associated with bacterial communities in which roseobacters dominate, although it remains unknown whether, when, and how these eukaryotes may have imprinted on the evolution of roseobacters. Addressing this question will require the availability of more *Roseobacter* genome sequences associated with specific eukaryotes and an improved understanding of the biology and ecology underlying *Roseobacter*-eukaryote interactions.

A few *Roseobacter* lineages adopted a free-living planktonic strategy or are evolving toward one. Representing the most abundant *Roseobacter* lineages among the free-living ocean bacteria, these groups appear to have experienced an evolutionary stage in which effective population sizes were reduced compared to those of other *Roseobacter* lineages. Their evolutionary modifications, such as reductions in genome size and G+C content, may therefore have occurred under relaxed selective constraints. Because of their distinct gene repertoire, these planktonic roseobacters are postulated to play roles in marine elemental transformations that differ from those of their eukaryote-associated relatives.

Molecular evolutionary approaches promise to generate insights into the history and mechanisms underlying the ecological divergence of dominant surface ocean bacterioplankton (41–43, 151, 156–159), with the caveat that different evolutionary approaches will have different time scales of resolution. We are just starting to appreciate the evolutionary patterns of the *Roseobacter* clade as a whole, and we still have limited knowledge of how the tremendous ecological diversity of individual members evolved. Previous genomic sequencing projects have focused on sampling taxa that diverged anciently, enabling the reconstruction of a nearly complete evolutionary history of roseobacters. Future studies will also obtain genomes of closely related lineages within species, which are required to understand the ecological and evolutionary factors driving ongoing speciation processes. In fact, investigation of strain-level genomic variation has already started for marine *Vibrio* lineages associated with various microscale ecological niches, and significant insights regarding how microbes diversify in the seemingly well-mixed ocean have been obtained (5, 7, 160). Two current constraints on the evolutionary analysis of the *Roseobacter* lineage are the limited number of genome sequences available for each specific ecological niche and the biases associated with focusing on readily cultured representatives. It is anticipated that *Roseobacter* isolates and single cells isolated from samples from various ocean habitats will continue to be sequenced and, along with metagenomic data sets, will fill crucial gaps in our understanding of the evolutionary ecology of this important marine bacterioplankton clade.

ACKNOWLEDGMENTS

We thank A. Z. Worden for providing the unpublished Monterey Bay metagenomic data set and C. English for assistance with graphics.

This research was funded in part by the Gordon and Betty Moore Foundation through grant GBMF538.01 and the National Science Foundation through grant OCE-1259562.

REFERENCES

1. **Giovannoni SJ, Britschgi TB, Moyer CL, Field KG.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature **345:**60–63. http://dx.doi.org/10.1038/345060a0.
2. **Rappé MS, Vergin K, Giovannoni SJ.** 2000. Phylogenetic comparisons of a coastal bacterioplankton community with its counterparts in open ocean and freshwater systems. FEMS Microbiol. Ecol. **33:**219–232. http://dx.doi.org/10.1111/j.1574-6941.2000.tb00744.x.
3. **Suzuki MT, Béjà O, Taylor LT, DeLong EF.** 2001. Phylogenetic analysis of ribosomal RNA operons from uncultivated coastal marine bacterioplankton. Environ. Microbiol. **3:**323–331. http://dx.doi.org/10.1046/j.1462-2920.2001.00198.x.
4. **Shapiro BJ, Polz MF.** 2014. Ordering microbial diversity into ecologically and genetically cohesive units. Trends Microbiol. **22:**235–247. http://dx.doi.org/10.1016/j.tim.2014.02.006.
5. **Cordero OX, Polz MF.** 2014. Explaining microbial genomic diversity in light of evolutionary ecology. Nat. Rev. Microbiol. **12:**263–273. http://dx.doi.org/10.1038/nrmicro3218.
6. **Cordero OX, Ventouras L-A, DeLong EF, Polz MF.** 2012. Public good dynamics drive evolution of iron acquisition strategies in natural bacterioplankton populations. Proc. Natl. Acad. Sci. U. S. A. **109:**20059–20064. http://dx.doi.org/10.1073/pnas.1213344109.
7. **Shapiro J, Friedman J, Cordero O, Preheim S, Timberlake S, Szabó G, Polz M, Alm E.** 2012. Population genomics of early events in the ecological differentiation of bacteria. Science **336:**48–51. http://dx.doi.org/10.1126/science.1218198.
8. **Giovannoni SJ, Cameron Thrash J, Temperton B.** 2014. Implications of streamlining theory for microbial ecology. ISME J. **8:**1553–1565. http://dx.doi.org/10.1038/ismej.2014.60.
9. **Moran MA, Belas R, Schell MA, Gonzalez JM, Sun F, Sun S, Binder BJ, Edmonds J, Ye W, Orcutt B, Howard EC, Meile C, Palefsky W, Goesmann A, Ren Q, Paulsen I, Ulrich LE, Thompson LS, Saunders E, Buchan A.** 2007. Ecological genomics of marine roseobacters. Appl. Environ. Microbiol. **73:**4559–4569. http://dx.doi.org/10.1128/AEM.02580-06.
10. **Suzuki MT, Preston CM, Chavez FP, DeLong EF.** 2001. Quantitative mapping of bacterioplankton populations in seawater: field tests across an upwelling plume in Monterey Bay. Aquat. Microb. Ecol. **24:**117–127. http://dx.doi.org/10.3354/ame024117.
11. **Gonzalez J, Moran M.** 1997. Numerical dominance of a group of marine bacteria in the alpha-subclass of the class Proteobacteria in coastal seawater. Appl. Environ. Microbiol. **63:**4237–4242.
12. **González JM, Simó R, Massana R, Covert JS, Casamayor EO, Pedrós-Alió C, Moran MA.** 2000. Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. Appl. Environ. Microbiol. **66:**4237–4246. http://dx.doi.org/10.1128/AEM.66.10.4237-4246.2000.
13. **Grossart H-P, Levold F, Allgaier M, Simon M, Brinkhoff T.** 2005. Marine diatom species harbour distinct bacterial communities. Environ. Microbiol. **7:**860–873. http://dx.doi.org/10.1111/j.1462-2920.2005.00759.x.
14. **Amin SA, Parker MS, Armbrust EV.** 2012. Interactions between diatoms and bacteria. Microbiol. Mol. Biol. Rev. **76:**667–684. http://dx.doi.org/10.1128/MMBR.00007-12.
15. **Alavi M, Miller T, Erlandson K, Schneider R, Belas R.** 2001. Bacterial community associated with Pfiesteria-like dinoflagellate cultures. Environ. Microbiol. **3:**380–396. http://dx.doi.org/10.1046/j.1462-2920.2001.00207.x.
16. **Jasti S, Sieracki ME, Poulton NJ, Giewat MW, Rooney-Varga JN.** 2005. Phylogenetic diversity and specificity of bacteria closely associated with *Alexandrium* spp. and other phytoplankton. Appl. Environ. Microbiol. **71:**3483–3494. http://dx.doi.org/10.1128/AEM.71.7.3483-3494.2005.
17. **Zubkov MV, Fuchs BM, Archer SD, Kiene RP, Amann R, Burkill PH.** 2001. Linking the composition of bacterioplankton to rapid turnover of dissolved dimethylsulphoniopropionate in an algal bloom in the North Sea. Environ. Microbiol. **3:**304–311. http://dx.doi.org/10.1046/j.1462-2920.2001.00196.x.
18. **Ashen JB, Goff LJ.** 2000. Molecular and ecological evidence for species specificity and coevolution in a group of marine algal-bacterial symbioses. Appl. Environ. Microbiol. **66:**3024–3030. http://dx.doi.org/10.1128/AEM.66.7.3024-3030.2000.
19. **Rao D, Webb JS, Holmström C, Case R, Low A, Steinberg P, Kjelle-**

berg S. 2007. Low densities of epiphytic bacteria from the marine alga *Ulva australis* inhibit settlement of fouling organisms. Appl. Environ. Microbiol. **73:**7844–7852. http://dx.doi.org/10.1128/AEM.01543-07.
20. **Apprill A, Marlow HQ, Martindale MQ, Rappe MS.** 2009. The onset of microbial associations in the coral Pocillopora meandrina. ISME J. **3:**685–699. http://dx.doi.org/10.1038/ismej.2009.3.
21. **Becker PT, Gillan DC, Eeckhaut I.** 2009. Characterization of the bacterial community associated with body wall lesions of *Tripneustes gratilla* (Echinoidea) using culture-independent methods. J. Invertebr. Pathol. **100:**127–130. http://dx.doi.org/10.1016/j.jip.2008.11.002.
22. **Barbieri E, Paster BJ, Hughes D, Zurek L, Moser DP, Teske A, Sogin ML.** 2001. Phylogenetic characterization of epibiotic bacteria in the accessory nidamental gland and egg capsules of the squid Loligo pealei (Cephalopoda: Loliginidae). Environ. Microbiol. **3:**151–167. http://dx.doi.org/10.1046/j.1462-2920.2001.00172.x.
23. **Collins AJ, LaBarre BA, Won BSW, Shah MV, Heng S, Choudhury MH, Haydar SA, Santiago J, Nyholm SV.** 2012. Diversity and partitioning of bacterial populations within the accessory nidamental gland of the squid *Euprymna scolopes*. Appl. Environ. Microbiol. **78:**4200–4208. http://dx.doi.org/10.1128/AEM.07437-11.
24. **Geng H, Belas R.** 2010. Molecular mechanisms underlying roseobacter-phytoplankton symbioses. Curr. Opin. Biotechnol. **21:**332–338. http://dx.doi.org/10.1016/j.copbio.2010.03.013.
25. **Wagner-Dobler I, Ballhausen B, Berger M, Brinkhoff T, Buchholz I, Bunk B, Cypionka H, Daniel R, Drepper T, Gerdts G, Hahnke S, Han C, Jahn D, Kalhoefer D, Kiss H, Klenk H-P, Kyrpides N, Liebl W, Liesegang H, Meincke L, Pati A, Petersen J, Piekarski T, Pommerenke C, Pradella S, Pukall R, Rabus R, Stackebrandt E, Thole S, Thompson L, Tielen P, Tomasch J, von Jan M, Wanphrut N, Wichels A, Zech H, Simon M.** 2010. The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a hitchhiker's guide to life in the sea. ISME J. **4:**61–77. http://dx.doi.org/10.1038/ismej.2009.94.
26. **Boettcher KJ, Geaghan KK, Maloy AP, Barber BJ.** 2005. *Roseovarius crassostreae* sp. nov., a member of the *Roseobacter* clade and the apparent cause of juvenile oyster disease (JOD) in cultured Eastern oysters. Int. J. Syst. Evol. Microbiol. **55:**1531–1537. http://dx.doi.org/10.1099/ijs.0.63620-0.
27. **Seyedsayamdost MR, Case RJ, Kolter R, Clardy J.** 2011. The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. Nat. Chem. **3:**331–335. http://dx.doi.org/10.1038/nchem.1002.
28. **Seyedsayamdost MR, Carr G, Kolter R, Clardy J.** 2011. Roseobacticides: small molecule modulators of an algal-bacterial symbiosis. J. Am. Chem. Soc. **133:**18343–18349. http://dx.doi.org/10.1021/ja207172s.
29. **Fernandes N, Case RJ, Longford SR, Seyedsayamdost MR, Steinberg PD, Kjelleberg S, Thomas T.** 2011. Genomes and virulence factors of novel bacterial pathogens causing bleaching disease in the marine red alga *Delisea pulchra*. PLoS One **6:**e27387. http://dx.doi.org/10.1371/journal.pone.0027387.
30. **Selje N, Simon M, Brinkhoff T.** 2004. A newly discovered Roseobacter cluster in temperate and polar oceans. Nature **427:**445–448. http://dx.doi.org/10.1038/nature02272.
31. **Brinkmeyer R, Knittel K, Jürgens J, Weyland H, Amann R, Helmke E.** 2003. Diversity and structure of bacterial communities in Arctic versus Antarctic pack ice. Appl. Environ. Microbiol. **69:**6610–6619. http://dx.doi.org/10.1128/AEM.69.11.6610-6619.2003.
32. **Junge K, Imhoff F, Staley T, Deming W.** 2002. Phylogenetic diversity of numerically important arctic sea-ice bacteria cultured at subzero temperature. Microb. Ecol. **43:**315–328. http://dx.doi.org/10.1007/s00248-001-1026-4.
33. **Lenk S, Moraru C, Hahnke S, Arnds J, Richter M, Kube M, Reinhardt R, Brinkhoff T, Harder J, Amann R, Musmann M.** 2012. Roseobacter clade bacteria are abundant in coastal sediments and encode a novel combination of sulfur oxidation genes. ISME J. **6:**2178–2187. http://dx.doi.org/10.1038/ismej.2012.66.
34. **Eloe EA, Malfatti F, Gutierrez J, Hardy K, Schmidt WE, Pogliano K, Pogliano J, Azam F, Bartlett DH.** 2011. Isolation and characterization of a psychropiezophilic alphaproteobacterium. Appl. Environ. Microbiol. **77:**8145–8153. http://dx.doi.org/10.1128/AEM.05204-11.
35. **Wang B, Tan T, Shao Z.** 2009. *Roseovarius pacificus* sp. nov., isolated from deep-sea sediment. Int. J. Syst. Evol. Microbiol. **59:**1116–1121. http://dx.doi.org/10.1099/ijs.0.002477-0.
36. **Moran MA, Buchan A, Gonzalez JM, Heidelberg JF, Whitman WB, Kiene RP, Henriksen JR, King GM, Belas R, Fuqua C, Brinkac L, Lewis M, Johri S, Weaver B, Pai G, Eisen JA, Rahe E, Sheldon WM, Ye W, Miller TR, Carlton J, Rasko DA, Paulsen IT, Ren Q, Daugherty SC, Deboy RT, Dodson RJ, Durkin AS, Madupu R, Nelson WC, Sullivan SA, Rosovitz MJ, Haft DH, Selengut J, Ward N.** 2004. Genome sequence of Silicibacter pomeroyi reveals adaptations to the marine environment. Nature **432:**910–913. http://dx.doi.org/10.1038/nature03170.
37. **Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE, Howard EC, King E, Oakley CA, Reisch CR, Rinta-Kanto JM, Sharma S, Sun S, Varaljay V, Vila-Costa M, Westrich JR, Moran MA.** 2010. Genome characteristics of a generalist marine bacterial lineage. ISME J. **4:**784–798. http://dx.doi.org/10.1038/ismej.2009.150.
38. **Kalhoefer D, Thole S, Voget S, Lehmann R, Liesegang H, Wollher A, Daniel R, Simon M, Brinkhoff T.** 2011. Comparative genome analysis and genome-guided physiological analysis of Roseobacter litoralis. BMC Genomics **12:**324. http://dx.doi.org/10.1186/1471-2164-12-324.
39. **Thole S, Kalhoefer D, Voget S, Berger M, Engelhardt T, Liesegang H, Wollherr A, Kjelleberg S, Daniel R, Simon M, Thomas T, Brinkhoff T.** 2012. *Phaeobacter gallaeciensis* genomes from globally opposite locations reveal high similarity of adaptation to surface life. ISME J. **6:**2229–2244. http://dx.doi.org/10.1038/ismej.2012.62.
40. **Buchan A, Gonzalez JM, Moran MA.** 2005. Overview of the marine Roseobacter lineage. Appl. Environ. Microbiol. **71:**5665–5677. http://dx.doi.org/10.1128/AEM.71.10.5665-5677.2005.
41. **Luo H, Csűros M, Hughes AL, Moran MA.** 2013. Evolution of divergent life history strategies in marine Alphaproteobacteria. mBio **4(4):**e00373-13. http://dx.doi.org/10.1128/mBio.00373-13.
42. **Luo H, Swan BK, Stepanauskas R, Hughes AL, Moran MA.** 2014. Evolutionary analysis of a streamlined lineage of surface ocean roseobacters. ISME J. **8:**1428–1439. http://dx.doi.org/10.1038/ismej.2013.248.
43. **Luo H, Löytynoja A, Moran MA.** 2012. Genome content of uncultivated marine roseobacters in the surface ocean. Environ. Microbiol. **14:**41–51. http://dx.doi.org/10.1111/j.1462-2920.2011.02528.x.
44. **Foster PG.** 2004. Modeling compositional heterogeneity. Syst. Biol. **53:**485–495. http://dx.doi.org/10.1080/10635150490445779.
45. **Lartillot N, Philippe H.** 2004. A Bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. Mol. Biol. Evol. **21:**1095–1109. http://dx.doi.org/10.1093/molbev/msh112.
46. **Galtier N.** 2001. Maximum-likelihood phylogenetic analysis under a covarion-like model. Mol. Biol. Evol. **18:**866–873. http://dx.doi.org/10.1093/oxfordjournals.molbev.a003868.
47. **Lanfear R, Calcott B, Ho SYW, Guindon S.** 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. Mol. Biol. Evol. **29:**1695–1701. http://dx.doi.org/10.1093/molbev/mss020.
48. **Brinkhoff T, Giebel H-A, Simon M.** 2008. Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. Arch. Microbiol. **189:**531–539. http://dx.doi.org/10.1007/s00203-008-0353-y.
49. **Slightom RN, Buchan A.** 2009. Surface colonization by marine roseobacters: integrating genotype and phenotype. Appl. Environ. Microbiol. **75:**6027–6037. http://dx.doi.org/10.1128/AEM.01508-09.
50. **Lang AS, Beatty JT.** 2000. Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. Proc. Natl. Acad. Sci. U. S. A. **97:**859–864. http://dx.doi.org/10.1073/pnas.97.2.859.
51. **Huang S, Zhang Y, Chen F, Jiao N.** 2011. Complete genome sequence of a marine roseophage provides evidence into the evolution of gene transfer agents in alphaproteobacteria. Virol. J. **8:**124. http://dx.doi.org/10.1186/1743-422X-8-124.
52. **Lang AS, Zhaxybayeva O, Beatty JT.** 2012. Gene transfer agents: phage-like elements of genetic exchange. Nat. Rev. Microbiol. **10:**472–482. http://dx.doi.org/10.1038/nrmicro2802.
53. **McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, Paul JH.** 2010. High frequency of horizontal gene transfer in the oceans. Science **330:**50. http://dx.doi.org/10.1126/science.1192243.
54. **Biers EJ, Wang K, Pennington C, Belas R, Chen F, Moran MA.** 2008. Occurrence and expression of gene transfer agent genes in marine bacterioplankton. Appl. Environ. Microbiol. **74:**2933–2939. http://dx.doi.org/10.1128/AEM.02129-07.
55. **Tang K-H, Feng X, Tang YJ, Blankenship RE.** 2009. Carbohydrate metabolism and carbon fixation in *Roseobacter denitrificans* OCh114. PLoS One **4:**e7233. http://dx.doi.org/10.1371/journal.pone.0007233.
56. **Cunliffe M.** 2011. Correlating carbon monoxide oxidation with cox

genes in the abundant marine Roseobacter clade. ISME J. **5:**685–691. http://dx.doi.org/10.1038/ismej.2010.170.
57. **González JM, Covert JS, Whitman WB, Henriksen JR, Mayer F, Scharf B, Schmitt R, Buchan A, Fuhrman JA, Kiene RP, Moran MA.** 2003. *Silicibacter pomeroyi* sp. nov. and *Roseovarius nubinhibens* sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. Int. J. Syst. Evol. Microbiol. **53:**1261–1269. http://dx.doi.org/10.1099/ijs.0.02491-0.
58. **Petersen J, Brinkmann H, Bunk B, Michael V, Päuker O, Pradella S.** 2012. Think pink: photosynthesis, plasmids and the Roseobacter clade. Environ. Microbiol. **14:**2661–2672. http://dx.doi.org/10.1111/j.1462-2920.2012.02806.x.
59. **Brinkhoff T, Bach G, Heidorn T, Liang L, Schlingloff A, Simon M.** 2004. Antibiotic production by a Roseobacter clade-affiliated species from the German Wadden Sea and its antagonistic effects on indigenous isolates. Appl. Environ. Microbiol. **70:**2560–2565. http://dx.doi.org/10.1128/AEM.70.4.2560-2565.2003.
60. **Berger M, Brock NL, Liesegang H, Dogs M, Preuth I, Simon M, Dickschat JS, Brinkhoff T.** 2012. Genetic analysis of the upper phenylacetate catabolic pathway in the production of tropodithietic acid by *Phaeobacter gallaeciensis*. Appl. Environ. Microbiol. **78:**3539–3551. http://dx.doi.org/10.1128/AEM.07657-11.
61. **Petersen J, Frank O, Göker M, Pradella S.** 2013. Extrachromosomal, extraordinary and essential—the plasmids of the Roseobacter clade. Appl. Microbiol. Biotechnol. **97:**2805–2815. http://dx.doi.org/10.1007/s00253-013-4746-8.
62. **Csűös M.** 2010. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. Bioinformatics **26:**1910–1912. http://dx.doi.org/10.1093/bioinformatics/btq315.
63. **Koblížek M, Zeng Y, Horák A, Oborník M.** 2013. Regressive evolution of photosynthesis in the Roseobacter clade, p 385–405. *In* Beatty JT (ed), Advances in botanical research, vol 66. Academic Press, San Diego, CA.
64. **Hold GL, Smith EA, Rappe MS, Maas EW, Moore ERB, Stroempl C, Stephen JR, Prosser JI, Birkbeck TH, Gallacher S.** 2001. Characterisation of bacterial communities associated with toxic and non-toxic dinofagellates: *Alexandrium* spp. and *Scrippsiella trochoidea*. FEMS Microbiol. Ecol. **37:**161–173. http://dx.doi.org/10.1111/j.1574-6941.2001.tb00864.x.
65. **Hasegawa Y, Martin JL, Giewat MW, Rooney-Varga JN.** 2007. Microbial community diversity in the phycosphere of natural populations of the toxic alga, *Alexandrium fundyense*. Environ. Microbiol. **9:**3108–3121. http://dx.doi.org/10.1111/j.1462-2920.2007.01421.x.
66. **Jones KL, Mikulski CM, Barnhorst A, Doucette GJ.** 2010. Comparative analysis of bacterioplankton assemblages from *Karenia brevis* bloom and nonbloom water on the west Florida shelf (Gulf of Mexico, USA) using 16S rRNA gene clone libraries. FEMS Microbiol. Ecol. **73:**468–485. http://dx.doi.org/10.1111/j.1574-6941.2010.00914.x.
67. **Zubkov MV, Fuchs BM, Archer SD, Kiene RP, Amann R, Burkill PH.** 2002. Rapid turnover of dissolved DMS and DMSP by defined bacterioplankton communities in the stratified euphotic zone of the North Sea. Deep Sea Res. Part 2 Top. Stud. Oceanogr. **49:**3017–3038. http://dx.doi.org/10.1016/S0967-0645(02)00069-3.
68. **Jackson GA.** 1987. Simulating chemosensory responses of marine microorganisms. Limnol. Oceanogr. **32:**1253–1266. http://dx.doi.org/10.4319/lo.1987.32.6.1253.
69. **Hughes AL.** 2012. Evolution of adaptive phenotypic traits without positive Darwinian selection. Heredity **108:**347–353. http://dx.doi.org/10.1038/hdy.2011.97.
70. **Gould SJ, Vrba ES.** 1982. Exaptation—a missing term in the science of form. Paleobiology **8:**4–15.
71. **Rappe MS, Giovannoni SJ.** 2003. The uncultured microbial majority. Annu. Rev. Microbiol. **57:**369–394. http://dx.doi.org/10.1146/annurev.micro.57.030502.090759.
72. **Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin K, Batista D, Bibbs L, Eads J, Richardson TH, Noordewier M.** 2005. Genome streamlining in a cosmopolitan oceanic bacterium. Science **309:**1242–1245. http://dx.doi.org/10.1126/science.1114057.
73. **Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ, Rappé MS.** 2012. Streamlining and core genome conservation among highly divergent members of the SAR11 clade. mBio **3**(5)**:**e00252-12. http://dx.doi.org/10.1128/mBio.00252-12.
74. **Polz MF, Hunt DE, Preheim SP, Weinreich DM.** 2006. Patterns and mechanisms of genetic and phenotypic differentiation in marine microbes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **361:**2009–2021. http://dx.doi.org/10.1098/rstb.2006.1928.
75. **Miller TR, Belas R.** 2006. Motility is involved in Silicibacter sp. TM1040 interaction with dinoflagellates. Environ. Microbiol. **8:**1648–1659. http://dx.doi.org/10.1111/j.1462-2920.2006.01071.x.
76. **Persson OP, Pinhassi J, Riemann L, Marklund B-I, Rhen M, Normark S, González JM, Hagström Å.** 2009. High abundance of virulence gene homologues in marine bacteria. Environ. Microbiol. **11:**1348–1357. http://dx.doi.org/10.1111/j.1462-2920.2008.01861.x.
77. **Cicirelli EM, Williamson H, Tait K, Fuqua C.** 2008. Acylated homoserine lactone signaling in marine bacterial systems, p 251–272. *In* Winans SC, Bassler BL (ed), Chemical communication among bacteria. ASM Press, Washington, DC.
78. **Hjelm M, Bergh Ø, Riaza A, Nielsen J, Melchiorsen J, Jensen S, Duncan H, Ahrens P, Birkbeck H, Gram L.** 2004. Selection and identification of autochthonous potential probiotic bacteria from turbot larvae (*Scophthalmus maximus*) rearing units. Syst. Appl. Microbiol. **27:**360–371. http://dx.doi.org/10.1078/0723-2020-00256.
79. **Fogg GE.** 1983. The ecological significance of extracellular products of phytoplankton photosynthesis. Bot. Mar. **26:**3–14.
80. **Edenborn HM, Litchfield CD.** 1987. Glycolate turnover in the water column of the New York Bight apex. Mar. Biol. **95:**459–467. http://dx.doi.org/10.1007/BF00409575.
81. **Leboulanger C, Oriol L, Jupin H, Desolas-gros C.** 1997. Diel variability of glycolate in the eastern tropical Atlantic Ocean. Deep Sea Res. Part 1 Oceanogr. Res. Pap. **44:**2131–2139. http://dx.doi.org/10.1016/S0967-0637(97)00090-3.
82. **Coughlan SJ, Al Hasan RH.** 1977. Studies of uptake and turnover of glycolic acid in the Menai Straits North Wales. J. Ecol. **65:**731–746. http://dx.doi.org/10.2307/2259376.
83. **Lau WWY, Keil RG, Armbrust EV.** 2007. Succession and diel transcriptional response of the glycolate-utilizing component of the bacterial community during a spring phytoplankton bloom. Appl. Environ. Microbiol. **73:**2440–2450. http://dx.doi.org/10.1128/AEM.01965-06.
84. **Lau WWY, Armbrust EV.** 2006. Detection of glycolate oxidase gene glcD diversity among cultured and environmental marine bacteria. Environ. Microbiol. **8:**1688–1702. http://dx.doi.org/10.1111/j.1462-2920.2006.01092.x.
85. **Hahnke S, Sperling M, Langer T, Wichels A, Gerdts G, Beardsley C, Brinkhoff T, Simon M.** 2013. Distinct seasonal growth patterns of the bacterium *Planktotalea frisia* in the North Sea and specific interaction with phytoplankton algae. FEMS Microbiol. Ecol. **86:**185–199. http://dx.doi.org/10.1111/1574-6941.12151.
86. **Seymour JR, Simó R, Ahmed T, Stocker R.** 2010. Chemoattraction to dimethylsulfoniopropionate throughout the marine microbial food web. Science **329:**342–345. http://dx.doi.org/10.1126/science.1188418.
87. **Varaljay VA, Gifford SM, Wilson ST, Sharma S, Karl DM, Moran MA.** 2012. Bacterial dimethylsulfoniopropionate degradation genes in the oligotrophic North Pacific Subtropical Gyre. Appl. Environ. Microbiol. **78:**2775–2782. http://dx.doi.org/10.1128/AEM.07559-11.
88. **Ribalet F, Intertaglia L, Lebaron P, Casotti R.** 2008. Differential effect of three polyunsaturated aldehydes on marine bacterial isolates. Aquat. Toxicol. **86:**249–255. http://dx.doi.org/10.1016/j.aquatox.2007.11.005.
89. **Paul C, Mausz M, Pohnert G.** 2013. A co-culturing/metabolomics approach to investigate chemically mediated interactions of planktonic organisms reveals influence of bacteria on diatom metabolism. Metabolomics **9:**349–359. http://dx.doi.org/10.1007/s11306-012-0453-1.
90. **Sañudo-Wilhelmy SA, Gómez-Consarnau L, Suffridge C, Webb EA.** 2014. The role of B vitamins in marine biogeochemistry. Annu. Rev. Mar. Sci. **6:**339–367. http://dx.doi.org/10.1146/annurev-marine-120710-100912.
91. **Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG.** 2005. Algae acquire vitamin B12 through a symbiotic relationship with bacteria. Nature **438:**90–93. http://dx.doi.org/10.1038/nature04056.
92. **Droop MR.** 2007. Vitamins, phytoplankton and bacteria: symbiosis or scavenging? J. Plankton Res. **29:**107–113. http://dx.doi.org/10.1093/plankt/fbm009.
93. **Mayali X, Franks PJS, Azam F.** 2008. Cultivation and ecosystem role of a marine Roseobacter clade-affiliated cluster bacterium. Appl. Environ. Microbiol. **74:**2595–2603. http://dx.doi.org/10.1128/AEM.02191-07.
94. **Christie PJ, Whitaker N, González-Rivera C.** 2014. Mechanism and structure of the bacterial type IV secretion systems. Biochim. Biophys. Acta **1843:**1578–1591. http://dx.doi.org/10.1016/j.bbamcr.2013.12.019.

95. **Webster NS, Negri AP, Munro MMHG, Battershill CN.** 2004. Diverse microbial communities inhabit Antarctic sponges. Environ. Microbiol. **6:**288–300. http://dx.doi.org/10.1111/j.1462-2920.2004.00570.x.
96. **Porsby CH, Nielsen KF, Gram L.** 2008. *Phaeobacter* and *Ruegeria* species of the *Roseobacter* clade colonize separate niches in a Danish turbot (*Scophthalmus maximus*)-rearing farm and antagonize *Vibrio anguillarum* under different growth conditions. Appl. Environ. Microbiol. **74:**7356–7364. http://dx.doi.org/10.1128/AEM.01738-08.
97. **Boardman CL, Maloy AP, Boettcher KJ.** 2008. Localization of the bacterial agent of juvenile oyster disease (*Roseovarius crassostreae*) within affected Eastern oysters (*Crassostrea virginica*). J. Invertebr. Pathol. **97:**150–158. http://dx.doi.org/10.1016/j.jip.2007.08.007.
98. **Ashen JB, Goff LJ.** 1996. Molecular identification of a bacterium associated with gall formation in the marine red alga *Prionitis lanceolata*. J. Phycol. **32:**286–297. http://dx.doi.org/10.1111/j.0022-3646.1996.00286.x.
99. **Boettcher KJ, Barber BJ, Singer JT.** 2000. Additional evidence that juvenile oyster disease is caused by a member of the *Roseobacter* group and colonization of nonaffected animals by *Stappia stellulata*-like strains. Appl. Environ. Microbiol. **66:**3924–3930. http://dx.doi.org/10.1128/AEM.66.9.3924-3930.2000.
100. **Case RJ, Longford SR, Campbell AH, Low A, Tujula N, Steinberg PD, Kjelleberg S.** 2011. Temperature induced bacterial virulence and bleaching disease in a chemically defended marine macroalga. Environ. Microbiol. **13:**529–537. http://dx.doi.org/10.1111/j.1462-2920.2010.02356.x.
101. **Cooney RP, Pantos O, Le Tissier MDA, Barer MR, O'Donnell AG, Bythell JC.** 2002. Characterization of the bacterial consortium associated with black band disease in coral using molecular microbiological techniques. Environ. Microbiol. **4:**401–413. http://dx.doi.org/10.1046/j.1462-2920.2002.00308.x.
102. **Koren O, Rosenberg E.** 2006. Bacteria associated with mucus and tissues of the coral *Oculina patagonica* in summer and winter. Appl. Environ. Microbiol. **72:**5254–5259. http://dx.doi.org/10.1128/AEM.00554-06.
103. **Lampert Y, Kelman D, Nitzan Y, Dubinsky Z, Behar A, Hill RT.** 2008. Phylogenetic diversity of bacteria associated with the mucus of Red Sea corals. FEMS Microbiol. Ecol. **64:**187–198. http://dx.doi.org/10.1111/j.1574-6941.2008.00458.x.
104. **Littman RA, Willis BL, Pfeffer C, Bourne DG.** 2009. Diversities of coral-associated bacteria differ with location, but not species, for three acroporid corals on the Great Barrier Reef. FEMS Microbiol. Ecol. **68:**152–163. http://dx.doi.org/10.1111/j.1574-6941.2009.00666.x.
105. **Rohwer F, Breitbart M, Jara J, Azam F, Knowlton N.** 2001. Diversity of bacteria associated with the Caribbean coral *Montastraea franksi*. Coral Reefs **20:**85–91. http://dx.doi.org/10.1007/s003380100138.
106. **Sunagawa S, DeSantis TZ, Piceno YM, Brodie EL, DeSalvo MK, Voolstra CR, Weil E, Andersen GL, Medina M.** 2009. Bacterial diversity and white plague disease-associated community changes in the Caribbean coral *Montastraea faveolata*. ISME J. **3:**512–521. http://dx.doi.org/10.1038/ismej.2008.131.
107. **Raina J-B, Tapiolas D, Willis BL, Bourne DG.** 2009. Coral-associated bacteria and their role in the biogeochemical cycling of sulfur. Appl. Environ. Microbiol. **75:**3492–3501. http://dx.doi.org/10.1128/AEM.02567-08.
108. **McKew BA, Dumbrell AJ, Daud SD, Hepburn L, Thorpe E, Mogensen L, Whitby C.** 2012. Characterization of geographically distinct bacterial communities associated with coral mucus produced by *Acropora* spp. and *Porites* spp. Appl. Environ. Microbiol. **78:**5229–5237. http://dx.doi.org/10.1128/AEM.07764-11.
109. **Lee OO, Yang J, Bougouffa S, Wang Y, Batang Z, Tian R, Al-Suwailem A, Qian P-Y.** 2012. Spatial and species variations in bacterial communities associated with corals from the Red Sea as revealed by pyrosequencing. Appl. Environ. Microbiol. **78:**7173–7184. http://dx.doi.org/10.1128/AEM.01111-12.
110. **Morrow KM, Moss AG, Chadwick NE, Liles MR.** 2012. Bacterial associates of two Caribbean coral species reveal species-specific distribution and geographic variability. Appl. Environ. Microbiol. **78:**6438–6449. http://dx.doi.org/10.1128/AEM.01162-12.
111. **Raina J-B, Dinsdale EA, Willis BL, Bourne DG.** 2010. Do the organic sulfur compounds DMSP and DMS drive coral microbial associations? Trends Microbiol. **18:**101–108. http://dx.doi.org/10.1016/j.tim.2009.12.002.
112. **Apprill A, Marlow HQ, Martindale MQ, Rappé MS.** 2012. Specificity of associations between bacteria and the coral *Pocillopora meandrina* during early development. Appl. Environ. Microbiol. **78:**7467–7475. http://dx.doi.org/10.1128/AEM.01232-12.
113. **Sharp KH, Distel D, Paul VJ.** 2012. Diversity and dynamics of bacterial communities in early life stages of the Caribbean coral *Porites astreoides*. ISME J. **6:**790–801. http://dx.doi.org/10.1038/ismej.2011.144.
114. **Ceh J, Raina J-B, Soo RM, van Keulen M, Bourne DG.** 2012. Coral-bacterial communities before and after a coral mass spawning event on Ningaloo Reef. PLoS One **7:**e36920. http://dx.doi.org/10.1371/journal.pone.0036920.
115. **Nissimov J, Rosenberg E, Munn CB.** 2009. Antimicrobial properties of resident coral mucus bacteria of *Oculina patagonica*. FEMS Microbiol. Lett. **292:**210–215. http://dx.doi.org/10.1111/j.1574-6968.2009.01490.x.
116. **Rypien KL, Ward JR, Azam F.** 2010. Antagonistic interactions among coral-associated bacteria. Environ. Microbiol. **12:**28–39. http://dx.doi.org/10.1111/j.1462-2920.2009.02027.x.
117. **Sekar R, Kaczmarsky LT, Richardson LL.** 2008. Microbial community composition of black band disease on the coral host *Siderastrea siderea* from three regions of the wider Caribbean. Mar. Ecol. Prog. Ser. **362:**85–98. http://dx.doi.org/10.3354/meps07496.
118. **Apprill A, Hughen K, Mincer T.** 2013. Major similarities in the bacterial communities associated with lesioned and healthy Fungiidae corals. Environ. Microbiol. **15:**2063–2072. http://dx.doi.org/10.1111/1462-2920.12107.
119. **Koski LB, Golding GB.** 2001. The closest BLAST hit is often not the nearest neighbor. J. Mol. Evol. **52:**540–542. http://dx.doi.org/10.1007/s002390010184.
120. **Iverson V, Morris RM, Frazar CD, Berthiaume CT, Morales RL, Armbrust EV.** 2012. Untangling genomes from metagenomes: revealing an uncultured class of marine Euryarchaeota. Science **335:**587–590. http://dx.doi.org/10.1126/science.1212665.
121. **Yooseph S, Nealson KH, Rusch DB, McCrow JP, Dupont CL, Kim M, Johnson J, Montgomery R, Ferriera S, Beeson K, Williamson SJ, Tovchigrechko A, Allen AE, Zeigler LA, Sutton G, Eisenstadt E, Rogers Y-H, Friedman R, Frazier M, Venter JC.** 2010. Genomic and functional adaptation in surface ocean planktonic prokaryotes. Nature **468:**60–66. http://dx.doi.org/10.1038/nature09530.
122. **Lauro FM, McDougald D, Thomas T, Williams TJ, Egan S, Rice S, DeMaere MZ, Ting L, Ertan H, Johnson J, Ferriera S, Lapidus A, Anderson I, Kyrpides N, Munk AC, Detter C, Han CS, Brown MV, Robb FT, Kjelleberg S, Cavicchioli R.** 2009. The genomic basis of trophic strategy in marine bacteria. Proc. Natl. Acad. Sci. U. S. A. **106:**15527–15533. http://dx.doi.org/10.1073/pnas.0903507106.
123. **Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, González JM, Luo H, Wright JJ, Landry ZC, Hanson NW, Thompson BP, Poulton NJ, Schwientek P, Acinas SG, Giovannoni SJ, Moran MA, Hallam SJ, Cavicchioli R, Woyke T, Stepanauskas R.** 2013. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. Proc. Natl. Acad. Sci. U. S. A. **110:**11463–11468. http://dx.doi.org/10.1073/pnas.1304246110.
124. **Hughes AL, Ota T, Nei M.** 1990. Positive Darwinian selection promotes charge profile diversity in the antigen-binding cleft of class I major-histocompatibility-complex molecules. Mol. Biol. Evol. **7:**515–524.
125. **Zhang J.** 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. J. Mol. Evol. **50:**56–68.
126. **Bragg JG, Hyder CL.** 2004. Nitrogen versus carbon use in prokaryotic genomes and proteomes. Proc. Biol. Sci. **271:**S374–S377. http://dx.doi.org/10.1098/rsbl.2004.0193.
127. **Grzymski JJ, Dussaq AM.** 2012. The significance of nitrogen cost minimization in proteomes of marine microorganisms. ISME J. **6:**71–80. http://dx.doi.org/10.1038/ismej.2011.72.
128. **Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, Cristina Souza de Oliveira T, Wagner Garcia J, Pellon de Miranda F, Henrique-Silva F.** 2011. Metagenomics of the water column in the pristine upper course of the Amazon river. PLoS One **6:**e23785. http://dx.doi.org/10.1371/journal.pone.0023785.
129. **Hershberg R, Petrov DA.** 2010. Evidence that mutation is universally biased towards AT in bacteria. PLoS Genet. **6:**e1001115. http://dx.doi.org/10.1371/journal.pgen.1001115.
130. **Hildebrand F, Meyer A, Eyre-Walker A.** 2010. Evidence of selection upon genomic GC-content in bacteria. PLoS Genet. **6:**e1001107. http://dx.doi.org/10.1371/journal.pgen.1001107.
131. **Balbi KJ, Rocha EPC, Feil EJ.** 2009. The temporal dynamics of slightly

deleterious mutations in Escherichia coli and Shigella spp. Mol. Biol. Evol. **26:**345–355. http://dx.doi.org/10.1093/molbev/msn252.
132. **Lind PA, Andersson DI.** 2008. Whole-genome mutational biases in bacteria. Proc. Natl. Acad. Sci. U. S. A. **105:**17878–17883. http://dx.doi.org/10.1073/pnas.0804445105.
133. **Morris JJ, Lenski RE, Zinser ER.** 2012. The Black Queen hypothesis: evolution of dependencies through adaptive gene loss. mBio **3**(2)**:** e00036-12. http://dx.doi.org/10.1128/mBio.00036-12.
134. **Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu W-T, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T.** 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature **499:**431–437. http://dx.doi.org/10.1038/nature12352.
135. **Luo H, Tolar BB, Swan BK, Zhang CL, Stepanauskas R, Ann Moran M, Hollibaugh JT.** 2014. Single-cell genomics shedding light on marine Thaumarchaeota diversification. ISME J. **8:**732–736. http://dx.doi.org/10.1038/ismej.2013.202.
136. **Galtier N, Gouy M.** 1995. Inferring phylogenies from DNA sequences of unequal base compositions. Proc. Natl. Acad. Sci. U. S. A. **92:**11317–11321. http://dx.doi.org/10.1073/pnas.92.24.11317.
137. **Jermiin LS, Ho SYW, Ababneh F, Robinson J, Larkum AWD.** 2004. The biasing effect of compositional heterogeneity on phylogenetic estimates may be underestimated. Syst. Biol. **53:**638–643. http://dx.doi.org/10.1080/10635150490468648.
138. **Nesnidal MP, Helmkampf M, Bruchhaus I, Hausdorf B.** 2010. Compositional heterogeneity and phylogenomic inference of metazoan relationships. Mol. Biol. Evol. **27:**2095–2104. http://dx.doi.org/10.1093/molbev/msq097.
139. **Foster PG, Cox CJ, Embley TM.** 2009. The primary divisions of life: a phylogenomic approach employing composition-heterogeneous methods. Philos. Trans. R. Soc. Lond. B Biol. Sci. **364:**2197–2207. http://dx.doi.org/10.1098/rstb.2009.0034.
140. **Herbeck JT, Degnan PH, Wernegreen JJ.** 2005. Nonhomogeneous model of sequence evolution indicates independent origins of primary endosymbionts within the Enterobacteriales (γ-Proteobacteria). Mol. Biol. Evol. **22:**520–532. http://dx.doi.org/10.1093/molbev/msi036.
141. **Dagan T, Blekhman R, Graur D.** 2006. The "domino theory" of gene death: gradual and mass gene extinction events in three lineages of obligate symbiotic bacterial pathogens. Mol. Biol. Evol. **23:**310–316. http://dx.doi.org/10.1093/molbev/msj036.
142. **Moran NA, Wernegreen JJ.** 2000. Lifestyle evolution in symbiotic bacteria: insights from genomics. Trends Ecol. Evol. **15:**321–326. http://dx.doi.org/10.1016/S0169-5347(00)01902-9.
143. **Woolfit M, Bromham L.** 2003. Increased rates of sequence evolution in endosymbiotic bacteria and fungi with small effective population sizes. Mol. Biol. Evol. **20:**1545–1555. http://dx.doi.org/10.1093/molbev/msg167.
144. **Dupont CL, Rusch DB, Yooseph S, Lombardo M-J, Richter RA, Valas R, Novotny M, Yee-Greenbaum J, Selengut JD, Haft DH, Halpern AL, Lasken RS, Nealson K, Friedman R, Venter JC.** 2012. Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. ISME J. **6:**1186–1199. http://dx.doi.org/10.1038/ismej.2011.189.
145. **Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, Ahlgren NA, Arellano A, Coleman M, Hauser L, Hess WR, Johnson ZI, Land M, Lindell D, Post AF, Regala W, Shah M, Shaw SL, Steglich C, Sullivan MB, Ting CS, Tolonen A, Webb EA, Zinser ER, Chisholm SW.** 2003. Genome divergence in two Prochlorococcus ecotypes reflects oceanic niche differentiation. Nature **424:**1042–1047. http://dx.doi.org/10.1038/nature01947.
146. **Dufresne A, Salanoubat M, Partensky F, Artiguenave F, Axmann IM, Barbe V, Duprat S, Galperin MY, Koonin EV, Le Gall F, Makarova KS, Ostrowski M, Oztas S, Robert C, Rogozin IB, Scanlan DJ, de Marsac NT, Weissenbach J, Wincker P, Wolf YI, Hess WR.** 2003. Genome sequence of the cyanobacterium Prochlorococcus marinus SS120, a nearly minimal oxyphototrophic genome. Proc. Natl. Acad. Sci. U. S. A. **100:**10020–10025. http://dx.doi.org/10.1073/pnas.1733211100.
147. **Dufresne A, Garczarek L, Partensky F.** 2005. Accelerated evolution associated with genome reduction in a free-living prokaryote. Genome Biol. **6:**R14. http://dx.doi.org/10.1186/gb-2005-6-2-r14.
148. **Biers EJ, Sun S, Howard EC.** 2009. Prokaryotic genomes and diversity in surface ocean waters: interrogating the global ocean sampling metagenome. Appl. Environ. Microbiol. **75:**2221–2229. http://dx.doi.org/10.1128/AEM.02118-08.
149. **Wilkins D, Lauro FM, Williams TJ, Demaere MZ, Brown MV, Hoffman JM, Andrews-Pfannkoch C, McQuaid JB, Riddle MJ, Rintoul SR, Cavicchioli R.** 2013. Biogeographic partitioning of Southern Ocean microorganisms revealed by metagenomics. Environ. Microbiol. **15:**1318–1333. http://dx.doi.org/10.1111/1462-2920.12035.
150. **Charlesworth B.** 2009. Effective population size and patterns of molecular evolution and variation. Nat. Rev. Genet. **10:**195–205. http://dx.doi.org/10.1038/nrg2526.
151. **Luo H, Swan BK, Stepanauskas R, Hughes AL, Moran MA.** 2014. Comparing effective population sizes of dominant marine alphaproteobacteria lineages. Environ. Microbiol. Rep. **6:**167–172. http://dx.doi.org/10.1111/1758-2229.12129.
152. **Woolfit M.** 2009. Effective population size and the rate and pattern of nucleotide substitutions. Biol. Lett. **5:**417–420. http://dx.doi.org/10.1098/rsbl.2009.0155.
153. **Ohta T.** 1992. The nearly neutral theory of molecular evolution. Annu. Rev. Ecol. Evol. Syst. **23:**263–286. http://dx.doi.org/10.1146/annurev.es.23.110192.001403.
154. **Fraser C, Alm EJ, Polz MF, Spratt BG, Hanage WP.** 2009. The bacterial species challenge: making sense of genetic and ecological diversity. Science **323:**741–746. http://dx.doi.org/10.1126/science.1159388.
155. **Sung W, Ackerman MS, Miller SF, Doak TG, Lynch M.** 2012. Drift-barrier hypothesis and mutation-rate evolution. Proc. Natl. Acad. Sci. U. S. A. **109:**18488–18492. http://dx.doi.org/10.1073/pnas.1216223109.
156. **Luo H, Hughes AL.** 2012. dN/dS does not show positive selection drives separation of polar-tropical SAR11 populations. Mol. Syst. Biol. **8:**625. http://dx.doi.org/10.1038/msb.2012.58.
157. **Luo H, Friedman R, Tang J, Hughes AL.** 2011. Genome reduction by deletion of paralogs in the marine cyanobacterium Prochlorococcus. Mol. Biol. Evol. **28:**2751–2760. http://dx.doi.org/10.1093/molbev/msr081.
158. **Kettler GC, Martiny AC, Huang K, Zucker J, Coleman ML, Rodrigue S, Chen F, Lapidus A, Ferriera S, Johnson J, Steglich C, Church GM, Richardson P, Chisholm SW.** 2007. Patterns and implications of gene gain and loss in the evolution of Prochlorococcus. PLoS Genet. **3:**e231. http://dx.doi.org/10.1371/journal.pgen.0030231.
159. **Osburne MS, Holmbeck BM, Coe A, Chisholm SW.** 2011. The spontaneous mutation frequencies of *Prochlorococcus* strains are commensurate with those of other bacteria. Environ. Microbiol. Rep. **3:**744–749. http://dx.doi.org/10.1111/j.1758-2229.2011.00293.x.
160. **Acinas SG, Klepac-Ceraj V, Hunt DE, Pharino C, Ceraj I, Distel DL, Polz MF.** 2004. Fine-scale phylogenetic architecture of a complex bacterial community. Nature **430:**551–554. http://dx.doi.org/10.1038/nature02649.
161. **Zhao Y, Wang K, Budinoff C, Buchan A, Lang A, Jiao N, Chen F.** 2009. Gene transfer agent (GTA) genes reveal diverse and dynamic *Roseobacter* and *Rhodobacter* populations in the Chesapeake Bay. ISME J. **3:**364–373. http://dx.doi.org/10.1038/ismej.2008.115.

Mary Ann Moran is a Distinguished Research Professor in the Department of Marine Sciences at the University of Georgia. She has earned degrees at Colgate University (B.A., 1977), Cornell University (M.S., 1982), and the University of Georgia (Ph.D., 1987). Dr. Moran is a fellow of the American Association for the Advancement of Science and the American Academy of Microbiology and a recipient of the American Society for Microbiology's D. C. White Research and Mentoring Award. She serves on the editorial boards of *Applied and Environmental Microbiology*, *Environmental Microbiology*, and *mBio*. Dr. Moran's research program focuses on the genetic basis of bacterial sulfur and carbon cycling in the ocean, with the goal of understanding the role of marine bacteria in the productivity of the ocean and the formation and flux of climatically active gases. Her research uses molecular microbial ecology and ecological genomics approaches to explore bacterial processes and their regulation in seawater.

Haiwei Luo received his M.S. in marine microbial ecology and Ph.D. in molecular evolution and bioinformatics from the University of South Carolina. During his doctoral study, he employed genome rearrangement and molecular sequence-based approaches to resolve evolutionary relationships of microbial and animal species and used population genetic models to address how genetic diversity evolved. He is currently a postdoctoral researcher in the Department of Marine Sciences at the University of Georgia. His current research interest includes understanding the role of selection, drift, mutation, and recombination in giving rise to the genetic diversity of marine microbes and resolving the controversial evolutionary relationships of marine microbial lineages. He will begin a position in the School of Life Sciences at the Chinese University of Hong Kong in January 2015, where he will continue to employ phylogenomic and population genetic approaches to understand microbial evolution and adaptation in the ocean.