Prochlorococcus: Advantages and Limits of Minimalism

Frédéric Partensky[1,2] and Laurence Garczarek[1,2]

[1]UPMC-Université Paris 06, Station Biologique, 29682 Roscoff cedex, France

[2]CNRS, UMR 7144, Oceanic Plankton Group, 29682 Roscoff cedex, France; email: partensky@sb-roscoff.fr

Annu. Rev. Mar. Sci. 2010. 2:305–31

First published online as a Review in Advance on September 28, 2009

The *Annual Review of Marine Science* is online at marine.annualreviews.org

This article's doi: 10.1146/annurev-marine-120308-081034

1941-1405/10/0115-0305$20.00

Key Words

DNA repair, evolution, genome streamlining, marine cyanobacteria, niche adaptation, oligotrophy, photosynthesis

Abstract

Prochlorococcus is the key phytoplanktonic organism of tropical gyres, large ocean regions that are depleted of the essential macronutrients needed for photosynthesis and cell growth. This cyanobacterium has adapted itself to oligotrophy by minimizing the resources necessary for life through a drastic reduction of cell and genome sizes. This rarely observed strategy in free-living organisms has conferred on *Prochlorococcus* a considerable advantage over other phototrophs, including its closest relative *Synechococcus*, for life in this vast yet little variable ecosystem. However, this strategy seems to reach its limits in the upper layer of the S Pacific gyre, the most oligotrophic region of the world ocean. By losing some important genes and/or functions during evolution, *Prochlorococcus* has seemingly become dependent on co-occurring microorganisms. In this review, we present some of the recent advances in the ecology, biology, and evolution of *Prochlorococcus*, which because of its ecological importance and tiny genome is rapidly imposing itself as a model organism in environmental microbiology.

INTRODUCTION

Cyanobacteria: prokaryotes capable of oxygenic photosynthesis, a process by which they use solar energy to fix carbon dioxide, with the release of oxygen

Cyanophages: viruses that specifically infect cyanobacterial cells

Cyanobacteria are the most ancient oxygenic phototrophs on Earth and have a key role in a variety of ecosystems. Indeed, free-living and symbiotic cyanobacteria are collectively ubiquitous, colonizing all environments reached by solar light, including the most inhospitable areas of the planet, such as hot springs, sand deserts, and nutrient-depleted areas of the world ocean (Garcia-Pichel et al. 2003, Garcia-Pichel et al. 2001). Despite the huge ecosystem they inhabit, the diversity of marine planktonic, free-living cyanobacteria is surprisingly simple, compared to, e.g., their benthic counterparts. Only a few genera are known so far, including the diazotrophs (i.e., cells able to fix atmospheric N_2) *Trichodesmium* and *Crocosphaera* and the two nondiazotrophs *Synechococcus* and *Prochlorococcus*. To the first category can be added one recently discovered and still uncultured organism (provisionally called UCYN-A) phylogenetically related to cyanobacteria but which has the astounding peculiarity of possessing neither photosystem II nor RuBisCo and thus has likely lost the ability to perform oxygenic photosynthesis (Zehr et al. 2008). Although all of these phototrophs are regular inhabitants of subtropical, oligotrophic areas, only *Prochlorococcus* and *Synechococcus* are ubiquitous in these zones, while others are found only episodically, e.g., associated to mesoscale eddies (Fong et al. 2008). Furthermore, the abundance of *Synechococcus* in these regions is typically two orders of magnitude (and its biomass ~20 times) lower than that of *Prochlorococcus*, which is the key photosynthetic organism of these gigantic areas (Partensky et al. 1999a, 1999b).

Many aspects of the biology, physiology, and ecology of *Prochlorococcus* have been discussed in earlier reviews that we recommend. Eleven years after the discovery of *Prochlorococcus*, Partensky et al. (1999b) wrote one of the first general overviews, which included some early molecular biology data. Other reviews focused on the comparative ecology of *Prochlorococcus* and *Synechococcus* (Partensky et al. 1999a, Scanlan & West 2002). Ting et al. (2002) covered aspects dealing with the photophysiology of these two organisms. Garcia-Fernandez et al. (2004) and Garcia-Fernandez & Diez (2004) described some features of nitrogen and carbon assimilation specific to *Prochlorococcus*. Partensky & Garczarek (2003) compared the photosynthetic apparatuses of several atypical cyanobacteria, including *Prochlorococcus*. Finally, more recent reviews by Coleman & Chisholm (2007) and Scanlan et al. (2009) dealt with ecological genomics of *Prochlorococcus* or marine picocyanobacteria in general. In the present review, we discuss one particular question emerging from the recent literature on *Prochlorococcus*, namely, the biological causes of its tiny cell and genome sizes and implications for its ecological success in warm, oligotrophic areas of the ocean. Numerous genomes of *Prochlorococcus* strains (12 published so far, but in fact 13 appear in databases) (**Table 1**), covering most of the ecotypic diversity existing within this genus (Kettler et al. 2007), as well as large metagenomic databases (mainly from surface waters) that include many sequences from natural *Prochlorococcus* populations and their phages (Rusch et al. 2007, Venter et al. 2004), have recently become available. This has brought new insights about the biology, functional diversity, and evolutionary patterns of this microorganism, highlighting the key role that lateral transfer, mostly via cyanophages, plays in the dynamics of its genome. With this wealth of biological and ecological information, this tiny and extremely abundant microorganism is rapidly imposing itself as a new model in microbial ecology, only two decades after its discovery.

PROCHLOROCOCCUS AND OLIGOTROPHY

The success of *Prochlorococcus* in nutrient-poor areas of the ocean and its considerable contribution to the chlorophyll (Chl) biomass—30–60% of the total Chl a is attributable to the *Prochlorococcus*-specific divinyl-Chl a (Chl a_2) in subtropical oligotrophic areas (DiTullio et al. 2003, Gieskes &

Table 1 Main characteristics of the different *Prochlorococcus* genomes sequenced to date

Strain	CCMP[a] no.	RCC[b] no.	Clade name	Genome size (Mb)	No. of protein-coding genes	GC%	Region	Genbank accession no.	References
MED4	1986	151	HLI	1.658	1716[c]/1763[d]/1929[e]	30.8	Mediterrranean Sea	BX548174	(Kettler et al. 2007, Rocap et al. 2003)
MIT9515	3148	—	HLI	1.704	1908	30.8	Eq. Pacific	CP000552	(Kettler et al. 2007)
MIT9301	2971	—	HLII	1.643	1907	31.4	Sargasso Sea	CP000576	(Kettler et al. 2007)
AS9601	2970	—	HLII	1.670	1926	31.3	Arabian Sea	CP000551	(Kettler et al. 2007)
MIT9215	2939	—	HLII	1.738	1989	31.1	Eq. Pacific	CP000825	(Kettler et al. 2007)
MIT9312	2777	—	HLII	1.709	1962	31.2	Gulf Stream	CP000111	(Coleman et al. 2006)
MIT9202	—	—	HLII	1.691 (unfin.)	1891	31.9	S Pacific	ACDW00000000	S.W. Chisholm et al., unpublished data
NATL1A	2968	—	LLI	1.865	2201	35.1	N Atlantic	CP000553	(Kettler et al. 2007)
NATL2A	2969	314	LLI	1.843	2158	35.0	N Atlantic	CP000095	(Kettler et al. 2007)
SS120	1375	154	LLII	1.751	1884[c]/1930[d]/1925[e]	36.4	Sargasso Sea	AE017126	(Dufresne et al. 2003, Kettler et al. 2007)
MIT9211	—	—	LLIII	1.689	1855	38.0	Eq. Pacific	CP000878	(Kettler et al. 2007)
MIT9303	2937	—	LLIV	2.683	3022	50.1	Sargasso Sea	CP000554	(Kettler et al. 2007)
MIT9313	2773	407	LLIV	2.411	2275[c]/2330[d]/2843[e]	50.7	Gulf Stream	BX548175	(Kettler et al. 2007, Rocap et al. 2003)

[a]CCMP: Provasoli-Guillard National Center for Culture of Marine Phytoplankton.
[b]RCC: Roscoff Culture Collection.
[c]Gene number in original publication.
[d]Gene number as recalculated by Dufresne et al. (2008).
[e]Gene number as recalculated by Kettler et al. (2007).

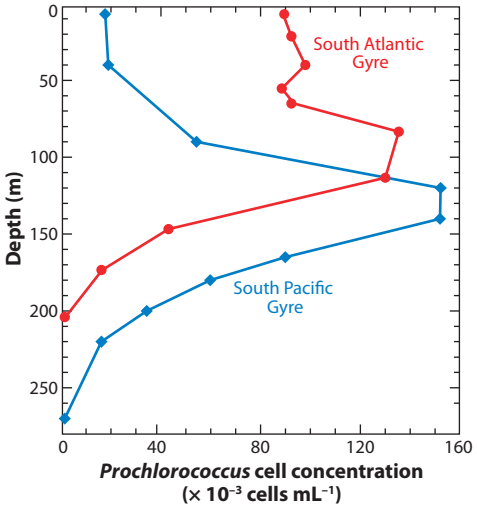


Figure 1

Comparative vertical profiles of *Prochlorococcus* cell concentrations as determined by flow cytometry in the S Atlantic gyre at 23° 54′ S, 25° W (station 57 of AMT13 cruise; data courtesy of S.W. Chisholm and Z.I. Johnson) and in the S Pacific gyre at 26° 18′ S, 114° 06′ E (station GYR of Biosope cruise; data courtesy of C. Grob and O. Ulloa). For the latter, *Prochlorococcus* cell concentrations in the upper 80 m were deduced from divinyl-Chl a concentrations, assuming a divinyl-Chl a cell content of 0.23 fg per cell (see Grob et al. 2007 for details).

Kraay 1983, Goericke & Welschmeyer 1993, Partensky et al. 1996, Ras et al. 2008, Veldhuis & Kraay 2004)—make it a key player in the carbon cycle on a global scale. Its concentration in the upper mixed layer of the S and N Atlantic (**Figure 1**), Indian, and N Pacific gyres is typically above 10^5 cells ml^{-1}. For instance, Campbell et al. (1997) followed the abundance of *Prochlorococcus* on a monthly basis for three consecutive years at the HOT station off Hawaii and determined that it was fairly uniform above 100 m with a median concentration of 1.76×10^5 cell ml^{-1}. However, in the S Pacific gyre, an area as large as the Mediterranean Sea and likely the most oligotrophic region of the world ocean (Claustre et al. 2008), *Prochlorococcus* abundance is comparatively low, with estimated mean cell concentrations below 3×10^4 cells ml^{-1} in the upper 80 m (Grob et al. 2007) (**Figure 1**).

Analyses at the GYR station, near Easter Island (**Figure 2*a***), of a diel cycle of *ftsZ* mRNA levels—a gene involved in cell division—furthermore suggest that, near the surface, *Prochlorococcus* populations are not synchronized by the light/dark cycle, revealing a perturbation of the cell cycle (**Figure 2*c***). Thus, we assume that those cells are growing more slowly than populations located at station EGY, which exhibit a peak of *ftsZ* expression at the end of the day (**Figure 2*c***), as expected (Holtzendorff et al. 2002). Unfortunately, no direct measurements of *Prochlorococcus* growth rates, which would confirm this hypothesis, have been reported thus far for this region. It must be stressed that *Prochlorococcus* concentrations integrated over the water column are not significantly lower in the S Pacific gyre than in other gyres ($\sim 1.6 \times 10^{13}$ cells m^{-2}, a comparable value to that found in the S Atlantic at a similar latitude) because *Prochlorococcus* cells are very abundant at depths to 200 m, with a density maximum at 120 m (**Figure 1**). However, the integrated productivity of *Prochlorococcus* in the S Pacific gyre is likely lower than in other areas. Bulk measurements indeed showed that integrated productivity in this gyre (154–203 mg C m^{-2} day^{-1}) is among the lowest ever reported in comparable marine ecosystems (Van Wambeke et al. 2008).

Gyre: a vast swirling vortex of wind and sea currents that occurs in all oceans with calm waters at the center

The low abundance of *Prochlorococcus* in the S Pacific mixed layer was possibly linked to particularly severe nitrogen deficiency in this region, since iron limitation was evidenced only at the border of the gyre (Behrenfeld & Kolber 1999, Bonnet et al. 2007). Other gyres are also nutrient-depleted but must be re-enriched by aeolian inputs from continental deserts frequently enough to sustain active growth of their tiniest photosynthetic inhabitants (Neuer et al. 2004). In contrast, in the S Pacific gyre, aeolian inputs are almost nil year-round and, surprisingly, Easter Island is apparently too small to have any notable "island effect," i.e., it seems not to contribute significantly in the nutrient enrichment of this oceanic area. UV radiation penetrates very deeply into the water column in this region (Morel et al. 2007) and could also be a contributing factor to the observed *Prochlorococcus* population growth inhibition. However, this is likely a consequence rather than a cause of the low Chl biomass encountered in this part of the world ocean.

Polovina et al. (2008) have recently shown that the surface of oceanic areas exhibiting the lowest chlorophyll concentrations (i.e., below 0.07 $\mu g\ L^{-1}$) is rapidly expanding (**Figure 2*a***), at an annual rate varying between 0.8 and 4.3%, depending on the gyre. These authors assumed that this phenomenon was a direct consequence of global warming, since surface water temperature concomitantly increased over the 9-year period examined, leading to a strong reinforcement of vertical stratification. It follows that the global productivity of the oceans is decreasing fairly rapidly. One may wonder if this phenomenon will consequently induce an increase of the relative contribution of *Prochlorococcus* to global productivity. Surprisingly, the answer is not necessarily yes, at least for the S Pacific. Indeed, an analysis of the pigment data from the Biosope cruise, which sampled a very large Chl gradient from the Marquesas Islands to the Chilean coast (**Figure 2*a***), revealed that the contribution of *Prochlorococcus* to the total Chl biomass (tChla) of (sub)surface waters with chlorophyll below 0.03 $\mu g\ L^{-1}$ is generally lower than 25% (**Figure 2*b***). The ratio of Chl a_2 to tChla can exceed 50% but only between 0.035 and 0.2 $\mu g\ L^{-1}$ tChla. Above this value, Chl a_2 concentration—which is directly related to densities of *Prochlorococcus* cells (see insert in **Figure 2*b***)—can still be high, but the contribution of larger sized phytoplankton groups to tChla becomes predominant. Nevertheless, even if global warming may indirectly provoke a decrease of *Prochlorococcus* contribution to the biomass and productivity in the S Pacific gyre (and possibly other gyres), it will also likely induce an expansion of the distribution of this key cyanobacterium toward higher latitudes in all oceans, so that its global contribution to the carbon cycle will most certainly increase in the near future.

tChla: total concentration of chlorophyll a (in $\mu g\ L^{-1}$)

LL: low-light-adapted

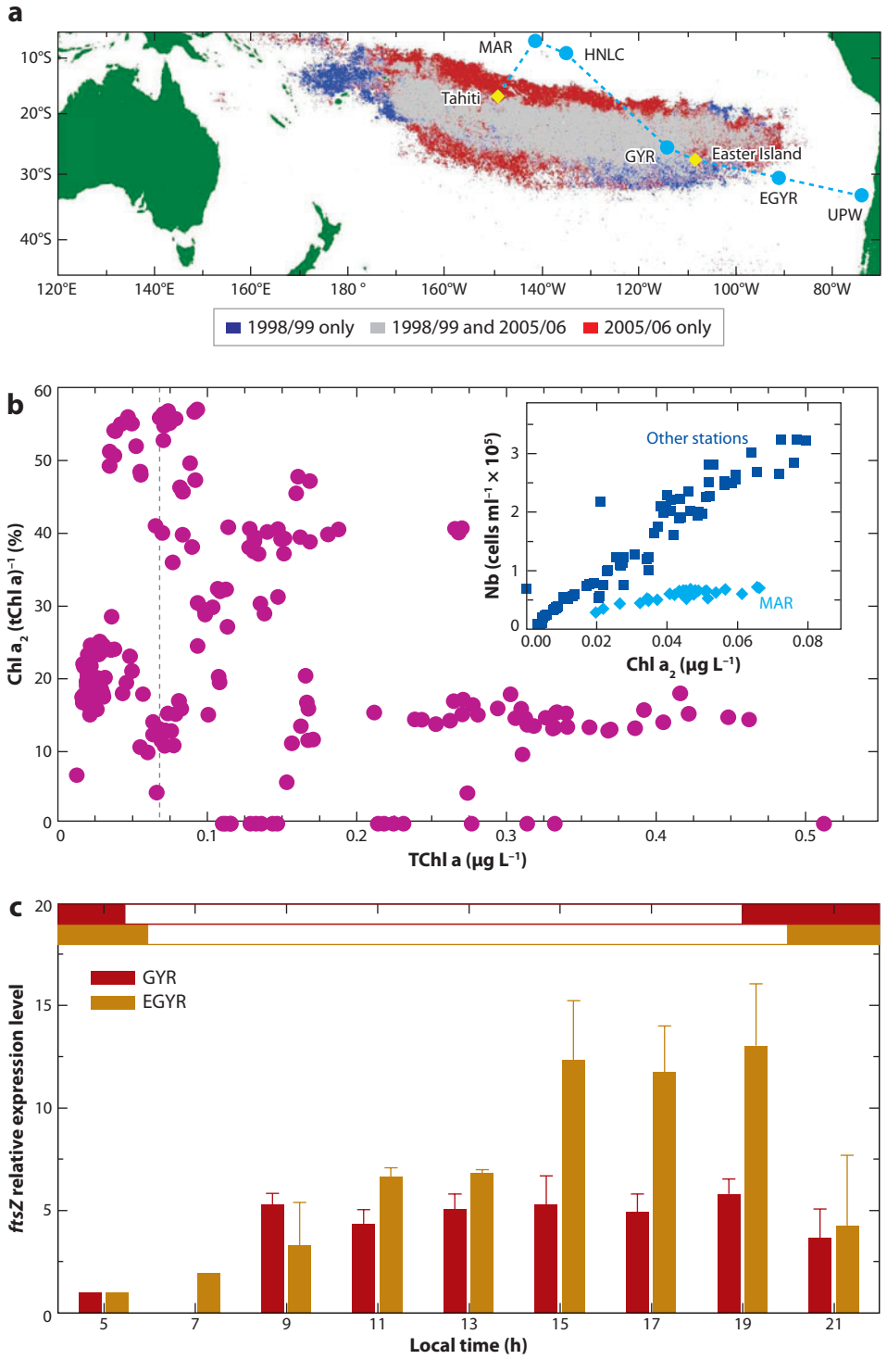
HL: high-light-adapted

ITS: 16S-23S ribosomal DNA internal transcribed spacer

ECOTYPIC DIFFERENTIATION

The presence of low- and high-light-adapted (hereafter LL and HL) ecotypes in *Prochlorococcus* was first suggested by flow cytometry evidence indicating bimodal populations in the red (Chl) fluorescence signal of cells from subsurface waters off Hawaii (Campbell & Vaulot 1993); this was later confirmed by physically sorting cells of each subpopulation and growing them independently (Moore et al. 1998). Isolates obtained from such sorted cells showed distinct light absorption properties due to differing (divinyl-) Chl a to b ratios as well as shifted growth irradiance maxima, with LL cells growing and photosynthesizing at (low) irradiances that could not support growth of HL cells, and conversely. These physiological discrepancies were further supported by significant differences at the genetic level, with HL strains clustering tightly together in 16S rRNA or 16S-23S ribosomal DNA internal transcribed spacer (ITS) trees, well apart from the phylogenetically more variable LL strains (Moore et al. 1998, Rocap et al. 2002, West & Scanlan 1999).

Since these early observations, several more *Prochlorococcus* strains have been isolated, allowing us to refine the initial picture. It is now clear that the HL group is the most recently evolved and comprises at least two main subclades, often called HLI and HLII (West & Scanlan 1999). These



have also been termed eMED4 and eMIT9312, respectively, where the prefix "e" is meant for "ecotype" and the following name is the type strain of the corresponding subclade (Ahlgren et al. 2006, Johnson et al. 2006, Zinser et al. 2006). Members from these two subclades/ecotypes were recently shown by quantitative polymerase chain reaction (qPCR) to have differential geographical distributions along a meridional transect in the Atlantic Ocean from SW Ireland to a station off the Argentina coast (AMT13 cruise; Johnson et al. 2006). HLII was found as the dominant ecotype in the upper mixed layer of stratified waters located between approximately 30°N and 30°S, whereas HLI populations dominated in moderately stratified or fully mixed waters at higher latitude (~30–40°S and 30–45°N) (**Figure 3**). These latitudinal distribution patterns were globally consistent with those observed using a dot blot hybridization approach during another Atlantic Meridional Transect (AMT) cruise (AMT15; Zwirglmaier et al. 2007) as well as during a circumnavigation in the Southern Hemisphere (BEAGLE cruise), where only surface samples were retrieved from the S Pacific, Atlantic, and Indian oceans in winter, late spring, and summer, respectively (Bouman et al. 2006). The differential distribution of HL ecotypes also applies to regional seas, since the Mediterranean Sea was found to be dominated by HLI (Garczarek et al. 2007), whereas the Gulf of Aqaba (Northern Red Sea) is populated by HLII (Zwirglmaier et al. 2008).

Ecotype: a genetically and physiologically unique population that is adapted to its local environment

Johnson et al. (2006) compared the temperature requirements of HLI and HLII strains and showed that the former could grow at temperatures as low as 11–15°C, whilst the latter could not. In contrast, HLII strains grew generally faster than HLI strains at high temperatures and still grew at 30°C, a temperature incompatible with the growth of HLI strains. Together with the differential distribution patterns, these data convincingly show that HLI and HLII truly correspond to different ecotypes adapted to the same light niche but distinct temperature niches. It must be stressed that, on a global scale, HLII are much more abundant than HLI, and this has somehow translated into a much larger number of HLII than HLI strains in culture collections (Rocap et al. 2002).

LL strains can also be separated into a number of clades, based on 16S rRNA or ITS phylogenies, and those clades were again assumed to correspond to as many ecotypes (Ahlgren et al.

Figure 2

(*a*) Expansion of S Pacific areas with lowest Chl a content in surface. Gray color represents the areas that exhibited surface Chl a concentrations ≤0.07 μg L^{-1} in both August 1998/99 and August 2005/06. Blue and red colors represent the extent of areas that had surface Chl a concentrations ≤0.07 μg L^{-1} only in August 1998/99 and only in August 2005/06, respectively. In the S Pacific Ocean, these low-Chl areas are expanding at an average 2.9% per year (adapted with permission from Polovina et al. 2008). The ~8000 km transect and main stations of the Biosope cruise from Tahiti to the Chilean coast are also shown on the map (MAR, HNLC, GYR, EGYR, and UPW correspond to stations sampled near the Marquesas Islands, in the high nutrient–low chlorophyll region, inside the gyre, east of the gyre and in the upwelling, respectively; see Claustre et al. 2008 for details). (*b*) Variation of the ratio of the *Prochlorococcus*-specific divinyl-Chl a (Chl a_2) to total Chl a (tChl a), as measured by high-pressure liquid chromatography (HPLC) analysis in the upper mixed layer along the Biosope transect. Note that the values >0.5 μg L^{-1} tChl a have been omitted for readability since Chl a_2 is generally undetectable in these areas, and that all samples showing a nil Chl a_2 to tChl a ratio are located east of the EGYR station. The dashed line represents the 0.07 μg L^{-1} threshold used by Polovina et al. (2008). The insert represents the relationship between *Prochlorococcus* cell concentration (as measured by flow cytometry) and the Chl a_2 concentration in seawater. Note that data for the MAR station do not follow the general trend (see Grob et al. 2007). Pigment data were obtained courtesy of J. Ras and H. Claustre. (*c*) Diel rhythm of the *ftsZ* gene expression at stations GYR and EGYR of the Biosope cruise, as measured by quantitative polymerase chain reaction (qPCR) (see Holtzendorff et al. 2002 for technical details). The *rnpB* gene was used as an internal standard to normalize the relative transcript level as described in Six et al. (2007a). All data on the graph are relative to the expression level at 5 AM. The two bars on top of the figure indicate the length of the light period at each station, using the same color code as that of the corresponding bar charts.

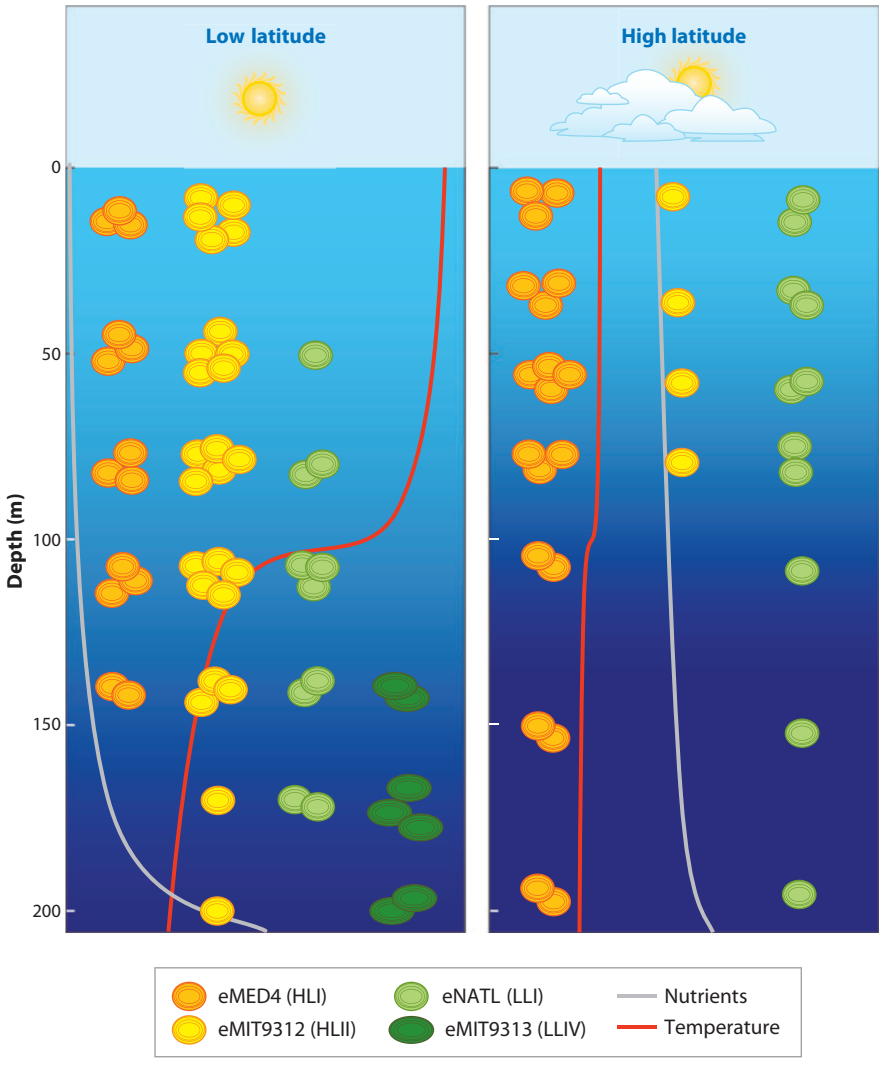


Figure 3

Schematic representation of typical vertical distributions of the four main *Prochlorococcus* ecotypes at low and high latitudes in the Atlantic Ocean. This drawing is based on real data from Johnson et al. (2006) at AMT13 cruise stations 57 (23° 54°S, 25°W) and 69 (36° 24°S, 37° 42°W). Each *Prochlorococcus* cell on the graph symbolizes one order of magnitude of the population cell concentration in nature, i.e., 1 cell corresponds to 10–99 cells mL^{-1}, 2 cells to 100–999 cells mL^{-1}, 3 cells to 1000–9999 cells mL^{-1}, etc.

2006, Johnson et al. 2006, Zinser et al. 2006). Two LL strains (MIT9303 and MIT9313), which were originally sorted by flow cytometry based on their high red fluorescence (Moore et al. 1998), are distinguishable from all other *Prochlorococcus* by their significantly larger cell and genome sizes and higher GC% (**Table 1**). Phylogenetically, this group (called LLIV clade, or eMIT9313 ecotype) is located at the base of the *Prochlorococcus* radiation. In the field, it was found to be abundant below the thermocline over the whole AMT13 transect, except at high latitudes (i.e., above 35°S/40°N), where it is virtually absent (Johnson et al. 2006). The LLI (or eNATL2A; hereafter, eNATL) ecotype, which was absent only at the southernmost stations of the transect, had a depth

AMT: Atlantic Meridional Transect

distribution intermediate between that of LLIV and HLII (i.e., just below the thermocline) in warm, stratified waters between 25°S and 40°N, with rare incursions into surface waters; whereas in mixed water columns at higher latitudes, it co-occurred with HLI in the whole euphotic layer, including surface waters (**Figure 2**). Thus, although initially classified as an LL ecotype, based on its fairly high divinyl Chl b to a ratio (Rocap et al. 2002), the eNATL group is in fact somehow intermediate between HL and LL ecotypes, which is consistent with its position in phylogenetic trees as the closest LL clade from the HL branch (Kettler et al. 2007) and with a number of genomic characteristics shared with one type or the other (see below). In the Mediterranean Sea, a study of *Prochlorococcus* diversity using restriction fragment length polymorphism (RFLP) analyses of the *pcb* gene, encoding the main antenna protein, revealed the occurrence of a genotype with a large depth distribution, which was tentatively attributed to eNATL (Garczarek et al. 2007). The distributions of two other ecotypes, eSS120 (LLII) and eMIT9211 (LLIII), were also analyzed by qPCR (Ahlgren et al. 2006, Johnson et al. 2006, Zinser et al. 2006) and both were found to be scarce at all stations examined. However, this should likely be re-examined since qPCR primers used for these two clades were both designed based on a single sequence (Ahlgren et al. 2006).

RFLP: restriction fragment length polymorphism; a molecular method by which the migration of DNA fragments is analyzed following exposure to specific restriction enzymes

OTU: operational taxonomic unit; a group of organisms that is defined by its sequence similarity

As already stated, the diversity of the LL group as a whole is much wider than that of the HL group. Since there are fewer LL strains than HL strains in culture collections (Rocap et al. 2002), the diversity of the LL group is currently underestimated. To better assess the true diversity of natural *Prochlorococcus* populations, Martiny et al. (2009c) recently cloned and sequenced a large number of 16S-23S rRNA ITS regions at different depths from 10 sites in the Atlantic and Pacific oceans. They found that, while most sequences were related to clades/ecotypes with cultured representatives, there were numerous novel sublineages within each clade. Additionally, more than 23% of the sequences from deep samples belonged to a novel, still uncultured *Prochlorococcus* clade (provisionally termed NC1), which was deeply branching within the LL group. Another study, aimed at characterizing the diversity of marine picocyanobacteria from the oxygen minimum zone in the eastern tropical Pacific using a T-RFLP analysis of the ITS region, also revealed the occurrence of two novel uncultivated clades, tentatively called LLV and LLVI (P. Lavin, B. Gonsalez, J.F. Santibanez, D.J. Scanlan, O. Ulloa, unpublished paper). Together with LLIV clade members, they contributed to ~90% of the *Prochlorococcus* diversity in the low-oxygen layer. Consistent with their ecology, both LLV and LLVI clades fall at the base of the *Prochlorococcus* radiation in ITS phylogenies in a branch shared with LLIV strains, and it is quite possible that they also possess a fairly large genome.

There is considerable microdiversity within *Prochlorococcus* ecotypes. In the Mediterranean Sea, Garczarek et al. (2007) showed that there were virtually no identical sequences—even taking the PCR error rate into account—in *pcb* gene clone libraries. In contrast, very similar RFLP patterns of *Hae*III-digested *pcb* genes could be found at given depths of very distant stations, even though these patterns varied extensively with depth at any given station. Thus, while *pcb* gene sequences provided information on the genotypic diversity of *Prochlorococcus* populations, RFLP patterns corresponded to an intermediate degree of genetic diversity between genotype and ecotype. The delimitation between these two levels is therefore delicate to clearly assign. Martiny et al. (2009c) found that when an operational taxonomic unit (OTU) was defined as a group of *Prochlorococcus* cells exhibiting 90% or more similarity of the ITS sequence, there were only 1–5 OTUs in surface and mid-depth samples (though more in deep samples), whereas with a more stringent definition (e.g., >97% similarity), the number of OTUs increased significantly in most samples and rarefaction curves did not saturate. Community composition was correlated to dispersal time only when OTUs were defined at a very high cut-off value (99.5% similarity), suggesting that *Prochlorococcus* cells may evolve faster than currents can mix them, creating local microdiversity.

Specific Features of the *Prochlorococcus* Genus

Photosystems: protein complexes made of several subunits that are converting the electromagnetic energy from sun into chemical energy usable by the cell

Phylogenetic analyses using a number of genes (16S rRNA, *rpoC1*, *ntcA*) or intergenic regions (ITS, *petB-D*) (see, e.g., Palenik 1994, Penno et al. 2006, Rocap et al. 2002, Urbach et al. 1998) converge to show that (*a*) marine picocyanobacteria (i.e., *Prochlorococcus* and *Synechococcus*) form a branch well separated from all other cyanobacteria, including freshwater *Synechococcus*, but also from other marine planktonic cyanobacteria of larger cell size, such as *Crocosphaera* and *Trichodesmium*, and (*b*) *Prochlorococcus* constitutes a monophyletic group within the marine *Synechococcus* radiation. The differentiation of the *Prochlorococcus* lineage involved some founder events unifying all members of this genus. These include the replacement of the large, extrinsic antenna complexes (phycobilisomes) still found in extant marine *Synechococcus* (see Six et al. 2007b for a review) by compact antennae located within thylakoid membranes in direct contact with photosystems (PSs) (Bibby et al. 2003, Garczarek et al. 2001, LaRoche et al. 1996). These complexes are composed of Chl-binding Pcb proteins, which are strongly related to the iron stress-induced protein IsiA. The latter was shown to form an 18-protein ring around PS I trimers in freshwater cyanobacteria grown under iron deplete (Fe-) conditions (Bibby et al. 2001a, Boekema et al. 2001) and a similar structure has been described in the LL *Prochlorococcus* strain MIT9313 under Fe- only and in SS120 under both Fe+ and Fe- conditions (Bibby et al. 2003, Bibby et al. 2001b).

In the early stages of the differentiation of the *Prochlorococcus* genus, the gene encoding this PSI-associated Pcb protein must have been duplicated and the product of one of the two gene copies evolved to become PS II–specific, an event that also occurred independently in the two other green oxyphotobacteria (often improperly called prochlorophytes) *Prochlorothrix* and *Prochloron*, as well as in the Chl d–containing prokaryote *Acaryochloris* (see Chen & Bibby 2005, Partensky & Garczarek 2003 for reviews). Such antenna exchange induced a drastic change in absorption properties of the *Prochlorococcus* cells, with the acquisition of divinyl derivatives of Chl a and b (so-called Chl a_2 and b_2; Goericke & Repeta 1992). With their unique suite of pigments, *Prochlorococcus* cells are highly specialized for the capture of blue wavelengths prevailing in oligotrophic waters. Synthesis of Chl b_2 from Chl a_2 is catalyzed by a *Prochlorococcus*-specific Chl a oxygenase (PcCao), which is only distantly related to the Cao protein found in higher plants and other green oxychlorobacteria (Satoh & Tanaka 2006). The presence of divinyl- instead of monovinyl-Chls has been suggested to result from the loss of the 3,8-divinyl protochlorophyllide *a* (*dvr*) gene (Kettler et al. 2007, Nagata et al. 2005), but several marine *Synechococcus* strains also seemingly lacking the *dvr* gene do produce normal (monovinyl-) Chl a (Dufresne et al. 2008, Scanlan et al. 2009). All *Prochlorococcus* also have a specific lycopene cyclase (CtrL-e) responsible for the synthesis of α-carotene (Hess et al. 2001, Stickforth et al. 2003). Despite low amounts in the cell (Goericke & Repeta 1992, Moore et al. 1995, Partensky et al. 1993), this pigment must be important for the correct functioning of the photosynthetic apparatus of *Prochlorococcus*.

In a recent comparison of 12 *Prochlorococcus* and 4 marine *Synechococcus* genomes, Kettler et al. (2007) identified only 13 clusters of orthologs that were supposedly specific to the *Prochlorococcus* genus. However, when adding 7 more marine *Synechococcus* genomes in the comparison, orthologs of many of these genes were in fact retrieved in one or several *Synechococcus* strains. **Table 2** lists genes that are present in all 12 published *Prochlorococcus* but none of the 11 *Synechococcus* sequences. Several of the orthologs listed by Kettler et al. (2007) are indeed *Prochlorococcus*-specific, but sequences found in HL and LL strains were in fact fairly distantly related to each other (i.e., potentially have different functions, possibly linked to adaptation to their specific niches) and have been placed in different clusters of orthologs by the clustering method used by Dufresne et al. (2008). Other genes such as orthologs of PMM1027 are only found in *Prochlorococcus*, but not

Table 2 List of *Prochlorococcus*-specific genes

Cyanorak cluster no.[a]	MED4 locus tag in MoL[b]	MED4 locus tag in Genbank[c]	Gene name	Product	References
173	PMED4_06771	PMM0627	*pcb*	PSII-associated Pcb antenna protein	Laroche et al. 1996, Bibby et al. 2003
2331	PMED4_08941	PMM0808	*PcCao*	Chlorophyll b synthase (Chlorophyllide a oxygenase)	Nagata et al. 2005
9152[d]	PMED4_06831	PMM0633	*crtL*-e	Lycopene ε-cyclase	Stickforth et al. 2003
3411	PMED4_06861	PMM0636		Isochorismatase-like hydrolase superfamily	Kettler et al. 2007
2026	PMED4_12891	PMM1135	*hli*	High-light-induced protein (HLIP)	≡ Cluster 10 in Bhaya et al. 2002
2025	PMED4_15741	PMM1385	*hli*	High-light-induced protein (HLIP)	≡ Clusters 12 + 19 in Bhaya et al. 2002
3473	PMED4_11001	PMM0983	—	Secreted protein (1 signal peptide predicted)	Kettler et al. 2007
3414	PMED4_07061	PMM0655	—	Conserved hypothetical protein	Kettler et al. 2007
2333	PMED4_11581	PMM1022	—	Conserved hypothetical protein	Kettler et al. 2007
2639	PMED4_12731	PMM1125	—	Conserved hypothetical protein	Kettler et al. 2007

[a]Number of the cluster of orthologs in the Cyanorak database (**http://www.sb-roscoff.fr/Phyto/cyanorak**).
[b]Locus tag of the *P. marinus* MED4 sequence in the Microbes-on-Line database.
[c]Locus tag of the *P. marinus* MED4 sequence in the Genbank database.
[d]Sequences of *Prochlorococcus crtL-e*, a lycopene cyclase that produces α-, β-, δ-, and ε-carotene when heterologously expressed in *Escherichia coli* (Stickforth et al. 2003), cluster apart from sequences of *Prochlorococcus crtL-b*, a lycopene cyclase that forms only β-carotene (Cyanorak cluster no. 86).

in all strains. On the other hand, some genes that are known to be *Prochlorococcus*-specific were not retrieved by the clustering program used by Kettler et al. (2007), e.g., PSII-specific *pcb* genes or *crtL-e*, likely because they have phylogenetically close paralogs (namely, PSI-specific *pcb* and *crtL-b* genes, respectively; see Garczarek et al. 2007, Stickforth et al. 2003) and the discrimination threshold was likely too stringent to split them apart. This observation suggests that **Table 2** is probably not exhaustive and that more *Prochlorococcus*-specific core genes will be found in future as a result of more refined genome comparisons. Also, some genes in **Table 2** have no or only broad predicted function and should be privileged targets for characterization, as this will help determine new traits specifically acquired during the differentiation of the *Prochlorococcus* genus.

Gene Loss Associated with the Differentiation of the *Prochlorococcus* Genus

One of the most striking events to have occurred during the evolution of the *Prochlorococcus* group is certainly the extensive genome streamlining that has affected most lineages. It must be stressed that *Prochlorococcus* is, with *Candidatus* Pelagibacter ubique, the first sequenced representative of the most abundant marine bacterial group SAR11 (Giovannoni et al. 2005), the only free-living organism known so far to have undergone such a genome reduction process, and this is likely related to adaptation to oligotrophic conditions prevailing in oceanic gyres (Dufresne et al. 2005). The first three *Prochlorococcus* strains to be sequenced (namely, the LL strains MIT9313 and SS120 and the HL strain MED4) showed widely different genome sizes (2.41, 1.75, and 1.66 Mbp, respectively), which at first sight had suggested a progressive decrease of this feature during *Prochlorococcus* evolution (Dufresne et al. 2005, Dufresne et al. 2003, Rocap et al. 2003). In fact,

*Erratum

*This PDF amended on (24 Feb. 2010): See explanation at http://arjournals.annualreviews.org/errata/marine

G+C%: percentage of the genome content composed of cytosine (C) and guanine (G)

the recent availability of many complete genomes representing all cultivated clades (**Table 1**) has shown that the only *Prochlorococcus* strains that have kept a relatively large genome size are the LLIV strains MIT9303 and MIT9313, i.e., those two strains isolated by cell sorting and not filtration (typically through 0.6 μm filters) coupled with serial dilutions, a method used for the other *Prochlorococcus* strains (Chisholm et al. 1992, Moore et al. 2007, Partensky et al. 1993). LLIV cells also exhibit larger cell dimensions and a ~2.2-fold larger volume than other *Prochlorococcus* (Ting et al. 2007). This suggests that cell volume is tightly linked to genome size and hence that these two features likely decreased concomitantly during evolution. The genome characteristics of LLIV strains resemble more those of marine *Synechococcus*—with which *Prochlorococcus* share a common ancestor—than those of other *Prochlorococcus* strains. Interestingly, although much more ancient and diversified, marine *Synechococcus* show a low variability of genome size (2.22–2.86 Mbp) and GC content (52.5–66.0 G+C%) compared to *Prochlorococcus* (1.64–2.68 Mbp and 30.8–50.7 G + C%; Dufresne et al. 2008, Kettler et al. 2007), and thus, seemingly, no *Synechococcus* lineage has undergone any significant genome streamlining.

Prior to the main streamlining process, which was associated with the occurrence of novel ecotypes (see below), the differentiation of the *Prochlorococcus* genus itself already involved the loss of a number of useful but dispensable core genes that were present in the genome of its last common ancestor with marine *Synechococcus*. Kettler et al. (2007) retrieved 140 genes that were absent in 12 *Prochlorococcus* but present in 4 *Synechococcus* taken as a reference. This list shrunk to 70 genes when Dufresne et al. (2008) compared 11 *Synechococcus* to 3 reference *Prochlorococcus* strains, so at least half of the genes in the previous list were in fact accessory in *Synechococcus* genomes. The set of *Synechococcus*-specific core genes (see Dufresne et al. 2008, add'l data file 1) will probably decrease again as more marine picocyanobacteria genomes are included in the comparison; it comprises all allophycocyanin and phycocyanin biosynthesis genes that became useless after the differentiation of Pcb antenna complexes. More surprisingly, a number of other photosynthetic genes have also been eliminated, including some genes involved in the carbon concentration mechanism, namely, homologs of *Synechocystis* sp. PCC 6803 *ndhD4* and *ndhF4* involved in low-affinity CO_2 uptake, the CO_2 hydration protein ChpX, and a putative carbonic anhydrase (Badger & Price 2003, Dufresne et al. 2008). Also missing are the two subunits of the ferredoxin-thioredoxin reductase as well as one specific ferredoxin and one thioredoxin, a set of genes that might be involved in the light-mediated regulation of Calvin cycle enzymes (Dufresne et al. 2008). *Prochlorococcus* also possess no *psbA* gene encoding the D1:2 isoform of the D1 protein, known to confer better resistance of PSII to photoinhibition. Instead, 1 to 3 *psbA* copies found in *Prochlorococcus* strains all encode a D1:1-like isoform (Garczarek et al. 2008).

Other metabolic pathways have also been affected by gene losses that occurred during the early stages of the *Prochlorococcus* genus differentiation. For instance, all *Prochlorococcus* lack *kaiA*, encoding one of the three components of the core oscillator of the circadian clock (Holtzendorff et al. 2008). Interestingly, a short remnant of the *kaiA* gene is in fact found upstream from *kaiB* in the LLIV strains, but this is likely a pseudogene since its predicted product (corresponding to the C-terminus of the protein) lacks many amino acids conserved in other KaiA. Implications of the lack of KaiA were studied in the HL *Prochlorococcus* strain PCC 9511 (Holtzendorff et al. 2008). It was shown to result in a complete loss (within 24 h) of the synchronization of the DNA replication timing and of the diel oscillation of the *psbA* gene expression (and probably the whole transcriptome) when cells previously entrained under a light:dark cycle were shifted to continuous light. It was concluded that *Prochlorococcus* likely possess a clock working on a 24-h basis in an hourglass-like fashion, rather than as a self-sustained oscillator (Holtzendorff et al. 2008).

Although all *Prochlorococcus* strains sequenced so far lack all genes linked to nitrate metabolism [including genes encoding the nitrate transporter NrtP (a.k.a. NapA), those encoding the nitrate

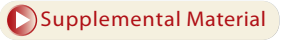
reductase NarB, and seven genes involved in the biosynthesis of a molybdenum cofactor of nitrate reductase], these genes are not present in the *Synechococcus*-specific core set. Indeed, one *Synechococcus* strain (RS9917) out of eleven sequenced ones also lacks this whole suite of genes, which was most likely present in the common ancestor of *Prochlorococcus* and *Synechococcus*. In fact, Casey et al. (2007) showed that *Prochlorococcus* cells sorted from deep populations in the Sargasso Sea displayed some specific NO_3^- uptake, suggesting that these natural populations may have kept (or reacquired by lateral transfer?) nitrate metabolism genes. Recently, Martiny et al. (2009b) found direct evidence of nitrate assimilation genes in uncultured *Prochlorococcus* by screening metagenome databases from surface oceanic waters.

Genes encoding nitrite reductase (NirA), its siroheme cofactor (CobA, a.k.a. CysG), and the putative nitrite transporter FocA, i.e., the minimal gene set for nitrite uptake and assimilation, have been maintained in four out of the six LL *Prochlorococcus* strains (MIT9313, MIT9303, and the two NATL strains). All other strains rely only on ammonium and organic nitrogen forms (urea, amino acids, and sometimes cyanate), as suggested by their nitrogen-related gene set and/or by growth experiments (Mary et al. 2008, Moore et al. 2002).

Genome Streamlining Associated with Ecotypic Differentiation

As mentioned above, the actual genome streamlining process seemingly started just after the differentiation of the LLIV branch from the common ancestor of all other *Prochlorococcus* lineages. Based on an estimated genome size of 2.5 Mbp for the ancestral *Prochlorococcus* genome (i.e., equivalent to the average size of extant *Synechococcus* and *Prochlorococcus* LLIV genomes) and a gene density of 1 gene every 874 bp (i.e., the average value for sequenced *Prochlorococcus*), one may assess the total gene loss as ~930 genes for HL and LLII/III strains, i.e., approximately one-third of the hypothetical ancestral genome.

However, only 130 genes are present both in all sequenced marine *Synechococcus* and in LLIV strains but are absent from all other *Prochlorococcus* genomes (see **Supplemental Table 1**; follow the **Supplemental Material** link from the Annual Reviews home page at **http://www.annualreviews.org**). This implies that while all these lineages have undergone massive gene losses, different lineages must have lost distinct sets of genes. Among the genes lacking in all streamlined *Prochlorococcus* genomes, the most notable are a number of DNA repair genes (see **Table 3** and corresponding discussion below), two genes (*psb*U/V) involved in the stabilization of the PS II oxygen-evolving complex, three genes encoding the subunits of the glycolate oxidase (i.e., a complex involved in photorespiration), and eight genes coding two distinct ABC transporters (including a Ggt homolog) that are probably used for the uptake of compatible solutes available in the immediate cell environment or to prevent leakage of accumulated sucrose (Scanlan et al. 2009).



It must be stressed that even if gene loss has clearly been the dominant process during evolution of these streamlined genomes, all of them have also continuously (but at a slower pace) reacquired genes by lateral transfer from other microorganisms co-occurring in their environment (Kettler et al. 2007). Given the absence of plasmids in *Prochlorococcus* cells, most of these exchanges must have proceeded via phages, though natural transformation cannot be excluded. Thus, in spite of their much reduced size, a significant degree of variability in gene composition can be found among reduced genomes, the so-called accessory or flexible genome accounting for approximately one-third of the genome in the most streamlined HLI/II and LLII/III strains (Kettler et al. 2007).

Sets of genes specifically retained or acquired by given ecotypes likely play a crucial role in adaptation to their particular niches. Examination of the set of genes present in HL but not LL strains (and vice versa) can provide a first assessment of this ecotypic differentiation

Table 3 List of genes potentially involved in DNA replication, recombination, and repair that are only present in a subset of *Prochlorococcus* strains

Pathway[a]	Gene name	Product	Type of mutation (in *E. coli* orthologs)[b]	Cyanorak cluster no.	Locus tag of one representative gene in MoL	Locus tag of one representative gene in Genbank	HL-I		HL-II				LL-I		LL-II	LL-III	LL-IV	
							MED4	MIT 9515	MIT 9301	AS 9601	MIT 9215	MIT 9312	NATL1A	NATL2A	SS120	MIT 9211	MIT 9303	MIT 9313
SSBR[c]–Direct repair	*phrA*	DNA photolyase	—	1330	PMED4_02901	PMM0285	1	1	1	1	1	1	1	1	0	0	0	0
	phrB/cry2	DNA photolyase (or cryptochrome?)	—	1584	NATL1_03651	NATL1_03651	0	1	1	1	1	0	0	0	0	0	0	0
	—	DNA photolyase-related protein	—	1460	PMED4_04701	PMM0425	1	1	1	1	1	1	1	1	0	0	0	0
	—	FAD domain-containing protein	—	3563[j]	PMED4_15291	PMM1360	1	1	1	1	1	1	1	1	0	0	0	0
	—	Alkylated DNA repair protein, AlkB superfamily	—	1486	P9313_10361	PMT0878	0	0	0	0	0	0	0	0	1	1	1	1
SSBR–BER[d]/NER[e]/MMR[f]	—	ATP-dependent DNA ligase	—	3613[j]	PMED4_18901	PMM1679	1	1	1	1	1	1	0	0	0	0	0	0
	—	ATP-dependent DNA ligase	—	1447	PMED4_08061	PMM0729	1	1	1	1	1	1	0	0	0	0	0	0
	—	NAD^+-dependent DNA ligase (adenylation domain)	—	9149[j]	PMED4_07121	PMM0659	1	2	2	2	2	1	2	2	0	0	0	0
	ligA	NAD^+-dependent DNA ligase	—	1185	P9313_27391	PMT2184	0	0	0	0	0	0	1	1	1	1	1	1
SSBR–MMR	*vsr*	DNA mismatch endonuclease of very short patch repair	G:C → A:T	–[j]	MIT9215_847	P9215_04531	0	1	0	1	1	1	0	0	0	0	0	0
	xseA	Exonuclease VII large subunit	—	1316	P9313_20731	PMT1641	0	0	0	0	0	0	1	1	1	1	1	1
	xseB	Exonuclease VII small subunit	—	1730	P9313_20741	PMT1642	0	0	0	0	0	0	1	1	1	1	1	1
SSBR–BER	*mutY*	A/G-specific adenine glycosylase	G:C → A:T	306	P9313_01441	PMT0135	0	0	0	0	0	0	1	1	1	1	1	1
	—	Pyrimidine dimer DNA glycosylase	—	2679[j]	P9313_09911	PMT0842	0	0	0	0	0	0	0	0	1	1	1	1
	nei	Endonuclease VIII	G:C → A:T	–[j]	NATL1_13271	NATL1_13271	0	0	0	0	0	0	1	1	0	0	0	0

SSBR–NER	*ybaZ*	6-*O*-alkylguanine DNA alkyltransferase-like	G:C → A:T[i]	1425	P9313_03171	PMT0269	0	0	0	0	0	0	0	0	0	0	1	1
SSBR–Other factors	*dprA*	DNA-protecting protein DprA	—	1400	P9313_21031	PMT1661	0	0	0	0	0	0	0	0	0	0	1	1
DSBR[g]–HR[h]	*recQ*	ATP-dependent DNA helicase	—	1211	P9313_02081	PMT0189	0	0	0	0	0	0	0	0	0	0	1	1
	recJ	Single-stranded DNA-specific exonuclease	—	1259	P9313_08931	PMT0761	0	0	0	0	0	0	1	1	1	1	1	1
	—	P-loop containing nucleoside triphosphate hydrolase, related to RecD (viral-like)	—	1296	P9313_14111	PMT1107	0	0	0	0	0	0	1	1	1	1	1	1
DSBR–Other factors	—	Lhr-like helicase	—	1446	PMED4_08051	PMM0728	1	1	1	1	1	1	0	0	0	0	0	0
	brpB	ATP-dependent helicase	—	1530	P9313_20701	PMT1639	0	0	0	0	0	0	0	0	0	0	1	1
	—	ATP-dependent helicase	—	1351	P9313_00371	PMT0037	0	0	0	0	0	0	0	0	0	0	1	1
	—	ATP-dependent helicase	—	1504	P9313_14791	PMT1168	0	0	0	0	0	0	0	0	0	0	1	1
	—	Superfamily II DNA/RNA helicases	—	413	P9313_27071	PMT2158	0	0	0	0	0	0	0	0	1	1	1	1
Other–Nudix hydrolases	—	NUDIX hydrolase superfamily	—	1443	P9313_13001	PMT1026	0	0	0	0	0	0	0	0	0	0	1	1
	nudD	GDP-mannose mannosyl hydrolase	—	–[j]	MIT9215_1849	P9215_14401	0	1	0	0	1	0	0	0	0	0	0	0
Other-Pyrimidine metabolism	*udk*	Putative uridine kinase	—	3476[j]	PMED4_11091	PMM0989	1	1	1	1	1	1	1	1	0	0	0	0
Other–DNA restriction	—	HNH family nuclease	—	9150[j]	PMED4_10981	PMM0982	1	1	1	1	1	1	0	0	0	0	0	0

[a] Genes are classified according to DNA repair pathways, mainly as described in the *Kyoto Encyclopedia of Genes and Genomes* (*KEGG*).
[b] When the orthologs of these genes have been shown in *E. coli* to be mutator genes (Miller 1998), the type of mutation involved is indicated.
[c] SSBR, single-strand breaks repair.
[d] BER, base excision repair.
[e] NER, nucleotide excision repair.
[f] MMR, methyl-directed mismatch repair.
[g] DSBR, double-strand breaks repair.
[h] HR, homologous recombination.
[i] This transversion effect was shown for *ada* and *ogt* genes and is likely to occur for *ybaZ* as well.
[j] This cluster of orthologs does not include any marine *Synechococcus* sequence, suggesting acquisition by lateral transfer.

Transversions: substitution of a purine for a pyrimidine, or vice versa

(Kettler et al. 2007, Rocap et al. 2003). However, the occurrence of the LLI ecotype/clade makes the story more complex. With genome sizes approximating 1.85 Mbp, representing ~250 genes more than in other streamlined genomes, NATL strains stand somehow apart. This slightly larger genome size and gene content is likely related in part to their intermediate (and hence more variable) habitat in the field (**Figure 3**). Indeed, NATL strains display a number of genetic characteristics that are otherwise either HL- or LL-specific. Among the characters they share with SS120 and MIT9211, NATL strains have a large number of *pcb* gene copies, encoding six PSII-associated and one PSI-associated Pcb proteins, whereas HL strains have either one (in MED4) but most often two *pcb* gene copies, one for each photosystem (Garczarek et al. 2007, Kettler et al. 2007). Like all LL strains, NATL strains have also kept all genes necessary to synthesize a complete phycoerythrin, whereas HL strains have kept only a small subset of these genes (Hess et al. 1996, Hess et al. 2001, Hess et al. 1999, Steglich et al. 2003, Ting et al. 2001). The putative ecological advantage for life at depth conferred by these two features is, however, not yet fully understood.

In contrast, NATL strains possess 41 high-light-induced protein (HLIP) genes (Kettler et al. 2007), i.e., many more than other LL strains (e.g., 6 in SS120 and 9 in MIT9313) but also significantly more than HL strains themselves (e.g., 22 in MED4, Bhaya et al. 2002; 24 in MIT9312, Coleman et al. 2006). This very large number of *hli* genes most likely translates the need for these cells to actively protect their photosystems, and especially the Pcb antennae that surround them (Bibby et al. 2003, Bibby et al. 2001b), against the deleterious effects of high light and UV stress when they are present in surface waters, notably at high latitude (**Figure 3**). By comparison, none of the marine *Synechococcus* have more than 14 *hli* genes (Palenik et al. 2006), despite the fact that they preferentially thrive in the upper mixed layer. Thus, phycobilisomes appear better suited to dissipate excess photon energy than Pcb's, likely because of their localization in the stromatic space, their ability to quickly disconnect from photosystems (Six et al. 2007a) and/or their efficient coupling with the orange carotenoid protein (OCP), which was found to play a crucial role in energy dissipation (Wilson et al. 2006). Many *hli* genes have seemingly been acquired by lateral transfers via cyanophages (Coleman et al. 2006, Kettler et al. 2007), as suggested by the fact that *Prochlorococcus* phages often contain *hli* genes (Lindell et al. 2004), whereas others have been acquired by gene duplication. In MED4, for instance, a 4-*hli* gene cluster is found in two copies (*hli06–09* and *hli16–19*) in the genome, and these have exactly the same sequence at the nucleotide level, suggesting that they result from a very recent duplication event (Bhaya et al. 2002).

The metabolic category in which one possibly finds the largest differentiation between the different ecotypes is DNA replication, recombination, and repair (**Table 3**), and these discrepancies have probably played a key role in genome evolution of the different ecotypes, as discussed later in this review. A rapid examination of **Table 3** shows that HL strains have preferentially retained (or recruited) genes related to photolyases and DNA ligases, whereas true LL strains possess a particularly large number of DNA helicases (besides the indispensable ones shared by all genomes). Again, NATL strains appear to have a finger in every pie since they possess DNA repair genes from both HL and LL ecotypes, plus one specific endonuclease VIII (Nei) homolog that might confer them an additional protection against G:C to A:T transversions. This large DNA repair gene set may translate the need of NATL strains for a particularly high level of protection against DNA damages.

Photolyases are photoreactive enzymes involved in the repair of pyrimidine (mainly thymine) dimers generated during exposure to UV radiation. Like MED4, NATL strains possess one true photolyase—characterized by the presence of two chromophore-binding domains: namely, an N-terminal 8-hydroxy-5-deazariboflavin domain and a C-terminal flavin adenine dinucleotide (FAD) domain—one FAD monodomain protein, and one uncharacterized photolyase-related protein.

Four out of six HL strains also possess an additional photolyase (or possibly a cryptochrome; see Goosen & Moolenaar 2008). In true LL strains, photolyases have been replaced by pyrimidine dimer DNA glycosylase, a (viral-like) nonphotoreactive analog of photolyases that has seemingly been recruited by lateral transfer from phages (Goosen & Moolenaar 2008).

DNA ligases catalyze the joining of breaks in double-stranded DNA, utilizing either NAD^+ or ATP as a cofactor. All LL strains (including NATL) possess a typical NAD^+-dependent DNA ligase (LigA). As in *Thermus filiformis* (Lee et al. 2000), this ligase comprises four domains: (*a*) a large N-terminal domain involved in adenylation, (*b*) an oligomeric-binding (OB) fold involved in DNA binding, (*c*) one domain constituted of a zinc finger motif and four helix-hairpin-helix (HhH) motifs (also involved in DNA binding), and (*d*) a BRCT domain—initially named after the C-terminal domain of a breast cancer susceptibility protein—which acts as a phosphoprotein-binding domain. A circular arrangement of these four domains in the closed conformation of the protein creates a large hole that can accommodate a double-stranded DNA molecule (Lee et al. 2000).

HL strains share with NATL strains the presence of one or two short, putative open reading frames comprising only the adenylation domain of a NAD^+-dependent DNA ligase (**Table 3**). Given the absence of any DNA-binding domain, it is not yet clear by what mechanism these putative enzymes could bind to DNA in order to perform their catalytic function. HL strains also possess two ATP-dependent DNA ligases, one found also in marine *Synechococcus* and the other specific to these strains. Contrary to the previous ligases, these comprise, like their homologs in viruses (Pascal 2008), both an adenylation domain and an OB domain—used to assist in formation of the ligase-AMP intermediate—and therefore they should be fully functional on their own.

While the whole set of genes required in the nucleotide excision repair (NER) pathway—including *uvrABC*, encoding the three subunits of the excinuclease; *uvrD*, encoding the ATP-dependent DNA helicase; and *polA*, encoding DNA polymerase I—is present in all *Prochlorococcus*, only a few enzymes of the methyl-directed mismatch (MMR) pathway, best characterized in *Escherichia coli* (Li 2008), can be readily identified in *Prochlorococcus* genomes. Only a subset of HL strains seems to possess a DNA mismatch endonuclease of very short patch repair (Vsr), a key enzyme for the repair of methylated cytidines that can be mutated into thymines, leading to G:C to A:T transversions. All LL strains possess genes encoding the two subunits of exonuclease VII, an enzyme which in *E. coli* is one of the four nucleases of the MMR process involved in eliminating mismatch-containing oligonucleotides. LL strains also possess two enzymes of the homologous recombination pathway (RecJ, another exonuclease and a protein related to RecD). It is not yet clear, given a habitat characterized by low photon fluxes and absence of UV light, why true LL strains (i.e., members of LLII/III/IV clades) have kept a fairly large set of DNA repair genes.

Adaptation to Local Niches at the Genotype or Population Level

All adaptive processes described above are either common to the whole *Prochlorococcus* genus or to a specific phototype (HL, eNATL, or true LL) and date back to the ancestors of these different groups. However, other metabolic processes seem to have been acquired much more recently. A well-documented case of recent adaptation is phosphorus (P) metabolism. Even though P is indispensable to cell survival—and therefore some P genes are present in all *Prochlorococcus* genomes, including those encoding the orthophosphate transport system PstABC and the cell wall–associated phosphate-binding protein PstS—a wide range of strategies exists in these organisms for scavenging, transporting, and assimilating the diverse forms of P (reviewed in Scanlan et al. 2009). This translates in *Prochlorococcus* strains into a large variability in the content of genes involved in P acquisition and regulation (Martiny et al. 2006, 2009a).

Genomic islands: part of the genomes with a hypervariable gene content, which may differ even between very closely related strains

For instance, MED4, a HLI strain isolated from the Mediterranean Sea, where P is notoriously limiting (Moutin et al. 2002), possesses 15 more genes involved in P metabolism (e.g., *pho*A encoding alkaline phosphatase, one porin gene involved in phosphate transport, and several regulation genes such as *pho*B/R and *ptr*A) than MIT9515, a HLI strain isolated from the tropical Pacific, a P replete area. Similar differences can be found among the different strains of the HLII clade (Martiny et al. 2006). These comparisons have suggested a direct relationship between the number of genes related to P metabolism and the P availability in the environment in which strains have been isolated. This interesting hypothesis was further confirmed by comparing metagenomic datasets from environments exhibiting different local P concentrations (Martiny et al. 2006, 2009a; Rusch et al. 2007). P-related genes were indeed much more numerous in regions displaying less than 0.1 μM phosphate (Sargasso and Caribbean Seas) than in more P-rich areas (E Pacific and Indian Oceans, Martiny et al. 2009a).

P-related genes are not distributed randomly within *Prochlorococcus* genomes. Most of them occur in two regions with highly variable gene content, so-called genomic islands (Coleman et al. 2006, Kettler et al. 2007, Martiny et al. 2006). One island called ISL5 was first revealed by whole genome comparison of HL strains MED4 and MIT9312 (Coleman et al. 2006). It was found to contain several uncharacterized genes that are strongly upregulated during P starvation, suggesting that they must have an important role in P metabolism. Other genes related to nutrition processes can be found in genomic islands, such as genes involved in transport or assimilation of nitrate, nitrite, amino acids, cyanate, or metal traces (Coleman et al. 2006, Kettler et al. 2007, Martiny et al. 2009b). Islands also contain genes involved in a variety of other processes, such as light stress response (e.g., most *hli* genes are found in genomic islands) or cell wall synthesis (e.g., glycosyltransferases or glycoside hydrolases). Indeed, it has been suggested that high variability of cell wall composition may constitute an efficient strategy against phages or other predators that often require specific recognition motifs or attachment sites (Kettler et al. 2007, Palenik et al. 2006). It is, however, worth noting that the majority of island genes, which are frequently unique or shared by few strains, have no known function.

While Kettler et al. (2007) found that newly acquired genes are preferentially located in islands in all streamlined *Prochlorococcus* genomes, they suggested that members of the LLIV clade may have adopted another strategy for integrating laterally transferred genes into their genomes, since their specific genes seemingly did not cluster into discernible islands. It must be noted, however, that Dufresne et al. (2008) managed to predict the occurrence of at least 11 islands in the genome of the LLIV strain MIT9313 by applying a bioinformatics method, based mainly on the deviation in tetranucleotide frequency, that proved successful in predicting islands in marine *Synechococcus* genomes. Thus, use of the sequence alignment tool BLAST to directly compare the LLIV genomes with field metagenomic data—when such data become available for the niche occupied by this ecotype—will likely be necessary to solve this controversy. Indeed, this approach proved very efficient to check for the validity of predicted island regions in both HL *Prochlorococcus* and marine *Synechococcus* genomes since there is generally much less recruitment of field gene fragments at the level of these islands than in conserved parts of the genomes (Dufresne et al. 2008, Kettler et al. 2007, Rusch et al. 2007).

Lastly, many genes contained in islands have been shown to be differentially expressed under a variety of conditions, including nutrient and light stresses, as well as during phage infection (Coleman et al. 2006, Lindell et al. 2007, Martiny et al. 2006, Steglich et al. 2006). Altogether, this implies that such genes must play a key role in the response to stresses that may appear in various combinations depending on the local environment. However, the precise dynamics of acquisition/loss of these specific genes and the mechanisms of specific inclusion into these particular

regions of the genome is not yet known and constitutes one of the most exciting challenges of modern microbial ecology.

Mechanisms Involved in Genome Streamlining

What factors were responsible for initiating the genome streamlining process that has affected most *Prochlorococcus* lineages is still a controversial question. It has been suggested that small cell size (and consequently small genome size) may be advantageous for life at depth, given that the cell-surface-to-volume ratio increases as cell size decreases, optimizing absorption of incident photons (Dufresne et al. 2005). Light absorption efficiency is also specifically improved for such tiny cells as *Prochlorococcus* due to low probability of photon scattering (Morel et al. 1993). *Prochlorococcus* genotypes with smaller cell and genome size could therefore have been favored by natural selection. Nevertheless, with the evidence that one of the most (if not the most) abundant *Prochlorococcus* populations at the bottom of the euphotic zone in extant oceans are representatives of the LLIV ecotype/clade (Johnson et al. 2006, Zinser et al. 2006), which have maintained a cell and genome size similar to that of *Synechococcus*, this hypothesis apparently does not hold.

Another possible explanation for the initiation of genome reduction is a relaxation of the selection on a number of genes that became dispensable in the low-light niche. In particular, absence of UV radiation in this niche may have allowed the degeneration and then elimination of some DNA repair genes. All *Prochlorococcus* strains with a streamlined genome also have an unusually high AT content (**Table 1**), and it has been hypothesized that this was related to the loss of one or several gene(s) involved in the repair of G:C to A:T transversions (Dufresne et al. 2005, Rocap et al. 2003). There are apparently no (annotated) genes belonging to the DNA replication and repair category among the 70 aforementioned *Synechococcus*-specific core genes, confirming that loss of genes belonging to this category occurred only after differentiation of the *Prochlorococcus* LLIV branch from its common ancestor with other *Prochlorococcus* lineages. At least seven genes involved in DNA repair are present in the two LLIV strains, but not in any streamlined genome (**Table 3**).

Six of these genes also belong to the *Synechococcus* core genome and the seventh, a NUDIX hydrolase, is missing in only one out of 11 *Synechococcus*. One of them, *ybaZ*, is related to the *ada* and *ogt* genes, with which it was confounded in previous genome analyses (Dufresne et al. 2005, Marais et al. 2008). Ada and Ogt methyltransferases, which are known to prevent G:C to A:T tranversions in *E. coli* by reverting potentially mutagenic 6-*O*-methylguanine adducts into guanine (Rye et al. 2008), are in fact seemingly missing in all marine picocyanobacteria. The alkyltransferase-like protein YbaZ does not have any alkyltransferase activity per se in *E. coli* but was suggested to tag nucleotides affected by large 6-*O*-alkylguanine adducts (such as ethyl- or propylguanine) before their excision by the nucleotide excision repair (NER) pathway (Mazon et al. 2009). Absence of *ybaZ* in all streamlined *Prochlorococcus* may thus have caused a decrease in the capacity of the NER pathway to eliminate these large, potentially mutagenic adducts, and possibly 6-*O*-methylguanine as well, given the absence of Ada and Ogt homologs.

Another phenomenon, which occurred concomitantly with genome reduction and AT-enrichment, was a global acceleration of the rate of evolution of protein-coding gene sequences, with the HL strain MED4 showing a twofold higher rate than the LLII strain SS120, itself having evolved twice as fast as the LLIV strain MIT9313 (Dufresne et al. 2005). This phenomenon, revealed by comparing the set of protein-coding genes shared by these different strains, affected most genes independently of their metabolic function. It has been suggested, though, that some highly conserved photosynthetic genes, such as the photosystem core proteins, may have escaped this general trend because natural selection acted to maintain the numerous and vital interactions

occurring between and within the supramolecular complexes constituting the photosynthetic apparatus (Shi et al. 2005). Similarly, the ribosomal RNA operon was not affected by the base composition drift (Dufresne et al. 2005), as can be clearly visualized on global plots representing GC content, where this region appears as a high GC% island in an ocean of high AT% nucleotides (see **Supplemental Figure 1**; follow the **Supplemental Material** link from the Annual Reviews home page at **http://www.annualreviews.org**).

Supplemental Material

Marais et al. (2008) proposed an interesting evolutionary scenario where loss of DNA repair genes in some genotypes may have caused them to become mutator strains, i.e., strains that have a much higher mutation rate than the rest of the population. Mutator strains may indeed acquire selective advantages that could allow them to conquer a new niche and/or overgrow other members of the population after a change in prevailing environmental conditions (Taddei et al. 1997). Although *ybaZ* is not yet known as a mutator (i.e., a gene which once inactivated provokes a sudden increase of the mutation rate of the cell), *ada* and *ogt* have been shown to have this potential in *E. coli* (Miller 1998) and we assume that *ybaZ* (and possibly some of the other missing DNA repair genes listed in **Table 3**) may have it as well. Marais et al. (2008) designed a simple mathematical model showing that for a very large population (as is the case for *Prochlorococcus* in the field), even a modest increase in mutation rate is sufficient to obtain a significant genome reduction of the total population over the long-term. For instance, they predicted that a 15% increase in mutation rate would lead to a ~33% decrease in genome size, equivalent to that assumed for the evolution of streamlined genomes (see above).

Marais et al. (2008) also suggested that modern *Prochlorococcus* populations with streamlined genomes may still possess high mutation rates. However, if this hypothesis were true, one would expect to observe, as in obligate intracellular bacteria (Frank et al. 2002), the accumulation of pseudogenes, i.e., genes with deleterious mutations and in the process of being eliminated. We know, from having manually checked the annotation of hundreds of genes of marine picocyanobacteria, that streamlined genomes have in fact extremely few pseudogenes (compared to, e.g., the LLIV strain MIT9313) and, in particular, very few genes interrupted by stop codons, which are predicted to be very frequent in a context where G:C to A:T transversions are frequent (Oller et al. 1993). Thus, if mutators ever occurred during *Prochlorococcus* evolution, we rather suggest that there have been several independent episodes of occurrence of mutator allele(s), followed each time by restoration of a lower mutation rate once the new population was adapted to its new environment, as predicted by the classical mutator theory (Denamur et al. 2000).

The first mutator event, potentially involving the inactivation then elimination of *ybaZ* and a number of other DNA repair genes, led to the occurrence of the LLII/III lineages. The triggering factors involved here remain unclear since the preferred ecological niches of these strains in the field—possibly located below that of the LLIV clade (see, e.g., Garczarek et al. 2007, Steglich et al. 2003)—are not as yet precisely established. A second mutator event may have led to the differentiation of the eNATL ecotype (LLI clade), allowing it to colonize a niche located higher in the water column (**Figure 3**). It is worth noting, though, that we have identified only one less (vertically inherited) DNA repair gene in LLI than in LLII/III strains (**Table 3**), a helicase belonging to the superfamily II with unknown mutator potential. The last episode, which could have been induced by the loss of several more DNA repair genes, may have occurred just before the differentiation of the HL branch. Among them, *mutY* is known to have mild mutator effects if inactivated in *Bacillus anthracis* (Zeibell et al. 2007).

Surprisingly, the last mutator episode seems to have had a notable effect on the GC content of HL strains but not on their genome size (**Table 1**). This observation supports the assumption that the genomes of modern *Prochlorococcus* HLI/II and LLII/III cells have reached a lower limit for a free-living photosynthetic organism (Dufresne et al. 2003). For the eNATL ecotype, which

is confronted with a more variable environment than its HLI/II and LLII/III/IV counterparts, it is possible that selection for maintaining more vertically inherited genes useful for life in such changing conditions is higher than in other streamlined genomes.

Besides acquisition of novel genes by lateral transfer, it has been recently suggested that *Prochlorococcus* may also be able to cooperate with microorganisms living in its immediate environment. Indeed, Morris et al. (2008) have shown that co-occurring bacteria may help *Prochlorococcus* cells survive oxidative stress by removing reactive oxygen species from the medium, a process involving activity of catalases or catalase/peroxidases. These discoveries suggest that the evolution of *Prochlorococcus* genomes cannot be considered apart from the surrounding metagenomes (i.e., the pool of genes accessible by lateral transfer) and the community of interacting organisms.

SUMMARY POINTS

1. With the increase of oligotrophic areas, the preferred niche of *Prochlorococcus* is globally increasing, thereby potentially enhancing its relative contribution to global ocean productivity. However, in the S Pacific gyre, nutrient limitation in the upper layer is seemingly too drastic for normal growth of *Prochlorococcus* populations.
2. Differentiation of the *Prochlorococcus* genus has involved more gene losses than creation. However, the actual genome streamlining process began only after the differentiation of the LLIV branch from the common ancestor of all other *Prochlorococcus*.
3. The dualistic concept of the HL and LL ecotypes that prevailed for many years in the literature on *Prochlorococcus* proved oversimplistic. Indeed, the eNATL ecotype (or LLI clade) definitely represents a third entity occupying an intermediate (more variable) niche compared to its HL/true LL counterparts. This specificity clearly translates into its genome composition, which gathers characteristics of the two other ecotypes.
4. The apparent dominance of the large genome-possessing LLIV ecotype at the bottom of the euphotic zone suggests that, even though genome streamlining began in this low-light/high-nutrient niche, a small cell and genome size is not an absolute requirement (nor seemingly the best strategy) for life in such a niche. In contrast, the remarkable abundance of HL populations in the upper, nutrient-poor areas of the world ocean shows that genome streamlining proved a very efficient strategy for colonizing this particular niche, in which *Prochlorococcus* always overgrows the much more ancient genus *Synechococcus.* Indeed, genome minimalism can allow *Prochlorococcus* (or other oceanic microorganisms, such as *Candidatus* Pelagibacter ubique) to propagate their genetic information with a minimal consumption of energy and limiting nutrients, such as N or P.
5. Even the most streamlined *Prochlorococcus* genomes have kept some plasticity and may adapt to changing conditions by recruiting genes (via phages or natural transformation) from other members of the community. These transfers are preferentially directed toward specific regions called genomic islands. Additionally, cooperative interactions with co-occurring organisms have been observed.
6. Loss of DNA repair genes likely had a critical effect on *Prochlorococcus* genome evolution. It has been proposed that disappearance of some of these genes created mutator strains, exhibiting an increased evolutionary rate. We suggest here that this event may have occurred several times during *Prochlorococcus* evolution.

FUTURE ISSUES

Although 13 genomes from cultured *Prochlorococcus* strains are currently available (**Table 1**) and have already brought remarkable insights about evolutionary processes in this genus and differentiation mechanisms between major ecotypes, a number of burning questions remain about this crucial component of the marine community, including the following:

1. Are LLII and LLIII ecotypes really minor components of the *Prochlorococcus* community or have they merely been underestimated?
2. Are there many more LL *Prochlorococcus* lineages in the field than currently known, as suggested by recent discoveries of novel uncultured lineages using metagenomics approaches?
3. How can we reduce potential errors in generalizing to field *Prochlorococcus* populations the results from comparative genome analyses based solely on cultured strains, as recently evidenced, e.g., for nitrate assimilation?
4. What are the major molecular processes allowing lateral gene transfer between *Prochlorococcus* lineages as well as with other members of the microbial community and at which rate do these exchanges occur?
5. Will *Prochlorococcus* ultimately adapt to the ultraoligotrophic conditions found in the upper layer of the S Pacific ocean (e.g., after another mutator event)?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by the Biosope project of the LEFE-CYBER program. We dedicate this paper to Alexis Dufresne, who developed the precious bioinformatics tools that were used for the comparative genome analyses mentioned in this review. Matthieu Barret and Anne Peyrat are also warmly thanked for their help with acquiring unpublished data shown in **Figure 2*c*** along with David J. Scanlan for critically reading the manuscript. We thank Zackary I. Johnson and Sallie W. Chisholm for kindly permitting us to use data from the AMT13 cruise and Joséphine Ras, Hervé Claustre, Carolina Grob, and Osvaldo Ulloa for sharing data from the Biosope cruise and (or) providing a manuscript prior to publication. Jeffrey Polovina is acknowledged for allowing us to reproduce a high-resolution image of his model of low-chlorophyll waters expansion (**Figure 2*a***).

LITERATURE CITED

Ahlgren NA, Rocap G, Chisholm SW. 2006. Measurement of *Prochlorococcus* ecotypes using real-time polymerase chain reaction reveals different abundances of genotypes with similar light physiologies. *Environ. Microbiol.* 8:441–54

Badger MR, Price GD. 2003. CO_2 concentrating mechanisms in cyanobacteria: Molecular components, their diversity and evolution. *J. Exp. Bot.* 54:609–22

Behrenfeld MJ, Kolber ZS. 1999. Widespread iron limitation of phytoplankton in the South Pacific Ocean. *Science* 283:840–43

Bhaya D, Dufresne A, Vaulot D, Grossman A. 2002. Analysis of the *hli* gene family in marine and freshwater cyanobacteria. *FEMS Microbiol. Lett.* 215:209–19

Bibby TS, Mary I, Nield J, Partensky F, Barber J. 2003. Low-light-adapted *Prochlorococcus* species possess specific antennae for each photosystem. *Nature* 424:1051–54

Bibby TS, Nield J, Barber J. 2001a. Iron deficiency induces the formation of an antenna ring around trimeric photosystem I in cyanobacteria. *Nature* 412:743–45

Bibby TS, Nield J, Partensky F, Barber J. 2001b. Oxyphotobacteria—Antenna ring around photosystem I. *Nature* 413:590

Boekema EJ, Hifney A, Yakushevska AE, Piotrowski M, Keegstra W, et al. 2001. A giant chlorophyll-protein complex induced by iron deficiency in cyanobacteria. *Nature* 412:745–48

Bonnet S, Guieu C, Bruyant F, Prasil O, Van Wambeke F, et al. 2007. Nutrient limitation of primary productivity in the Southeast Pacific (BIOSOPE cruise). *Biogeosciences* 5:215–25

Bouman HA, Ulloa O, Scanlan DJ, Zwirglmaier K, Li WK, et al. 2006. Oceanographic basis of the global surface distribution of *Prochlorococcus* ecotypes. *Science* 312:918–21

Campbell L, Liu HB, Nolla HA, Vaulot D. 1997. Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991–1994 ENSO event. *Deep-Sea Res. I* 44:167–92

Campbell L, Vaulot D. 1993. Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii (station ALOHA). *Deep-Sea Res.* 40:2043–60

Casey JR, Lomas MW, Mandecki J, Walker DE. 2007. *Prochlorococcus* contributes to new production in the Sargasso Sea deep chlorophyll maximum. *Geophys. Res. Lett.* 34:L10604, doi:10.1029/2006GL028725

Chen M, Bibby TS. 2005. Photosynthetic apparatus of antenna-reaction centres supercomplexes in oxyphotobacteria: Insight through significance of Pcb/IsiA proteins. *Photosynth. Res.* 86:165–73

Chisholm SW, Frankel SL, Goericke R, Olson RJ, Palenik B, et al. 1992. *Prochlorococcus marinus* nov. gen. nov. sp.: An oxyphototrophic marine prokaryote containing divinyl chlorophyll *a* and *b*. *Arch. Microbiol.* 157:297–300

Claustre H, Sciandra A, Vaulot D. 2008. Introduction to the special section bio-optical and biogeochemical conditions in the South East Pacific in late 2004: The BIOSOPE program. *Biogeosciences* 5:679–91

Coleman ML, Chisholm SW. 2007. Code and context: *Prochlorococcus* as a model for cross-scale biology. *Trends Microbiol.* 15:399–407

Coleman ML, Sullivan MB, Martiny AC, Steglich C, Barry K, et al. 2006. Genomic islands and the ecology and evolution of *Prochlorococcus*. *Science* 311:1768–70

Denamur E, Lecointre G, Darlu P, Tenaillon O, Acquaviva C, et al. 2000. Evolutionary implications of the frequent horizontal transfer of mismatch repair genes. *Cell* 103:711–21

DiTullio GR, Geesey ME, Jones DR, Daly KL, Campbell L, Smith WO. 2003. Phytoplankton assemblage structure and primary productivity along 170°W in the South Pacific Ocean. *Mar. Ecol. Prog. Ser.* 255:55–80

Dufresne A, Garczarek L, Partensky F. 2005. Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome Biol.* 6:R14

Dufresne A, Ostrowski M, Scanlan DJ, Garczarek L, Mazard S, et al. 2008. Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol.* 9:R90

Dufresne A, Salanoubat M, Partensky F, Artiguenave F, Axmann IM, et al. 2003. Genome sequence of the cyanobacterium *Prochlorococcus marinus* SS120, a nearly minimal oxyphototrophic genome. *Proc. Natl. Acad. Sci. USA* 100:10020–25

Fong AA, Karl DM, Lukas R, Letelier RM, Zehr JP, Church MJ. 2008. Nitrogen fixation in an anticyclonic eddy in the oligotrophic North Pacific Ocean. *ISME J.* 2:663–76

Frank AC, Amiri H, Andersson SG. 2002. Genome deterioration: Loss of repeated sequences and accumulation of junk DNA. *Genetica* 115:1–12

Garcia-Fernandez JM, de Marsac NT, Diez J. 2004. Streamlined regulation and gene loss as adaptive mechanisms in *Prochlorococcus* for optimized nitrogen utilization in oligotrophic environments. *Microbiol. Mol. Biol. Rev.* 68:630–38

Garcia-Fernandez JM, Diez J. 2004. Adaptive mechanisms of nitrogen and carbon assimilatory pathways in the marine cyanobacteria *Prochlorococcus*. *Res. Microbiol.* 155:795–802

Garcia-Pichel F, Belnap J, Neuer S, Schanz F. 2003. Estimates of cyanobacterial biomass and its distribution. *Algol. Stud.* 109:213–28

Garcia-Pichel F, Lopez-Cortes A, Nubel U. 2001. Phylogenetic and morphological diversity of cyanobacteria in soil desert crusts from the Colorado Plateau. *Appl. Environ. Microbiol.* 67:1902–10

Garczarek L, Dufresne A, Blot N, Cockshutt AM, Peyrat A, et al. 2008. Function and evolution of the *psbA* gene family in marine *Synechococcus*: *Synechococcus* sp. WH7803 as a case study. *ISME J.* 2:937–53

Garczarek L, Dufresne A, Rousvoal S, West NJ, Mazard S, et al. 2007. High vertical and low horizontal diversity of *Prochlorococcus* ecotypes in the Mediterranean Sea in summer. *FEMS Microbiol. Ecol.* 60:189–206

Garczarek L, van der Staay GWM, Hess WR, Le Gall F, Partensky F. 2001. Expression and phylogeny of the multiple antenna genes of the low-light-adapted strain *Prochlorococcus marinus* SS120 (Oxyphotobacteria). *Plant Mol. Biol.* 46:683–93

Gieskes WWC, Kraay GW. 1983. Unknown chlorophyll a derivatives in the North Sea and the tropical Atlantic Ocean revealed by HPLC analysis. *Limnol. Oceanogr.* 28:757–66

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, et al. 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–45

Goericke R, Repeta DJ. 1992. The pigments of *Prochlorococcus marinus*: The presence of divinyl chlorophyll *a* and *b* in a marine prochlorophyte. *Limnol. Oceanogr.* 37:425–33

Goericke R, Welschmeyer NA. 1993. The marine prochlorophyte *Prochlorococcus* contributes significantly to phytoplankton biomass and primary production in the Sargasso Sea. *Deep-Sea Res.* 40:2283–94

Goosen N, Moolenaar GF. 2008. Repair of UV damage in bacteria. *DNA Repair* 7:353–79

Grob C, Ulloa O, Claustre H, Huot Y, Alarcon G, Marie D. 2007. Contribution of picoplankton to the total particulate organic carbon concentration in the eastern South Pacific. *Biogeosciences* 4:837–52

Hess WR, Partensky F, van der Staay GWM, Garcia-Fernandez JM, Boerner T, Vaulot D. 1996. Coexistence of phycoerythrin and a chlorophyll a/b antenna in a marine prokaryote. *Proc. Natl. Acad. Sci. USA* 93:11126–30

Hess WR, Rocap G, Ting CS, Larimer F, Stilwagen S, et al. 2001. The photosynthetic apparatus of *Prochlorococcus*: Insights through comparative genomics. *Photosynth. Res.* 70:53–71

Hess WR, Steglich C, Lichtlé C, Partensky F. 1999. Phycoerythrins of the oxyphotobacterium *Prochlorococcus marinus* are associated to the thylakoid membrane and are encoded by a single large gene cluster. *Plant Mol. Biol.* 40:507–21

Holtzendorff J, Marie D, Post AF, Partensky F, Rivlin A, Hess WR. 2002. Synchronized expression of *ftsZ* in natural *Prochlorococcus* populations of the Red Sea. *Environ. Microbiol.* 4:644–53

Holtzendorff J, Partensky F, Mella D, Lennon JF, Hess WR, Garczarek L. 2008. Genome streamlining results in loss of robustness of the circadian clock in the marine cyanobacterium *Prochlorococcus marinus* PCC 9511. *J. Biol. Rhythms* 23:187–99

Johnson ZI, Zinser ER, Coe A, McNulty NP, Woodward EM, Chisholm SW. 2006. Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* 311:1737–40

Kettler G, Martiny AC, Huang K, Zucker J, Coleman ML, et al. 2007. Patterns and implications of gene gain and loss in the evolution of *Prochlorococcus*. *PLoS Genet.* 3:e231

LaRoche J, van der Staay GW, Partensky F, Ducret A, Aebersold R, et al. 1996. Independent evolution of the prochlorophyte and green plant chlorophyll *a/b* light-harvesting proteins. *Proc. Natl. Acad. Sci. USA* 93:15244–48

Lee JY, Chang C, Song HK, Moon J, Yang JK, et al. 2000. Crystal structure of NAD(+)-dependent DNA ligase: Modular architecture and functional implications. *EMBO J.* 19:1119–29

Li GM. 2008. Mechanisms and functions of DNA mismatch repair. *Cell Res.* 18:85–98

Lindell D, Jaffe JD, Coleman ML, Futschik ME, Axmann IM, et al. 2007. Genome-wide expression dynamics of a marine virus and host reveal features of coevolution. *Nature* 449:83–86

Lindell D, Sullivan MB, Johnson ZI, Tolonen AC, Rohwer F, Chisholm SW. 2004. Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc. Natl. Acad. Sci. USA* 101:11013–18

Marais GAB, Calteau A, Tenaillon O. 2008. Mutation rate and genome reduction in endosymbiotic and free-living bacteria. *Genetica* 134:205–10

Martiny AC, Coleman ML, Chisholm SW. 2006. Phosphate acquisition genes in *Prochlorococcus* ecotypes: Evidence for genome-wide adaptation. *Proc. Natl. Acad. Sci. USA* 103:12552–57

Martiny AC, Huang Y, Li W. 2009a. Occurrence of phosphate acquisition genes in *Prochlorococcus* cells from different ocean regions. *Environ. Microbiol.* 11:1340–47
Martiny AC, Kathuria S, Berube PM. 2009b. Widespread metabolic potential for nitrite and nitrate assimilation among *Prochlorococcus* ecotypes. *Proc. Natl. Acad. Sci. USA* 106:10787–92
Martiny AC, Tai AP, Veneziano D, Primeau F, Chisholm SW. 2009c. Taxonomic resolution, ecotypes and the biogeography of *Prochlorococcus*. *Environ. Microbiol.* 11:823–32
Mary I, Garczarek L, Tarran GA, Kolowrat C, Terry MJ, et al. 2008. Diel rhythmicity in amino acid uptake by *Prochlorococcus*. *Environ. Microbiol.* 10:2124–31
Mazon G, Philippin G, Cadet J, Gasparutto D, Fuchs RP. 2009. The alkyltransferase-like *ybaZ* gene product enhances nucleotide excision repair of O6-alkylguanine adducts in *E. coli*. *DNA Repair* 8:697–703
Miller JH. 1998. Mutators in *Escherichia coli*. *Mutat. Res.* 409:99–106
Moore LR, Coe A, Zinser ER, Saito MA, Sullivan MB, et al. 2007. Culturing the marine cyanobacterium *Prochlorococcus*. *Limnol. Oceanogr. Methods* 5:353–62
Moore LR, Goericke R, Chisholm SW. 1995. Comparative physiology of *Synechococcus* and *Prochlorococcus*: Influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Mar. Ecol. Prog. Ser.* 116:259–75
Moore LR, Post AF, Rocap G, Chisholm SW. 2002. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Limnol. Oceanogr.* 47:989–96
Moore LR, Rocap G, Chisholm SW. 1998. Physiology and molecular phylogeny of coexisting *Prochlorococcus* ecotypes. *Nature* 393:464–67
Morel A, Ahn Y-H, Partensky F, Vaulot D, Claustre H. 1993. *Prochlorococcus* and *Synechococcus*: A comparative study of their optical properties in relation to their size and pigmentation. *J. Mar. Res.* 51:617–49
Morel A, Gentili B, Claustre H, Babin M, Bricaud A, et al. 2007. Optical properties of the "clearest" natural waters. *Limnol. Oceanogr.* 52:217–29
Morris JJ, Kirkegaard R, Szul MJ, Johnson ZI, Zinser ER. 2008. Facilitation of robust growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria. *Appl. Environ. Microbiol.* 74:4530–34
Moutin T, Thingstad TF, Van Wambeke F, Marie D, Slawyk G, et al. 2002. Does competition for nanomolar phosphate supply explain the predominance of the cyanobacterium *Synechococcus*? *Limnol. Oceanogr.* 47:1562–67
Nagata N, Tanaka R, Satoh S, Tanaka A. 2005. Identification of a vinyl reductase gene for chlorophyll synthesis in *Arabidopsis thaliana* and implications for the evolution of *Prochlorococcus* species. *Plant Cell* 17:233–40
Neuer S, Torres Padron ME, Gelado Caballero MD, Rueda MJ, Hernandez Brito J, et al. 2004. Dust deposition pulses to the eastern subtropical North Atlantic gyre: Does ocean's biogeochemistry respond? *Glob. Biogeochem. Cycles* 18:GB4020, doi: 10.1029/2004GB002228
Oller AR, Fijalkowska IJ, Schaaper RM. 1993. The *Escherichia coli* galK2 papillation assay: Its specificity and application to seven newly isolated mutator strains. *Mutat. Res.* 292:175–85
Palenik B. 1994. Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. *Appl. Environ. Microbiol.* 60:3212–19
Palenik B, Ren QH, Dupont CL, Myers GS, Heidelberg JF, et al. 2006. Genome sequence of *Synechococcus* CC9311: Insights into adaptation to a coastal environment. *Proc. Natl. Acad. Sci. USA* 103:13555–59
Partensky F, Blanchot J, Lantoine F, Neveux J, Marie D. 1996. Vertical structure of picophytoplankton at different trophic sites of the tropical northeastern Atlantic Ocean. *Deep-Sea Res. I* 43:1191–213
Partensky F, Blanchot J, Vaulot D. 1999a. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: A review. In *Marine Cyanobacteria*, ed. L Charpy, A Larkum, pp. 457–75. Monaco: Musée Océanographique
Partensky F, Garczarek L. 2003. The photosynthetic apparatus of chlorophyll *b*- and *d*-containing Oxychlorobacteria. In *Photosynthesis in Algae*, ed. AWD Larkum, SE Douglas, JA Raven, pp. 29–62. Dordrecht, the Netherlands: Kluwer Academic
Partensky F, Hess WR, Vaulot D. 1999b. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* 63:106–27
Partensky F, Hoepffner N, Li WKW, Ulloa O, Vaulot D. 1993. Photoacclimation of *Prochlorococcus* sp. (Prochlorophyta) strains isolated from the North Atlantic and the Mediterranean Sea. *Plant Physiol.* 101:285–96

Pascal JM. 2008. DNA and RNA ligases: Structural variations and shared mechanisms. *Curr. Opin. Struct. Biol.* 18:96–105

Penno S, Lindell D, Post AF. 2006. Diversity of *Synechococcus* and *Prochlorococcus* populations determined from DNA sequences of the N-regulatory gene *ntcA*. *Environ. Microbiol.* 8:1200–11

Polovina JJ, Howell EA, Abecassis M. 2008. Ocean's least productive waters are expanding. *Geophys. Res. Lett.* 35, LO3618, doi:10.1029/2007GL031745

Ras J, Claustre H, Uitz J. 2008. Spatial variability of phytoplankton pigment distributions in the subtropical South Pacific Ocean: Comparison between in situ and predicted data. *Biogeosciences* 5:353–69

Rocap G, Distel DL, Waterbury JB, Chisholm SW. 2002. Resolution of *Prochlorococcus* and *Synechococcus* ecotypes by using 16S–23S ribosomal DNA internal transcribed spacer sequences. *Appl. Environ. Microbiol.* 68:1180–91

Rocap G, Larimer FW, Lamerdin J, Malfatti S, Chain P, et al. 2003. Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* 424:1042–47

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, et al. 2007. The Sorcerer II Global Ocean Sampling expedition: Northwest Atlantic through Eastern Tropical Pacific. *Plos Biol.* 5:398–431

Rye PT, Delaney JC, Netirojjanakul C, Sun DX, Liu JZ, Essigmann JM. 2008. Mismatch repair proteins collaborate with methyltransferases in the repair of O-6-methylguanine. *DNA Repair* 7:170–76

Satoh S, Tanaka A. 2006. Identification of chlorophyllide a oxygenase in the *Prochlorococcus* genome by a comparative genomic approach. *Plant Cell Physiol.* 47:1622–29

Scanlan DJ, Ostrowski M, Mazard S, Dufresne A, Garczarek L, et al. 2009. Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.* 73:249–99

Scanlan DJ, West NJ. 2002. Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microbiol. Ecol.* 40:1–12

Shi T, Bibby TS, Jiang L, Irwin AJ, Falkowski PG. 2005. Protein interactions limit the rate of evolution of photosynthetic genes in cyanobacteria. *Mol. Biol. Evol.* 22:2179–89

Six C, Joubin L, Partensky F, Holtzendorff J, Garczarek L. 2007a. UV-induced phycobilisome dismantling in the marine picocyanobacterium *Synechococcus* sp. WH8102. *Photosynth. Res.* 92:75–86

Six C, Thomas JC, Garczarek L, Ostrowski M, Dufresne A, et al. 2007b. Diversity and evolution of phycobilisomes in marine *Synechococcus* spp.: A comparative genomics study. *Genome Biol.* 8:R259

Steglich C, Futschik M, Rector T, Steen R, Chisholm SW. 2006. Genome-wide analysis of light sensing in *Prochlorococcus*. *J. Bacteriol.* 188:7796–806

Steglich C, Post AF, Hess WR. 2003. Analysis of natural populations of *Prochlorococcus* spp. in the northern Red Sea using phycoerythrin gene sequences. *Environ. Microbiol.* 5:581–90

Stickforth P, Steiger S, Hess WR, Sandmann G. 2003. A novel type of lycopene epsilon-cyclase in the marine cyanobacterium *Prochlorococcus marinus* MED4. *Arch. Microbiol.* 179:409–15

Taddei F, Radman M, Maynard-Smith J, Toupance B, Gouyon PH, Godelle B. 1997. Role of mutator alleles in adaptive evolution. *Nature* 387:700–2

Ting CS, Hsieh C, Sundararaman S, Mannella C, Marko M. 2007. Cryo-electron tomography reveals the comparative three-dimensional architecture of *Prochlorococcus*, a globally important marine cyanobacterium. *J. Bacteriol.* 189:4485–93

Ting CS, Rocap G, King J, Chisholm SW. 2001. Phycobiliprotein genes of the marine photosynthetic prokaryote *Prochlorococcus*: Evidence for rapid evolution of genetic heterogeneity. *Microbiology* 147:3171–82

Ting CS, Rocap G, King J, Chisholm SW. 2002. Cyanobacterial photosynthesis in the oceans: The origins and significance of divergent light-harvesting strategies. *Trends Microbiol.* 10:134–42

Urbach E, Scanlan DJ, Distel DL, Waterbury JB, Chisholm SW. 1998. Rapid diversification of marine picophytoplankton with dissimilar light-harvesting structures inferred from sequences of *Prochlorococcus* and *Synechococcus* (Cyanobacteria). *J. Mol. Evol.* 46:188–201

Van Wambeke F, Obernosterer I, Moutin T, Duhamel S, Ulloa O, Claustre H. 2008. Heterotrophic bacterial production in the eastern South Pacific: Longitudinal trends and coupling with primary production. *Biogeosciences* 5:157–69

Veldhuis MJW, Kraay GW. 2004. Phytoplankton in the subtropical Atlantic Ocean: Towards a better assessment of biomass and composition. *Deep Sea Res. I* 51:507–30

Venter CJ, Remington K, Heidelberg JF, Halpern AL, Rusch D, et al. 2004. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* 307:66–74

West NJ, Scanlan DJ. 1999. Niche-partitioning of *Prochlorococcus* populations in a stratified water column in the eastern North Atlantic Ocean. *Appl. Environ. Microbiol.* 65:2585–91

Wilson A, Ajlani G, Verbavatz JM, Vass I, Kerfeld CA, Kirilovsky D. 2006. A soluble carotenoid protein involved in phycobilisome-related energy dissipation in cyanobacteria. *Plant Cell* 18:992–1007

Zehr JP, Bench SR, Carter BJ, Hewson I, Niazi F, et al. 2008. Globally distributed uncultivated oceanic N2-fixing cyanobacteria lack oxygenic photosystem II. *Science* 322:1110–12

Zeibell K, Aguila S, Yan Shi V, Chan A, Yang H, Miller JH. 2007. Mutagenesis and repair in *Bacillus anthracis*: The effect of mutators. *J. Bacteriol.* 189:2331–38

Zinser ER, Coe A, Johnson ZI, Martiny AC, Fuller NJ, et al. 2006. *Prochlorococcus* ecotype abundances in the North Atlantic Ocean as revealed by an improved quantitative PCR method. *Appl. Environ. Microbiol.* 72:723–32

Zwirglmaier K, Heywood JL, Chamberlain K, Woodward EMS, Zubkov MV, Scanlan DJ. 2007. Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ. Microbiol.* 9:1278–90

Zwirglmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, et al. 2008. Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ. Microbiol.* 10:147–61

Annual Review of
Marine Science

Volume 2, 2010

Contents

Errata

An online log of corrections to *Annual Review of Marine Science* articles may be found at http://marine.annualreviews.org/errata.shtml