Prochlorococcus: Advantages and Limits of Minimalism

Frédéric Partensky¹,² and Laurence Garczarek¹,²

¹UPMC-Université Paris 06, Station Biologique, 29682 Roscoff cedex, France
²CNRS, UMR 7144, Oceanic Plankton Group, 29682 Roscoff cedex, France; email: partensky@sb-roscoff.fr

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 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 (García-Pichel et al. 2003, García-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 N₂) *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. García-Fernandez et al. (2004) and García-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 pico-cyanobacteria 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.

**PRO Chlorococcus 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 a₂) in subtropical oligotrophic areas (DiTullio et al. 2003, Gieskes &
Table 1  Main characteristics of the different Prochlorococcus genomes sequenced to date

<table>
<thead>
<tr>
<th>Strain</th>
<th>CCMP(^a) no.</th>
<th>RCC(^b) no.</th>
<th>Clade name</th>
<th>Genome size (Mb)</th>
<th>No. of protein-coding genes</th>
<th>GC%</th>
<th>Region</th>
<th>Genbank accession no.</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MED4</td>
<td>1986</td>
<td>151</td>
<td>HLI</td>
<td>1.658</td>
<td>1716(^c)/1763(^d)/1929(^e)</td>
<td>30.8</td>
<td>Mediterranean Sea</td>
<td>BX548174</td>
<td>(Kettler et al. 2007, Rocap et al. 2003)</td>
</tr>
<tr>
<td>MIT9515</td>
<td>3148</td>
<td>—</td>
<td>HLI</td>
<td>1.704</td>
<td>1908</td>
<td>30.8</td>
<td>Eq. Pacific</td>
<td>CP000552</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>MIT9301</td>
<td>2971</td>
<td>—</td>
<td>HLII</td>
<td>1.643</td>
<td>1907</td>
<td>31.4</td>
<td>Sargasso Sea</td>
<td>CP000576</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>AS9601</td>
<td>2970</td>
<td>—</td>
<td>HLII</td>
<td>1.670</td>
<td>1926</td>
<td>31.3</td>
<td>Arabian Sea</td>
<td>CP000551</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>MIT9215</td>
<td>2919</td>
<td>—</td>
<td>HLII</td>
<td>1.738</td>
<td>1989</td>
<td>31.1</td>
<td>Eq. Pacific</td>
<td>CP000825</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>MIT9312</td>
<td>2777</td>
<td>—</td>
<td>HLII</td>
<td>1.709</td>
<td>1962</td>
<td>31.2</td>
<td>Gulf Stream</td>
<td>CP000111</td>
<td>(Coleman et al. 2006)</td>
</tr>
<tr>
<td>MIT9202</td>
<td>—</td>
<td>—</td>
<td>HLII</td>
<td>1.691 (unfin.)</td>
<td>1891</td>
<td>31.9</td>
<td>S Pacific</td>
<td>ACDW00000000</td>
<td>S.W. Chisholm et al., unpublished data</td>
</tr>
<tr>
<td>NATL1A</td>
<td>2968</td>
<td>—</td>
<td>LLI</td>
<td>1.865</td>
<td>2201</td>
<td>35.1</td>
<td>N Atlantic</td>
<td>CP000553</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>NATL2A</td>
<td>2969</td>
<td>314</td>
<td>LLI</td>
<td>1.843</td>
<td>2158</td>
<td>35.0</td>
<td>N Atlantic</td>
<td>CP000095</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>SS120</td>
<td>1375</td>
<td>154</td>
<td>LLII</td>
<td>1.751</td>
<td>1884(^d)/1930(^d)/1925(^e)</td>
<td>36.4</td>
<td>Sargasso Sea</td>
<td>AE017126</td>
<td>(Dufresne et al. 2003, Kettler et al. 2007)</td>
</tr>
<tr>
<td>MIT9211</td>
<td>—</td>
<td>—</td>
<td>LLIII</td>
<td>1.689</td>
<td>1855</td>
<td>38.0</td>
<td>Eq. Pacific</td>
<td>CP000878</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>MIT9303</td>
<td>2917</td>
<td>—</td>
<td>LLIV</td>
<td>2.683</td>
<td>3022</td>
<td>50.1</td>
<td>Sargasso Sea</td>
<td>CP000554</td>
<td>(Kettler et al. 2007)</td>
</tr>
<tr>
<td>MIT9313</td>
<td>2773</td>
<td>407</td>
<td>LLIV</td>
<td>2.411</td>
<td>2275(^d)/2330(^d)/2843(^e)</td>
<td>50.7</td>
<td>Gulf Stream</td>
<td>BX548175</td>
<td>(Kettler et al. 2007, Rocap et al. 2003)</td>
</tr>
</tbody>
</table>

\(^a\)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 Bisoop cruise; data courtesy of C. Grob and O. Ullao). 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).

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

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 10⁵ 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 × 10⁵ 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 × 10⁴ cells ml⁻¹ in the upper 80 m (Grob et al. 2007) (Figure 1).

Analyses at the GYR station, near Easter Island (Figure 2a), 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 2c). 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 2c), 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 (~1.6 × 10¹³ 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).
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 2a), 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 Bioscope cruise, which sampled a very large Chl gradient from the Marquesas Islands to the Chilean coast (Figure 2a), revealed that the contribution of *Prochlorococcus* to the total Chl biomass ($t$Chla) of (sub)surface waters with chlorophyll below 0.03 $\mu$g L$^{-1}$ is generally lower than 25% (Figure 2b). The ratio of Chl $a_1$ to $t$Chla can exceed 50% but only between 0.035 and 0.2 $\mu$g L$^{-1}$ $t$Chla. Above this value, Chl $a_2$ concentration—which is directly related to densities of *Prochlorococcus* cells (see insert in Figure 2b)—can still be high, but the contribution of larger sized phytoplankton groups to $t$Chla 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.

**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).
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 (Garezarek et al. 2007), whereas the Gulf of Aqaba (Northern Red Sea) is populated by HLII (Zwirglmaier et al. 2008).

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. 2006).

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

---

**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 Bioscope 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 a₂) to total Chl a (tChl a), as measured by high-pressure liquid chromatography (HPLC) analysis in the upper mixed layer along the Bioscope transect. Note that the values >0.5 μg L⁻¹ tChl a have been omitted for readability since Chl a₂ is generally undetectable in these areas, and that all samples showing a nil Chl a₂ 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 a₂ 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 Bioscope 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.

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...
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, eS120 (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).

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
populations, Martiny et al. (2009c) recently cloned and sequenced a large
diversity using restriction fragment length polymorphism (RFLP) analyses
of the ITS region, also revealed
the occurrence of two novel uncultivated clades, tentatively called LLV and LLVI (P. Lavin,
clad 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 HaeIII-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.
Photosystems: protein complexes made of several subunits that are converting the electromagnetic energy from sun into chemical energy usable by the cell

**GENOMICS AND EVOLUTION**

**Specific Features of the Prochlorococcus Genus**

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 a2 and b2; 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 b2 from Chl a2 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, Stickfort 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. | MED4 locus tag in Mol|^a^ | MED4 locus tag in Genbank|^b^ | Gene name | Product | References |
|---------------------|----------------------|------------------------|-----------------|-----------|-----------|------------|
| 2331                | PMED4,08941          | PMM0808                | PCCao           | Chlorophyll b synthase (Chlorophyllide a oxygenase) | Nagata et al. 2005 |
| 9152d               | 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                | bhi             | High-light-induced protein (HLIP) | ≡ Cluster 10 in Bhaya et al. 2002 |
| 2025                | PMED4,15741          | PMM1385                | bhi             | 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 |

^aNumber of the cluster of orthologs in the Cyanorak database (http://www.sb-roscoff.fr/Phyto/cyanorak).  
^bLocus tag of the P. marinus MED4 sequence in the Microbes-on-Line database.  
^cLocus tag of the P. marinus MED4 sequence in the Genbank database.  
^dSequences 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,
of the genome content percentage G+C%:

316 Partensky

and GC content (52.5–66.0 G+C%) compared to ancient and diversified, marine common ancestor—than those of other Prochlorococcus of LLIV strains resemble more those of marine these two features likely decreased concomitantly during evolution. The genome characteristics (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 G+C 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 shrank 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 CO2 uptake, the CO2 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 pchA gene encoding the D1:2 isoform of the D1 protein, known to confer better resistance of PSII to photoinhibition. Instead, 1 to 3 pchA copies found in Prochlorococcus strains all encode a D1:1-like isoform (Garzarek 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 pchA 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* genomes), 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 (*psbU/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 photorepiration), 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

<table>
<thead>
<tr>
<th>Pathwaya</th>
<th>Gene name</th>
<th>Type of mutation</th>
<th>Cyonsorah Locus tag of one representative gene in MoL</th>
<th>Locus tag of one representative gene in Genbank</th>
<th>HL-I</th>
<th>HL-II</th>
<th>LL-I</th>
<th>LL-II</th>
<th>LL-III</th>
<th>LL-IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSBR -- Direct repair</td>
<td>pheA</td>
<td>DNA photolyase</td>
<td>—</td>
<td>1.33</td>
<td>PMED-Q1.3901</td>
<td>PMED0195</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- Direct repair</td>
<td>pheB/creB</td>
<td>DNA photolyase (or cryptochrome?)</td>
<td>—</td>
<td>1.984</td>
<td>NATL1.01.651</td>
<td>NATL1.01.651</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- Direct repair</td>
<td>—</td>
<td>DNA photolyase - related protein</td>
<td>—</td>
<td>1.46</td>
<td>PMED-Q2.6701</td>
<td>PMED0945</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- Direct repair</td>
<td>—</td>
<td>FAD domain-containing protein</td>
<td>—</td>
<td>3.63j</td>
<td>PMED-Q1.5291</td>
<td>PMED3140</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- Direct repair</td>
<td>—</td>
<td>Alkylated DNA repair protein, AlkB superfamily</td>
<td>—</td>
<td>1.46</td>
<td>P9313.03.61</td>
<td>PMED0878</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>ATP-dependent DNA ligase</td>
<td>—</td>
<td>3.613j</td>
<td>PMED-Q1.9001</td>
<td>PMED0679</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>ATP-dependent DNA ligase</td>
<td>—</td>
<td>1.447</td>
<td>PMED-Q2.9001</td>
<td>PMED0379</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>NAD5-dependent DNA ligase (oxytetron domain)</td>
<td>—</td>
<td>9.149j</td>
<td>PMED-Q1.7121</td>
<td>PMED0839</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>xseA</td>
<td>Exonuclease VII large subunit</td>
<td>—</td>
<td>1.185</td>
<td>P9313.27.79</td>
<td>PMED2184</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>xseB</td>
<td>Exonuclease VII small subunit</td>
<td>—</td>
<td>1.185</td>
<td>P9313.27.79</td>
<td>PMED2184</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>mutY</td>
<td>A/G-specific adenine glycosylase</td>
<td>G:C → A:T</td>
<td>3.31</td>
<td>MT9215.9447</td>
<td>P9313.20.9351</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>exoA</td>
<td>Exonuclease VII large subunit</td>
<td>—</td>
<td>1.16</td>
<td>P9313.20.71</td>
<td>PMED1641</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>exoB</td>
<td>Exonuclease VII small subunit</td>
<td>—</td>
<td>1.16</td>
<td>PMED2184</td>
<td>PMED1642</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>mutT</td>
<td>A/G-specific adenine glycosylase</td>
<td>G:C → A:T</td>
<td>3.31</td>
<td>P9313.20.71</td>
<td>PMED1641</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>—</td>
<td>Pyrimidine DNA glycosylase</td>
<td>—</td>
<td>2.679j</td>
<td>P9313.20.71</td>
<td>PMED1642</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>SSBR -- BERc/d</td>
<td>mutS</td>
<td>Exonuclease VIII</td>
<td>G:C → A:T</td>
<td>3.31</td>
<td>NATL1.13.271</td>
<td>NATL1.13.271</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

a) Following Partensky & Garczarek (2010).
<table>
<thead>
<tr>
<th>Category</th>
<th>Gene</th>
<th>Function Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SSBR-NER</td>
<td>phiB</td>
<td>6-0-alkylguanine DNA alkyltransferase-like</td>
</tr>
<tr>
<td></td>
<td>dprA</td>
<td>DNA-protecting protein DprA</td>
</tr>
<tr>
<td></td>
<td>recQ</td>
<td>ATP-dependent DNA helicase</td>
</tr>
<tr>
<td></td>
<td>recJ</td>
<td>Single-stranded DNA-specific exonuclease</td>
</tr>
<tr>
<td></td>
<td>Plosp</td>
<td>Single-stranded nucleoside triphosphate hydrolase, related to RecD (viral-like)</td>
</tr>
<tr>
<td></td>
<td>dprA</td>
<td>ATP-dependent helicase</td>
</tr>
<tr>
<td></td>
<td>recQ</td>
<td>ATP-dependent helicase</td>
</tr>
<tr>
<td></td>
<td>recJ</td>
<td>ATP-dependent helicase</td>
</tr>
<tr>
<td></td>
<td>Lhr-like</td>
<td>ATP-dependent helicase</td>
</tr>
<tr>
<td></td>
<td>hprB</td>
<td>ATP-dependent helicase</td>
</tr>
<tr>
<td></td>
<td>NUDIX</td>
<td>DNA/RNA helicases</td>
</tr>
<tr>
<td></td>
<td>nudD</td>
<td>GDP-mannose mannose hydrolase</td>
</tr>
<tr>
<td></td>
<td>udk</td>
<td>Putative uridine kinase</td>
</tr>
<tr>
<td></td>
<td>HNH family</td>
<td>DNA/RNA helicases</td>
</tr>
</tbody>
</table>

**Notes:**

- Genes are classified according to DNA repair pathways, mainly as described in the Kyoto Encyclopedia of Genes and Genomes (KEGG).
- 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.
- SSBR, single-strand breaks repair.
- BER, base excision repair.
- NER, nucleotide excision repair.
- MMR, methyl-directed mismatch repair.
- DSBR, double-strand breaks repair.
- HR, homologous recombination.

*This transversion effect was shown for *ada* and *ogt* genes and is likely to occur for *phiB* as well.*
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 bli 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 bli 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 bli 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 bli genes (Lindell et al. 2004), whereas others have been acquired by gene duplication. In MED4, for instance, a 4-bli gene cluster is found in two copies (bli06–09 and bli16–19) 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., \( phoA \) encoding alkaline phosphatase, one porin gene involved in phosphate transport, and several regulation genes such as \( phoB/R \) and \( ptrA \)) 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 \( \mu \text{M} \) 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).

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 (Zeilbell 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 HLII 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 oversimplified. 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 2c 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 2a).

LITERATURE CITED


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


## Contents

**Paleophysical Oceanography with an Emphasis on Transport Rates**  
*Peter Huybers and Carl Wunsch* ................................................................. 1

**Advances in Estuarine Physics**  
*Parker MacCready and W. Rockwell Geyer* ............................................... 35

**The Effect of Submarine Groundwater Discharge on the Ocean**  
*Willard S. Moore* .................................................................................. 59

**Marine Ecomechanics**  
*Mark W. Denny and Brian Gaylord* .......................................................... 89

**Sea Surface Temperature Variability: Patterns and Mechanisms**  
*Clara Deser, Michael A. Alexander, Shang-Ping Xie, and Adam S. Phillips* .......... 115

**Contemporary Sea Level Rise**  
*Anny Cazenave and William Llovel* .......................................................... 145

**Estimation of Anthropogenic CO2 Inventories in the Ocean**  
*Christopher L. Sabine and Toste Tansui* ................................................... 175

**Ocean Deoxygenation in a Warming World**  
*Ralph F. Keeling, Arne Körtzinger, and Nicolas Gruber* .............................. 199

**Archaeology Meets Marine Ecology: The Antiquity of Maritime Cultures and Human Impacts on Marine Fisheries and Ecosystems**  
*Jon M. Erlandson and Torben C. Rick* ....................................................... 231

**The Ecology of Seamounts: Structure, Function, and Human Impacts**  
*Malcolm R. Clark, Ashley A. Rowden, Thomas Schlacher, Alan Williams, Mireille Consalvey, Karen I. Stocks, Alex D. Rogers, Timothy D. O’Hara, Martin White, Timothy M. Shank, and Jason M. Hall-Spencer* .............................. 253

**Microbial Provinces in the Subseafloor**  
*Matthew O. Schrank, Julie A. Huber, and Katrina J. Edwards* ................. 279

**Prochlorococcus: Advantages and Limits of Minimalism**  
*Frédéric Partensky and Laurence Garczarek* ............................................. 305

**Oceanographic and Biogeochemical Insights from Diatom Genomes**  
*Chris Bowler, Assaf Vardi, and Andrew E. Allen* ....................................... 333
Genetic Perspectives on Marine Biological Invasions
   Jonathan B. Geller, John A. Darling, and James T. Carlton ..................... 367

Biocomplexity in Mangrove Ecosystems

What Can Ecology Contribute to Ecosystem-Based Management?
   Simon F. Thrush and Paul K. Dayton ...................................................... 419

Bioluminescence in the Sea
   Steven H.D. Haddock, Mark A. Moline, and James F. Case ........................... 443

Errata

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