Metabolic evolution and the self-organization of ecosystems

Rogier Braakman,a,b,1 Michael J. Follows,a and Sallie W. Chisholma,c,t

*Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139; †Department of Earth, Atmospheric, and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139; and ‡Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

Contributed by Sallie W. Chisholm, February 22, 2017 (sent for review November 28, 2016; reviewed by Ron Milo and John A. Raven)

Metabolism mediates the flow of matter and energy through the biosphere. We examined how metabolic evolution shapes ecosystems by reconstructing it in the globally abundant oceanic phytoplankter Prochlorococcus. To understand what drove observed evolutionary patterns, we interpreted them in the context of its population dynamics, growth rate, and light adaptation, and the size and macromolecular and elemental composition of cells. This multilevel view suggests that, over the course of evolution, there was a steady increase in Prochlorococcus’s metabolic rate and excretion of organic carbon. We derived a mathematical framework that suggests these adaptations lower the minimal subsistence nutrient concentration of cells, which results in a drawdown of nutrients in oceanic surface waters. This, in turn, increases total ecosystem biomass and promotes the coevolution of all cells in the ecosystem. Additional reconstructions suggest that Prochlorococcus and the dominant cooccurring heterotrophic bacterium SAR11 form a coevolved mutualism that maximizes their collective metabolic rate by recycling organic carbon through complementary excretion and uptake pathways. Moreover, the metabolic codependencies of Prochlorococcus and SAR11 are highly similar to those of chloroplasts and mitochondria within plant cells. These observations lead us to propose a general theory relating metabolic evolution to the self-amplification and self-organization of the biosphere. We discuss the implications of this framework for the evolution of Earth’s biogeochemical cycles and the rise of atmospheric oxygen.

Metabolism sustains the nonequilibrium chemical order of the biosphere by continually supplying the energy and building blocks of all cells on Earth (1–5). Here we ask: How does cellular metabolic evolution shape the mass and energy flows of ecosystems? The oceanic phytoplankter Prochlorococcus (6), the most abundant photosynthetic cell on Earth (7, 8), provides an ideal model system for addressing this question. Prochlorococcus and its deeper-branching sister lineage marine Synechococcus make up the marine picocyanobacteria and have a characteristic biogeography (9). Prochlorococcus “ecotypes” have geographically (10, 11) and seasonally (12) dynamic populations. To understand the observed patterns, we developed an evolutionary framework that illuminates the driving forces that produced many of the features of Prochlorococcus. Using this framework, we argue that the evolutionary patterns manifested in the Prochlorococcus collective reflect fundamental processes shaping the coevolution of the chemistry of the oligotrophic oceans and the microbial ecosystems they harbor.

Author contributions: R.B. designed research; R.B. performed research; R.B. contributed new reagents/analytic tools; R.B. developed the model; M.J.F. contributed to model development; R.B. and S.W.C. analyzed data; and R.B., M.J.F., and S.W.C. wrote the paper.

Reviews: R.M., Weizmann Institute; and J.A.R., University of Dundee, U.K.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

1To whom correspondence may be addressed. Email: rogierbraakman@gmail.com or chisholm@mit.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1619573114/-/DCSupplemental.

PNAS PLUS | EVOLUTION | EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES

www.pnas.org/cgi/doi/10.1073/pnas.1619573114

Published online March 27, 2017 | E3091–E3100
The remodeling of *Prochlorococcus*' metabolic core includes the disruption of photospiration and the TCA cycle (Fig. 3 and SI Appendix, Fig. S1), raising the possibility that intermediates of truncated pathways are excreted from the cell. Phytoplankton commonly excrete organic carbon as an outlet of excess reducing power under nutrient limitation or intense light (24). Analogously, when facing a large energy supply from organic carbon, some heterotrophs will effectively drain reducing power into the environment by excreting partially oxidized organic carbon rather than fully oxidizing it to CO$_2$ (25–28), whereas some photoheterotrophs use CO$_2$-fixation as a sink for excess reducing power (29). For photosynthetic cells in the oligotrophic surface oceans, where the solar energy supply may commonly outpace the nutrient supply (Fig. 1), the combination of increased CO$_2$ fixation and increased excretion of organic carbon could thus well be under strong selection. This is consistent with observations that *Prochlorococcus* has the most efficient carbon-concentrating mechanism (30) and highest known rate of CO$_2$ fixation per photosynthetic pigment (31) of any phytoplankton, even though its small size and slow growth (6, 32) suggest a relatively small carbon flux requirement. Lastly, selection to rid the system of excess reducing power is consistent with the acquisition of the Plastoquinol Terminal Oxidase (PTOX) in the LLI and HL clades of *Prochlorococcus* (Fig. 3 and SI Appendix, Table S1), which adds an additional outlet for excess reducing power in their photosynthetic electron transfer chain (SI Appendix, Fig. S1) (33–36).

That the evolution of *Prochlorococcus* permanently increased the excretion of organic carbon in its late-branching strains is also consistent with limited laboratory observations. At moderate light levels in nutrient replete medium, strains from HL clades that dominate surface waters (Fig. 1) excrete up to 20% of fixed carbon, and recently diverging strains excrete the most (37). Cells excrete significant amounts of glycolate (37), one of the dead ends in the metabolic network (SI Appendix, Fig. S1). P-starved cells excrete slightly less carbon (37), but this could be due to the coincident cessation of growth. Nevertheless, a higher fraction of the carbon excreted by P-starved cells consists of glycolate and other small carboxylic acids (37). Many phytoplankton excrete glycolate under intense light or nutrient limitation (24, 38), but retain the capacity to recycle it by using the three-subunit iron–sulfur protein glycolate oxidase (GlcDEF, rxn 13 in SI Appendix, Fig. S1), shown to be the enzymatic workhorse for this function in cyanobacteria (23). However, this gene is absent in all but the deepest-branching LLIV clade of *Prochlorococcus* (SI Appendix, Table S1), suggesting a permanent opening of this pathway early in its evolution (Fig. 3). Finally, the bulk of organic carbon excreted by *Synechococcus* consists of polysaccharides (39), which are commonly excreted by nutrient-limited microbes (26, 40), suggesting that these compounds could similarly act as a redox safety valve in *Prochlorococcus*, which dominates in one of the most nutrient-poor environments on Earth (7–9) (Fig. 1).

We further examined the possibility that the evolution of *Prochlorococcus* increased its excretion of organic carbon as an outlet of excess reducing power (Fig. 3) through additional genomic analyses. Functionally related and coexpressed genes are commonly located near each other, so we searched the genomic neighborhoods of core metabolic genes for transporters across clades (SI Appendix, Fig. S2). We identified chromosome rearrangements repositioning a series of transporters, including three export and one import transporters, near key metabolic genes, consistent with selection acting to fortify the control of transport pathways (SI Appendix, Fig. S2). Chromosome rearrangements are seen in freshwater picocyanobacteria (SI Appendix, Fig. S2), again suggesting that these pathways came under selection before the emergence of the marine lineages and the loss of photospiration in the LLI/III clade of *Prochlorococcus* (Fig. 3 and SI Appendix, Fig. S1). This analysis suggests that, in addition to glycolate, pyruvate and citrate (or isocitrate) are exported, whereas malate is imported (SI Appendix, Figs. S1 and S2).

Furthermore, because environmentally driven variations in the expression of genes can give insight into their function [i.e., “reverse ecology” (41)], and the metabolism of *Prochlorococcus* is highly choreographed to the diel light:dark cycle (42, 43), we examined the gene expression of a HL-adapted strain grown under a diel cycle (42). All putative export pathway genes have mRNA expression maxima at sunrise (SI Appendix, Fig. S2), consistent with their acting as redox safety valves, activated when the supply of reducing power increases at dawn. Similarly,

**Fig. 1.** Typical relative abundance distributions of *Prochlorococcus* ecotypes as a function of depth and accompanying light intensity and nutrient concentration profiles in stratified oceanic waters. Ecotype populations are geographically and temporally dynamic, but in warm, stable water columns return to this same depth-differentiated state (12). The deepest branching ecotypes are most abundant at the bottom of the euphotic zone, where nutrient concentrations are high and light energy low. The most recently diverging ecotypes are most abundant near the surface, where the reverse is true (10–14). HL, high-light-adapted; LL, low-light-adapted.

**Fig. 2.** Illustration of approach to metabolic reconstructions. Phylometabolic trees reflect the evolution of metabolic network phenotypes because they integrate constraints from phylogenetics and metabolism (15, 16). All sequenced genomes within a given clade are searched for the presence/absence of enzymes catalyzing the reactions of different pathways. Mapping pathway variability patterns onto phylogenies of the clade suggests the order of metabolic innovations. In this example, three alternative pathways (pink, yellow, and blue) connect essential and universal pathways (black). Genes for the yellow pathway are nearly universally distributed (Inset), suggesting that it is the ancestral pathway, with the pink and blue pathways deriving from it. Maintaining continuity of flux results in trees of functional phenotypes. Biochemical differences between alternative pathways (e.g., ATP/trace metal requirements or oxygen sensitivities of their enzymes) suggest evolutionary driving forces.
the expression of PTOX, the most immediate outlet for excess reducing power in the electron transfer chain (SI Appendix, Fig. S1), is maximal at midday (35, 42). Collective evidence thus suggests that the evolution of Prochlorococcus steadily added redox outlets, leading to increased excretion of organic carbon. Lastly, expression of the putative malate uptake pathway peaks at dusk (SI Appendix, Fig. S2), suggesting that the rest of the ecosystem may return some organic carbon to Prochlorococcus at night.

What drove Prochlorococcus’ increased excretion of organics over the course of its evolution? If maintaining redox balance was the driving force, why not enhance the ability to lower photosynthetic electron flux (44, 45), which many photosynthesizers use as the driving force, why not enhance the ability to lower photosynthetic electron flux? A decreasing nutrient flux density during the evolution of Prochlorococcus thus appears to reflect a more general trend that increases the total cellular throughput of electrons (Fig. 3). Various genes related to the protection and repair of light damage have been added in HL clades of Prochlorococcus (52, 53), suggesting that later divergences were primarily related to fortifying cellular machinery to allow cells to generate an increased electron flux at the highest photon fluxes near the surface (Fig. 3).

In addition to an increasing electron flux density ($$\nu_e$$), integrating evidence across levels suggests the evolution of Prochlorococcus also decreased its limiting nutrient flux density, $$\nu_n$$, [moles of $n$ (g dry weight)$^{-1}$ time$^{-1}$)]. For steady-state growth under constant conditions, $$\nu_n$$ can be parameterized as:

$$\nu_n = \mu Q_n$$  

[2]

where $\mu$ is the specific growth rate (time$^{-1}$) and $Q_n$ is the mass-normalized cellular quota of limiting nutrients [moles of $n$ (g dry weight)$^{-1}$]. A decreasing nutrient flux density during the evolution of Prochlorococcus is suggested by a decrease in its maximal intrinsic growth rate $\mu_{max}$ (time$^{-1}$) relative to Synechococcus
E3094
(32) and an increased efficiency in its use of limiting nutrients. The latter is inferred from many features of Prochlorococcus, including less investment of N and P in the genome by reducing its size and guanine–cytosine content (60), decreased use of amino acids with N-rich side chains (61), a switch from P- to sulfolipid membranes (62, 63), and less Fe use in metabolism and photosynthetic machinery (33, 34) (SI Appendix, Fig. S1).

These changes suggest a decreasing Qe for N, Fe, and P, the three main limiting nutrients in the oligotrophic oceans (64). An increasing \( \nu_e \) coinciding with a decreasing \( \nu_i \) over the course of Prochlorococcus evolution can be expressed as a single variable: the electron-to-nutrient flux ratio, \( \nu_e/\nu_i \) (Fig. 3).

Electron Flux and Nutrient Acquisition. What are the selective pressures maximizing \( \nu_e/\nu_i \) in Prochlorococcus? Photosynthetic electron flux transfers solar energy into metabolism via cofactors such ATP and NAD(P)H (65). An increased electron flux [jmol electrons (g dry weight)\(^{-1}\) time\(^{-1}\)] thus suggests an increased metabolic rate [kJ of absorbed solar energy (g dry weight)\(^{-1}\) time\(^{-1}\)]. This principle is exemplified by the increasing metabolic activity and ATP/ADP ratios of plant chloroplasts and cyanobacteria when they shift from darkness into light (66, 67). Furthermore, the highest \( \nu_e/\nu_i \) phenotypes are the most recently diverging clades (Fig. 3), dominating near the surface where the solar energy supply is greatest and nutrient levels are lowest (Fig. 1), suggesting an advantage to higher metabolic rates at lower nutrient concentrations.

How metabolic rate affects nutrient uptake can be understood from Michaelis–Menten kinetics, which under strong nutrient limitation \( [n] \ll K_{M,n} \) simplifies to (68) (SI Appendix, SI Text):

\[
\nu_i = [n] a_i^n = V_{max} [n] / K_{M,n},
\]

where \( [n] \) is the nutrient concentration (moles of \( n \) L\(^{-1}\)) and \( a_i^n \) is the specific nutrient affinity [L (g dry weight)\(^{-1}\) time\(^{-1}\)]. Specific affinity indicates how strongly cells absorb limiting nutrients (analogous to the pumping speed of a vacuum pump) and is equivalent to the saturated maximal nutrient uptake rate \( V_{max} \) [moles of \( n \) (g dry weight)\(^{-1}\) time\(^{-1}\)] over the Michaelis constant \( K_{M,n} \) (moles of \( n \) L\(^{-1}\)) (68) (SI Appendix, Fig. S3). \( V_{max} \) reflects the maximum handling rate for absorbed nutrients and is proportional to the metabolic rate (68, 69). The free energy cost \( \Delta G \) (kJ mol\(^{-1}\)) of nutrient transport scales with natural log of the ratio of internal to external nutrient concentrations (65) (SI Appendix, SI Text) and can become very large in the extremely nutrient-poor oligotrophic oceans (69). For example, the free energy cost of phosphate uptake in the oligotrophic oceans is far greater than the free energy gain of ATP hydrolysis unless the ATP/ADP ratio is increased drastically from commonly assumed metabolite concentrations of 1 mM for ATP, ADP, and P\(_i\) (SI Appendix, SI Text). An analogous, but slightly different, situation occurs for ammonia (NH\(_3\)) uptake, which has the potential for a major futile cycle in the oligotrophic ocean, because its conjugate base NH\(_4\)\(^+\) passively diffuses out of the cell (70) (SI Appendix, SI Text). It is therefore thought that cells pool their internal [NH\(_4\)\(^+\)] at the minimal viable value (71), which in turn similarly requires a significant increase in the ATP/(ADP \( \times P_i\)) ratio to drive forward glutamine synthesis (glutamate + NH\(_3\) + ATP = glutamine + ADP + P\(_i\)), the central highway for nitrogen into metabolism (65). (See SI Appendix, SI Text for detailed calculations and discussion.) Finally, while kinetic bottleneck reactions can be driven forward by increasing the ATP/(ADP \( \times P_i\)) ratio, this simultaneously increases the free energy cost of ATP synthesis, therefore requiring a greater proton motive force, and thus ultimately a greater photosynthetic electron flux (69).

The principles just outlined suggest that innovations increasing \( \nu_e \) may allow the minimal subsistence concentration of limiting nutrients \( [n]^* \) (the lowest value of \([n]\) at which net positive growth is possible). When growth and loss processes are relatively rapid and tightly coupled, microbial strains with the lowest \( [n]^* \) dominate (72), suggesting that selection should favor such innovations. To understand how selection to lower \( [n]^* \) shapes cells, we can substitute an expression for \( V_{max} \) that assumes reversible kinetics (SI Appendix, SI Text) into Eq. 3:

\[
[n] = K_{M,n} V_{max} \nu_i = K_{M,n} V_{max} \frac{\nu_i}{[E]^*} 1 - \frac{\nu_i}{[E]^*} e^{-r \nu_i / \Delta T},
\]

where \([E]^*\) is the enzyme concentration [moles of enzyme (g dry weight)\(^{-1}\)], \( k^*\) is the rate constant [moles of \( n \) (moles of enzyme)\(^{-1}\) time\(^{-1}\)], \( R \) is the gas constant (J K\(^{-1}\) mol\(^{-1}\)), and \( T \) is temperature (K). Eq. 4 suggests that there are two strategies for lowering \( [n]^* \) (SI Appendix, Fig. S3). First, cells can modify the kinetics (decreasing \( K_{M,n} \) and/or increasing \( k^*\) and) thermodynamics (decreasing \( \Delta G \) of their metabolism, the latter by increasing \( \nu_i \) as we have just argued. Second, cells can lower their \([n]^*\) by lowering their required flux of limiting nutrients \( \nu_i \), (Eq. 4), which can be achieved by lowering their minimal \( Q_e \) [i.e., streamlining (73)] or their \( \mu \) (Eq. 2), both of which are apparent in the evolution of Prochlorococcus as discussed above. Optimizing kinetics/thermodynamics and decreasing \( \nu_i \) can work synergistically, and both are helped by a decrease in cell mass, as is observed along the Prochlorococcus phylogeny (6, 59). That is, selection to lower \( \nu_i \) allows a decrease in total cell mass by minimizing the mass dedicated to nonessential components (73).

The amount of metabolic enzyme and photosynthetic machinery is kept fixed, this decrease in total cell mass would increase both \( [E]^* \) (68) and \( \sigma_{PSII} \) (and thus \( \nu_e \), which makes \( \Delta G \) more negative). For nitrogen, increasing \( \nu_i \) in turn provides an additional avenue for lowering \( Q_e \) (and thus \( \nu_e \)), because pathways with a more negative \( \Delta G \) require less protein biomass for a given flux (74). Thus, maximizing \( \nu_e/\nu_i \) lowers the \([n]^*\) of cells by making the \( \Delta G \) more negative (driving kinetic bottleneck reactions forward), thereby increasing the nutrient affinity (68, 69). As a result, the evolution of Prochlorococcus has driven nutrients to vanishingly low levels (<0.1 nM) in the oligotrophic oceans (64).

Benefits of Excreting Organic Carbon. If selection to lower \([n]^*\) drives the maximization of \( \nu_e/\nu_i \), why should it lead to an increased excretion of organic carbon? We argue this ultimately emerges from mass and energy conservation. That is, in the presence of kinetic bottlenecks, cells can drive up their ATP/ADP ratio by increasing their ATP supply rate, but to maintain steady state, they must also increase ATP consumption rates (49, 69). The same argument applies to the NAD(P)H/NAD(P) ratio. This is consistent with observations that some nutrient-limited aerobic chemoheterotrophs have increased glucose uptake rates, increased levels of respiration, increased excretion of polysaccharides (whose synthesis from glucose requires ATP consumption), and a high flux through various ATP-consuming futile cycles (25, 26, 49). For photosynthetic cell CO\(_2\)-fixation is a major sink for ATP and NAD(P)H, but cells are limited in the carbon flux density, \( \nu_C \), they can accommodate. This limit is proportional to growth rate (Eq. 2), and, because selection to lower \([n]^*\) favors relatively slow-growing cells (Eq. 4), it is lower in the oligotrophic oceans. Thus, we argue that excreted organic carbon represents a kind of “carbon exhaust” that allows cells to maximize their nutrient affinity by increasing their metabolic rate above limits arising from carbon saturation.

This general mechanism is further illustrated by expanding Eq. 4. We assume that the electron flux must support carbon fixation sufficient to build biomass at a rate dictated by \( \nu_{\nu_i} \), that biomass has an elemental stoichiometry \( Q_{O_X}/Q_{C_C} \), that the efficiency of fixation is \#C/#e (carbon atoms electron\(^{-1}\)), and that a fraction \((0<\chi<1)\) of carbon is excreted/respired. \#C/#e depends on the oxidation states of the carbon source, biomass, and excreted carbon. For example, if the carbon source is CO\(_2\) and biomass and excreted carbon have the oxidation state CH\(_3\)O\(_2\), then \#C/#e is 1/4 (i.e., CO\(_2\) + 4\(^{+}\) + 4H\(^{+}\) = CH\(_3\)O\(_2\) + 4H\(^{+}\))
H₂O). Photoprotective mechanisms like the water–water cycle of PTOX (SI Appendix, Fig. S1) act to lower #C/#c. Together, these assumptions lead to the expression (see also SI Appendix, SI Text):

\[
\left[ n \right] = \frac{K_{M,n}}{[E][k^+]^{1 - \eta}} = \frac{K_{M,n} Q_n \nu_n \left( \# C / \# e \right) (1 - \beta)}{[E][k^+] Q_C (1 - e^{-\Delta G^{RT}/RT})},
\]

Thus, while increasing metabolic rate (i.e., lowering \( \Delta G, \beta \)) lowers \( [n] \), it also acts to increase \( [n] \) by increasing the carbon flux density, unless excess carbon is excreted \( \beta > 0 \) and increasing faster than \( \nu_n (\# C / \# e) \) (Eq. 5). Selection to lower \( [n] \) would thus favor a decrease in \( \nu_n \) and simultaneous increases in \( \nu_e \) and \( \beta \) (carbon excretion), exactly as we observed for the evolution of Prochlorococcus (Fig. 3).

**Metabolic Rate, Ecosystem Biomass, and Coevolutionary Dynamics.** What are the evolutionary and ecological consequences of Eq. 5? Focusing on the evolution of Prochlorococcus, it suggests that the layered population structure observed in stable water columns (Fig. 1) (10–14) reflects the sequential evolution of new ecotypes near the surface, each with increasing metabolic rates, drawing down limiting nutrients and restricting ancestral ecotypes to ever deeper waters (Fig. 4). This type of evolutionary dynamic in which a key innovation causes a population to expand along a shifted adaptive landscape produces adaptive radiations (75, 76) and has been argued to be the dominant contributor to global microbial diversity (77). This is consistent with observations that the broadly defined ecotypes of Prochlorococcus consist of hundreds of stable subpopulations that diverged millions of years ago (78). These are “niche constructing” adaptive radiations, since the shift in the adaptive landscape arises because new ecotypes chemically modify their own environment (Fig. 4) (79).

Subpopulations form clusters defined by the acquisition of small cassettes of genes involved in electron transfer, redox stress, and/or synthesis of membrane polysaccharides (78). Increased production of the latter is a recognized outlet of excess reducing power (25, 26, 40). In addition to providing an electron sink, membrane polysaccharides also mediate biological interactions (e.g., with phytoplankton, grazers, or other bacteria), suggesting that an increase in electron flux could have produced ecological feedbacks that drove further differentiation within Prochlorococcus ecotypes (78) and in those with which it interacts (80). Lastly, this framework suggests that by promoting the fixation and excretion of increasing amounts of organic carbon as a by-product of increasing the nutrient affinity, the evolution of Prochlorococcus increased the long-term steady-state concentrations of dissolved organic carbon (DOC) in the oligotrophic oceans (Fig. 4), which are much higher (by a factor of up to ~2) than those observed in the ocean (81). Because several compounds that Prochlorococcus may be excreting are known iron-binding ligands, including polysaccharides (82) and carboxylic acids like citrate (83), it may, in turn, also play a key role in buffering trace metal bioavailability in these environments.

When extended to the ecosystem level, Eq. 5 provides a mechanistic view of how the evolution of all microbial cells in the open ocean is interconnected via the chemically coevolving environment (79, 84): Any innovation in any lineage that increases the energy flux and lowers nutrient concentrations (Fig. 4) pushes all other lineages to follow suit and adopt similar innovations. Unlike in the classic “Red Queen hypothesis,” in which evolution is a zero-sum game (85), here, the evolutionary dynamic increases resource capture, and thereby biomass, of the ecosystem. [Similarly, a palaeontological survey of the distribution and body size of marine animals also led to the conclusion that the energy flux and biomass of ecosystems increases over evolutionary time (86).] The imprint of this collective dynamic can be seen in the broadly convergent features of all oceanic microbes. Slow growth, small cell size, streamlined genomes and proteomes, and use of nonphospholipid membranes—signature features of Prochlorococcus—are observed across both autotrophic and heterotrophic microbes in the oligotrophic oceans (61, 73, 87–89). Some of the photosynthetic machinery modifications that increased the electron flux of Prochlorococcus (Fig. 3) are also seen in diatoms and picoeukaryotes (33, 90, 91).

Many oceanic heterotrophs in turn supplement their energy supply by feeding on organics by capturing sunlight with proteorhodopsins (92) and possess electron drains in their respiratory electron transfer chains (93). In general, heterotrophic growth in the oligotrophic oceans favors metabolic rate over efficiency—the fraction of carbon taken up from the environment that is converted into biomass is one of the lowest of any aquatic environment on Earth (94). These observations are all consistent with the notion that the maximization of \( \nu_e / \nu_n \), and thereby a lowering of steady-state nutrient concentrations (Eq. 5 and Fig. 4), has been a general evolutionary driving force in ocean ecosystems.

**Emergence of Mutualism in Oceanic Microbial Ecosystems.** Finally, if lowering \( [n] \) increases the excretion of organic carbon into the environment, it could produce new opportunities for cooccurring heterotrophs, like the ubiquitous and abundant SAR11 (73, 95). SAR11 requires pyruvate and either glycate, glycine, serine, or glycine betaine (96); the latter four all feed into the same pathway that starts from glycate (96). Furthermore, coastal SAR11 strains can replace pyruvate with glucose metabolized via glycolysis, whereas open ocean strains lack glycolysis and have an obligate requirement for pyruvate (97). Similar adaptations are not observed in freshwater strains of SAR11 (98). Thus, oceanic SAR11 populations have evolved a dependency on exactly the compounds (pyruvate and glycate) that our metabolic reconstructions suggest emerged as excretion pathways in Prochlorococcus (Fig. 3 and SI Appendix, Fig. S1).

We mapped the distribution of metabolic genes across clades (SI Appendix, Table S2) onto the phylogeny of marine SAR11 clades to further reconstruct the evolution of their metabolic core (96–98) (Fig. 5 and SI Appendix, Fig. S4). We aimed to resolve the innovations of open ocean lineages and look for evidence of selection on pathway controls by looking for transporters in the vicinity of metabolic genes. As in Prochlorococcus (Fig. 3 and SI Appendix, Table S1), the metabolic core of SAR11 (SI Appendix, Fig. S4) evolved through a sequence of innovations (Fig. 5 and SI Appendix, Table S2), including the step-wise completion of the glyoxylate shunt and the well-documented switch...
from Emden–Meyerhoff–Parnass (EMP) to Entner–Doudoroff (ED) glycolysis (97). Glycolysis is disrupted in the 1.3A clade (97) (Fig. 5 and SI Appendix, Table S2), which is most abundant in surface waters where the Prochlorococcus HLII clade dominates (Fig. 1) (99, 100). Glycolate uptake is also lost in this SAR11 clade (Fig. 5 and SI Appendix, Table S2), which suggests that pyruvate produced by the Prochlorococcus HLII ecotype may have become its central source of carbon and energy (97). Finally, similar to Prochlorococcus (SI Appendix, Fig. S2), chromosomal rearrangements positioned a series of transporter proteins near key metabolic genes in SAR11 (SI Appendix, Fig. S5), consistent with selection on the control of transport pathways. We identified putative import transporters for pyruvate and glycolate, and putative export transporters for citrate and malate, the latter exclusive to the IA clade (Fig. 5 and SI Appendix, Fig. S5). The emergence of malate export in SAR11 would provide a potential source for the emergent malate uptake pathway of Prochlorococcus that is putatively activated at night (Fig. 3 and SI Appendix, Figs. S1 and S2).

Furthermore, the glyoxylate shunt is a TCA cycle bypass activated under redox stress in some microbes (101–103), while the evolutionary switch from EMP to the higher-rate ED variant of glycolysis has, in other microbes, been attributed to an increased energy supply (74). These observations are consistent with selection acting to maximize $\nu_e/\nu_n$ in both systems and thereby producing pathways that transfer pyruvate and glycolate from Prochlorococcus to SAR11 and malate from SAR11 to Prochlorococcus. Oligotrophic waters have nanomolar concentrations of pyruvate and glycolate, with midday maxima for the latter, consistent with biological cross-feeding synchronized with the input of sunlight, although abiogenic photochemistry may also contribute (104, 105).

Additional evidence for metabolic mutualism in these systems comes from observations regarding hydrogen peroxide (HOOH), a by-product of biological electron transport, photochemistry, and other abiotic processes (106). Prochlorococcus and some later-diverging clades of SAR11 have lost HOOH-detoxifying catalase (SI Appendix, Tables S1 and S2), and Prochlorococcus grows better in the presence of bacteria retaining catalase (107, 108). This led to the “Black Queen” hypothesis, which argues that subpopulations of ecosystems can save essential nutrients by giving up inevitably shared functions (such as detoxifying the freely diffusible HOOH), so long as they are preserved by others in the ecosystem (109). Our reconstructions suggest that the loss of catalase in Prochlorococcus and SAR11 coincides with increased excretion of organic carbon, which provides carbon and energy for catalase-containing bacteria (109).

These observations suggest that metabolic mutualisms are self-amplifying feedback loops (110) that maximize the collective $\nu_e/\nu_n$ and thus total productivity, of ecosystems. Specifically, recycling otherwise wasted electrons through complementary excretion/uptake pathways increases the average $\nu_e$ of participating cells. Similarly, the loss of functions that are shared and require limiting nutrients (e.g., iron in catalase) in some members of the community decreases the average $\nu_n$. Mutualisms thus raise the $\nu_e/\nu_n$ of ecosystems—and lower the subsistence nutrient requirements of their cells (Eq. 5 and Fig. 4)—beyond what is possible for individual lineages in isolation. Because excreting organic carbon lowers the minimal subsistence nutrient requirements of individual cells, mutualisms of this kind are emergent properties of ecosystems and avoid public goods dilemmas (48). An upper bound may exist on the maximization of $\nu_e/\nu_n$ due to the minimal requirements of being an autonomous cell. As the smallest photosynthetic cell (6), Prochlorococcus may be closest to this limit, reinforcing the notion that it has a central role in shaping the features of ecosystems in the surface oceans.

Are Plant Cells Microscopic Analogs of Oceanic Microbial Ecosystems? It occurred to us that the metabolic organization of oceanic microbial ecosystems and plant cells may bear a striking resemblance (Fig. 6): Intermediates of lower glycolysis and photorespiration are central conduits of electron transfer from Prochlorococcus to SAR11 and from chloroplasts to mitochondria in plant cells, while TCA cycle intermediates facilitate electron transfer in the opposite direction in both systems (111–113). Similar patterns emerge at other levels of organization. Microbes other than Prochlorococcus/SAR11 in ocean communities, and organelles other than chloroplasts/mitochondria in plant cells, appear central to peroxide detoxification (108, 115). The heterotrophic bacteria SARS6 and SAR116 may be important for this function in the oligotrophic oceans, because they both possess catalase (114–116) and are abundant in warm stratified waters (117). Prochlorococcus and chloroplasts both have PTOX as an electron drain in their photosynthetic electron transfer chain, whereas SAR11 and mitochondria both have alternative oxidase (AOX) as an electron drain in their respiratory electron transfer chain (93, 113). Furthermore, like chloroplasts, Prochlorococcus uses chlorophyll b in addition to chlorophyll a, which has not been observed in cyanobacteria other than Prochloron and Prochlorotrix (6). Finally, organelles of plant cells have also undergone reductive genome evolution, and, for mitochondria, it has been argued that this increased the cellular power density of eukaryotes (118)—similar to what we argued for oceanic microbial ecosystems. The extensive convergence of plant cells and oceanic microbial ecosystems highlights the constraints that metabolism imposes on the large-scale structure of evolution (5) and suggests that the metabolic codependencies of eukaryotic organelles can evolve without the physical intimacy of endosymbiosis.

Biospheric Self-Amplification and the Rise of Atmospheric Oxygen. We have proposed that maximizing the metabolic rate of cells lowers their minimal subsistence nutrient requirements and that this is achieved by maximizing the cellular electron-to-nutrient flux ratio $(\nu_e/\nu_n)$, while increasing the excretion of organic carbon (Eq. 5). This leads to an evolutionary dynamic that increases total ecosystem biomass (Fig. 4) and paves the way for self-amplifying feedback loops that recycle organic carbon (Fig. 6) and reinforce the maximization of cellular metabolic rate at the ecosystem level. It has been argued that the hierarchical organization of pathways within metabolism reflects the outgrowth of self-amplifying feedbacks that increased the free energy consumption of the emerging biosphere (5, 119). Our framework extends those arguments into the world of phenotypically differentiated cells and microbial ecosystems, and is consistent with the theorem that the flow of energy through the biosphere promotes its self-organization into chemical cycles (2).
Our framework has implications for Earth history. If biospheric self-amplification driven by the sun enhanced the burial of organics and carbonates (120, 121) simply by increasing their production, it would help explain the drawdown of atmospheric \( \text{CO}_2 \) and the rise of atmospheric \( \text{O}_2 \) across several stages of Earth history (122, 123). Perhaps not coincidentally, marine picocyanobacteria are estimated to have emerged near the transition from the Neoproterozoic (1,000–541 Ma) to the Phanerozoic (541 Ma to present) (124, 125), when sediments indicate a higher metabolic rate and correspondingly greater ability to mobilize Fe and P under oxidizing conditions, and is associated with a higher metabolic rate and correspondingly greater ability to mobilize Fe and P under oxidizing conditions, and is associated with a higher metabolic rate and correspondingly greater ability to mobilize Fe and P under oxidizing conditions. Extant oceanic microbes surmount this negative feedback on photosynthetic electron transfer through reduced cellular Fe demands (33, 90, 91) (SI Appendix, Fig. S1), and through Fe-ligation by DOC (145), including polysaccharides (82), citrate (83), and other carboxylic acids, all of which may be excreted by marine picocyanobacteria (37, 39) (SI Appendix, Fig. S1). Polysaccharides and small carboxylic acids also enhance the dissolution of minerals (144), and minerals in wind-blown dust are a major source of Fe (145) and P (146) to the surface oceans. Thus, we hypothesize that the evolution of marine picocyanobacteria (Fig. 4) increased both the bioavailability and the overall supply of iron under aerobic conditions and helped transform the oceans from an anoxic state rich in free iron (135, 136) to an oxygenated state (131, 133) with DOC-bound iron. This positive iron-self-amplification feedback, in concert with an increased metabolic rate, was critical in pushing the marine biosphere past a major evolutionary bottleneck and paved the way for an expansion of oceanic oxygenic photosynthesis and a rise in atmospheric \( \text{O}_2 \) (123).

Sedimentary and genomic records suggest several additional positive feedbacks that could have pushed forth the Phanerozoic oxygenation of the ocean. Increased Fe bioavailability under aerobic conditions coinciding with the drawdown of nitrogen (Fig. 4) would have created opportunities for \( \text{N}_2 \)-fixers, while ocean oxygenation would have lifted their \( \text{Mo} \)-limitation by suppressing euxinia (137, 138), together increasing the supply of nitrogen to the oceans (124, 138, 139). This is consistent with the suggested overlap in the rise of marine picocyanobacteria and planktonic \( \text{N}_2 \)-fixers (124, 125). Furthermore, sediments suggest an increase in oceanic P levels after the Neoproterozoic, and it was argued that this was because a drop in Fe concentrations lessened the scavenging of P (140). We add that if enhanced DOC-dissolution of minerals from dust enhanced the oceanic iron supply under aerobic conditions, as we argued above, it could have also increased the P supply (140, 146), thus contributing to the reconstructed rise in P levels (140).

These scenarios are similar to those of how the rise of land plants impacted the Earth system. Nutrient-limited plants leach Fe, P, and other nutrients from rocks by excreting small carboxylic acids from their roots (which suggests an increased metabolic rate—Eq. 5) (147). It has therefore been argued that plant colonization of the continents during the Phanerozoic increased the weathering of rocks, and, in turn, the precipitation of carbonates, which, together with the increased burial of plant-derived organics, resulted in a drawdown of atmospheric \( \text{CO}_2 \) and a rise in atmospheric \( \text{O}_2 \) (148–150).

The convergent metabolic evolution of oceanic microbial ecosystems and land plants (Fig. 6), which, we have argued, may have impacted the Earth system in similar ways, suggests the temporal profile of the Earth's oxygenation (122, 123) may be constrained by two biological stages. The first consisted of the expansion of self-amplifying cyanobacterial \( \text{O}_2 \)-photosynthesis in shallow aquatic environments and is associated with the GOE (122, 123). The second consisted of the global expansion of eukaryotic or "eukaryote-like" \( \text{O}_2 \)-photosynthesis—both onto the continents and into the deep open ocean (Fig. 6)—with a higher metabolic rate and correspondingly greater ability to mobilize Fe and P under oxidizing conditions, and is associated with...
with the Neoproterozoic Oxidation Event (122, 123, 132, 142). Genomic studies estimate that chloroplast endosymbiosis leading to the rise of all photosynthetic eukaryotes occurred between the late Paleoproterozoic (2,500–1,600 Ma) and the early Neoproterozoic (1,100–900 Ma) (139, 140). Paleontological studies in turn find a significantly increased fossil diversity of eukaryotes (including photosynthetic eukaryotes) in the Neoproterozoic (155, 156), whose increasing body sizes and fecal pellets could have moreover strengthened the export and burial of organic carbon from the oceans (157, 158).

As a final note, one could argue that the emergence of modern human societies is a variant of the general framework we propose. As our populations expanded and extracted ever more electrons from fossil fuels, we have increased global CO2, while draining away natural resources and global O2 (albeit a small amount relative to the contemporary inventory for the latter) (159, 160). In the process, we have become increasingly socially, technologically, and economically interconnected, analogous to what we have observed in the evolution of oceanic microbial ecosystems and plant cells. As in those systems, this has increased our collective ability to harvest more difficult-to-access natural resources. Managing the biogeochemical perturbation that our global emergence is imposing on the Earth system is one of humanity’s greatest challenges.

Materials and Methods

We analyzed 56 genomes of cyanobacteria representative of the diversity of this clade (161) obtained from the UniProt website, 56 genomes of Prochlorococcus and marine Synechococcus (162), and 16 SAR11 genomes (163). We reconstructed phylomaculae trees (Fig. 2) by mapping the distributions of metabolic genes onto the phylogenies of Prochlorococcus (Figs. 1 and 3) and SAR11 (Fig. 5).

ACKNOWLEDGMENTS.

We thank the reviewers for critical and insightful reviews, and we thank Jamie Becker, Shane Hogle, and John Casey for comments on an earlier version of our manuscript. R. B. is a Simons Foundation Fellow of the Life Sciences Research Foundation. This work was supported by Simons Foundation Grant SCOPE 329108 (to S.W.C. and M.J.F.); Gordon and Betty Moore Foundation Grants 3778 (to M.J.F.) and 495.01 (to S.W.C.); and by the NSF Center for Microbiological Oceanography Research and Education (S.W.C.). R.B. dedicates this work to the memory of Harold Morowitz, a pioneer in the study of biological energy flow, whose theories and approaches to studying life were central in shaping the core ideas of this work.

8. Morowitz HJ, Raether in the study of biological energy flow, whose theories and approach to studying life were central in shaping the core ideas of this work.


Supporting Information for
“Metabolic evolution and the self-organization of ecosystems”

R. Braakman*, M.J. Follows, S.W. Chisholm**
*rogierbraakman@gmail.com, **chisholm@mit.edu

SI Text

Glossary of terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a^0$</td>
<td>Specific nutrient affinity</td>
<td>L (g dry weight)$^{-1}$ time$^{-1}$</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Fraction of the cellular carbon flux density that is excreted</td>
<td>unitless, ranges between 0 and 1</td>
</tr>
<tr>
<td>$# C / # e$</td>
<td>Efficiency of carbon fixation</td>
<td>carbon atoms electron$^{-1}$</td>
</tr>
<tr>
<td>$\Delta_r G$</td>
<td>Free energy of reaction</td>
<td>kJ mol$^{-1}$</td>
</tr>
<tr>
<td>$[E]$</td>
<td>Enzyme concentration</td>
<td>moles of enzyme (g dry weight)$^{-1}$</td>
</tr>
<tr>
<td>$F$</td>
<td>Faraday constant</td>
<td>$9.6485 \times 10^4$ coulombs mol$^{-1}$</td>
</tr>
<tr>
<td>$l$</td>
<td>Light intensity</td>
<td>µmol photons m$^{-2}$ time$^{-1}$</td>
</tr>
<tr>
<td>$K_{M,n}$</td>
<td>Whole-cell Michaelis constant</td>
<td>moles of n L$^{-1}$</td>
</tr>
<tr>
<td>$k^+$</td>
<td>Rate constant</td>
<td>moles of n (moles of enzyme)$^{-1}$ time$^{-1}$</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Specific growth rate</td>
<td>time$^{-1}$</td>
</tr>
<tr>
<td>$\mu_{max}$</td>
<td>Maximal intrinsic growth rate</td>
<td>time$^{-1}$</td>
</tr>
<tr>
<td>$[n]_0$</td>
<td>Environmental concentration of nutrient $n$</td>
<td>moles of n L$^{-1}$</td>
</tr>
<tr>
<td>$[n]^*$</td>
<td>Minimal subsistence concentration of nutrient $n$</td>
<td>moles of n L$^{-1}$</td>
</tr>
<tr>
<td>$P_n$</td>
<td>Membrane permeability of nutrient $n$</td>
<td>µm time$^{-1}$</td>
</tr>
<tr>
<td>$\Phi_{PSII}$</td>
<td>Quantum efficiency of PSII</td>
<td>electrons photon$^{-1}$</td>
</tr>
<tr>
<td>$\Delta \psi$</td>
<td>Membrane potential</td>
<td>mV</td>
</tr>
<tr>
<td>$Q_n$</td>
<td>Mass-normalized cell quota of nutrient $n$</td>
<td>µm$^{-1}$</td>
</tr>
<tr>
<td>$R$</td>
<td>Gas constant</td>
<td>J K$^{-1}$ mol$^{-1}$</td>
</tr>
<tr>
<td>$r_{cell}$</td>
<td>Cell radius</td>
<td>µm</td>
</tr>
<tr>
<td>$\sigma_{PSII}$</td>
<td>Mass-normalized whole-cell PSII absorption cross-section</td>
<td>m$^2$ of PSII (g dry weight)$^{-1}$</td>
</tr>
<tr>
<td>$T$</td>
<td>Temperature</td>
<td>K</td>
</tr>
<tr>
<td>$v_e$</td>
<td>Cellular electron flux density</td>
<td>µmol electrons (g dry weight)$^{-1}$ time$^{-1}$</td>
</tr>
<tr>
<td>$v_n$</td>
<td>Nutrient uptake flux density</td>
<td>moles of n (g dry weight)$^{-1}$ time$^{-1}$</td>
</tr>
<tr>
<td>$V_{cyt}$</td>
<td>Mass-normalized cytoplasmic volume</td>
<td>mL (g dry weight)$^{-1}$</td>
</tr>
<tr>
<td>$V_{max}$</td>
<td>Maximal nutrient uptake flux density</td>
<td>moles of n (g dry weight)$^{-1}$ time$^{-1}$</td>
</tr>
<tr>
<td>$Z$</td>
<td>Nutrient charge</td>
<td>unitless</td>
</tr>
</tbody>
</table>

Hypotheses for benefits of draining excess reducing power into the environment

Studies of oceanic phytoplankton have led to several hypotheses of why cells may drain excess reducing power into the environment rather than lowering their photosynthetic electron flux. One suggestion is that slow, nutrient-limited growth increases the need for ATP-heavy maintenance metabolism relative to more NAD(P)H-heavy de novo biosynthesis, and that the water-water cycle mediated by PTOX allows cells to increase their relative ATP/NAD(P)H supply [1]. Another argument is that low iron concentrations in the open ocean promote the more iron-efficient modified electron flow facilitated by PTOX [2-4], which is consistent with the loss of several iron-sulfur enzymes in the rest of Prochlorococcus’ metabolic core (Fig. S1). However, neither hypothesis clarifies why Prochlorococcus should acquire ATP-consuming pathways involved in the excretion of organic carbon prior to innovations in the electron transfer chain that would appear most effective at increasing both iron efficiency and the relative supply of ATP.

Other hypotheses for why a photoautotrophic cell might drain excess reducing power into the environment come from studies on the excretion of organic carbon by heterotrophic microbes. One proposes that excreting organic carbon facilitates a less efficient, but intrinsically faster, mode of ATP production that requires less protein, allowing cells with an abundant energy supply to achieve faster growth by allocating more protein to biosynthesis [5-7]. This is consistent with Prochlorococcus using the Entner-Doudoroff (ED) rather than the Emden-Meyerhoff-Parnas (EMP) variant of glycolysis to catabolize glucose as a supplementary energy source.
alongside photosynthesis [8], since ED glycolysis has a higher intrinsic rate and lower protein cost than EMP glycolysis [6]. However, Prochlorococcus is characterized by a slow growth rate [9,10], and it is unclear why its cells should supplement phototrophy with chemoheterotrophy if the energy supply outpaces the nutrient supply (Fig. 1 in the main text). It is similarly unclear how excreting organic carbon would lower the protein requirement of ATP production during photosynthesis. Another suggestion is that excreting organic carbon and other forms of "energy spilling" (e.g., various enzymatic futile cycles) allow nutrient-limited cells to maintain high energy throughput of their metabolism so they can rapidly 'reignite' growth once limitation is lifted [11,12]. However, while the deeper-branching Synechococcus is found throughout the surface oceans and can exhibit rapid blooms in response to nutrient influxes, particularly in coastal regions [13,14], Prochlorococcus is restricted to the low nutrient oligotrophic oceans [15-17] and maintains a fairly steady growth rate [18]. Thus, while various hypotheses provide clues about relevant driving forces, they leave us without a completely satisfying explanation for the metabolic evolution of Prochlorococcus (Fig. 3 in the main text and surrounding discussion).

**Nutrient uptake and nutrient affinity**

Evidence suggests that the increasing excretion of organic carbon over the course of evolution in Prochlorococcus is part of a general trend that increases the cellular metabolic rate, and that this may benefit the uptake of limiting nutrients (see main text for discussion). To understand how metabolic rate and nutrient uptake are related, we begin by examining basic principles of nutrient uptake. We can consider nutrient uptake by a cell as the reaction \( n \rightleftharpoons n_v \) where \( n \) is the nutrient in the environment at concentration \([n]\) and \( n_v \) is nutrient internal to the cell at concentration \([n]_i\). The flux density of nutrient uptake \( v_n \) is given by the standard Michaelis-Menten equation:

\[
v_n = \frac{V_{max}[n]}{K_{M,n} + [n]} \tag{1}
\]

where \( V_{max} \) is the maximal saturated uptake rate of the transport reaction and \( K_{M,n} \) is the Michaelis constant. Under extreme nutrient-limitation \([n] \ll K_{M,n}\) Eq. 1 simplifies to [19]:

\[
v_n = V_{max}[n]/K_{M,n} = a_n^0[n] \tag{2}
\]

where \( a_n^0 \) is the specific nutrient affinity. The relationship between \( v_n, a_n^0, V_{max} \) and \( K_{M,n} \) is shown visually in Fig. S3. To understand how cellular growth is affected by nutrient-limitation, we can rewrite Eq. 2 as:

\[
[n] = v_n/a_n^0 = v_n K_{M,n}/V_{max} \tag{3}
\]

The minimal subsistence nutrient concentration \([n]^*\) is the lowest steady-state value of \([n]\) at which cells can achieve net positive growth. Thus, cells can lower their \([n]^*\) through two basic adaptations: increasing their specific affinity \( a_n^0 \) or decreasing their required nutrient flux density \( v_n \) (Eq. 3 and Fig. S3). Since the required nutrient flux density is given by \( v_n = \mu Q_n \) (Eq. 2 in the main text), where \( \mu \) is the specific growth rate and \( Q_n \) is the mass-normalized cellular quota of nutrient \( n \), cells can therefore adapt to nutrient-limitation by lowering \( \mu \) and/or \( Q_n \) (blue arrows in Fig. S3). Alternatively, since the nutrient affinity is equal to \( V_{max}/K_{M,n} \), cells can adapt to nutrient-limitation by increasing \( V_{max} \) and/or decreasing \( K_{M,n} \), which increases \( v_n \) (and thereby \( \mu \)) at a given nutrient concentration (red arrows in Fig. S3). It is the selection for an increasing \( V_{max} \) (and thus \( a_n^0 \)) that we argue explains the benefit of increasing metabolic rate under nutrient-limitation, as we discuss next.

**Relating nutrient affinity and metabolic rate**

To understand how metabolic rate affects \( V_{max} \) and thereby the minimal subsistence nutrient concentration of cells (Eq. 3), we perform a thought experiment using basic principles of reaction kinetics. The scenario outlined above considers only the initial uptake reaction \( n \rightleftharpoons n_v \), but in reality that reaction is part of a large network of reactions (metabolism) that collectively processes the entire conversion from external nutrient to biomass. In principle cells can thus increase the rate of nutrient uptake by increasing the rate of the
downstream metabolic network [19,20], which can be further understood by considering a simplified metabolic network consisting of two reversible reactions:

\[ n \rightleftharpoons n_i \]  
\[ n_i \rightleftharpoons M \] (4) (5)

The reaction rate of reaction 5 (the “downstream metabolic network” that converts nutrient \( n \) to biomass building block \( M \)) is given by the standard reversible Michaelis-Menten equation:

\[ v_{n_i \rightarrow M} = v^+ - v^- = \frac{k^+[n_i]/K_{M,n} - k^-[M]/K_{M,M}}{1 + [n_i]/K_{M,n} + [M]/K_{M,M}} \] (6)

Where \( v^+ \) and \( v^- \) are the forward and backward rates, \( k^+ \) and \( k^- \) are the forward and backward rate constants, \([E]\) is the enzyme concentration and \([M]\) is the concentrations of \( M \) inside the cell. The Gibbs free energy of reaction of reaction 5 is given by [21,22]:

\[ \Delta_r G = -RT \ln \left( \frac{[n_i]K_{M,M}K_{eq}}{[M]K_{M,n}} \right) \] (7)

Where \( R \) is the gas constant, \( T \) is the temperature, and \( K_{eq} = k^+/k^- \) is the equilibrium constant. If we assume that strong environmental nutrient-limitation leads to strong nutrient-limitation of the internal assimilation reaction ([\( n_i \] << \( K_{M,n} \)) and we make the additional simplifying assumption that enzyme active sites for the backward reaction are also far from saturated ([\( M \] << \( K_{M,M} \)) we can rewrite Eq. 6 to [21-23]:

\[ v_{n_i \rightarrow M} = \frac{k^+[n_i]}{K_{M,n}} \left( 1 - e^{\Delta_r G/RT} \right) \] (8)

Eq. 8 allows us to consider how increasing the metabolic rate of cells impacts the nutrient uptake rate. In a closed system, reactions will tend to equilibrate toward a state with no net flux and \( \Delta_r G = 0 \) [24]. A given starting ratio \([n_i]/[M]\) will thus undergo a more significant decrease for reactions with an intrinsically more negative \( \Delta_r G \) (Eq. 7). This same principle holds in open systems, except that there equilibrium is prevented by a steady flux of \( n \) in and \( M \) out of the system. Thus, in an open system, a reaction with a more negative \( \Delta_r G \) (e.g. due to a high value of \( K_{eq} \), or coupling the reaction to a more exothermic second reaction) will have a higher intrinsic rate (Eq. 8) and will settle into a steady state with a lower ratio \([n_i]/[M]\) (Eq. 7). But, because the rate of reaction 4 also depends on \([n_i]\), a more negative \( \Delta_r G \) for reaction 5 thereby creates a “pull” on reaction 4 by making its \( \Delta_r G \) more negative, thus increasing the rate of nutrient uptake into the cell and in turn the overall flux from \( n \) to \( M \). Comparing the forms of Eq. 8 and Eq. 2, we can see that for a reversible system, \( a_n^0 \) and \( V_{max} \) are given by:

\[ a_n^0 = \frac{[E]k^+}{K_{M,n}} \left( 1 - e^{\Delta_r G/RT} \right) \] (9)

\[ V_{max} = [E]k^+ \left( 1 - e^{\Delta_r G/RT} \right) \] (10)

In real systems, nutrient transport (rxn 4) is often an active process. Indeed, active transport becomes essential under nutrient-limitation as the following examples show, which thus slightly modifies the considerations above. Active transport drives up \([n_i]\) relative to \([n]\), thereby driving forward downstream metabolism (rxn 5) without the need to increase the energetic driving force on the latter. However, active transport also requires an energetic driving force, which scales with the natural log of the ratio of internal and external nutrient concentrations [25]. Active transport thus follows the same basic principle as stimulating passive uptake by energetically driving downstream metabolism, with the biggest difference being where in the network the driving force is applied. Moreover, the energetic driving of uptake and of downstream reactions can work in tandem to increase the overall rate of nutrient assimilation. This is illustrated using the Gibbs free energy for the nutrient transport reaction (rxn 4), which is given by [25]:
\[ \Delta_r G = RT \ln \left( \frac{[n]_{i,\text{in}}}{[n]_{i,\text{out}}} \right) + ZF \Delta \psi \]  

(11)

where \( Z \) is the unit charge of the transported nutrient, \( F \) is the Faraday constant and \( \Delta \psi \) is the membrane potential (mV). Thus, internally accumulating nutrients (first term in right hand side of Eq. 11) to commonly assumed physiological concentrations of 1 mM from near-vanishing background levels (<0.1 mM), as is the case for N, P & Fe in much of the oligotrophic oceans [26], carries a free energy cost \( \Delta_r G > 40 \) kJ/mol. This cost is then combined with the free energy cost of transport against the electric gradient (second term in right hand side of Eq. 11) to give the total free energy cost. For example, assuming a typical bacterial membrane potential of 120 mV (negative inside) [25], the total free energy cost of transporting phosphate (HPO\(_4^{2-}\); \( Z = -2 \) in seawater of pH = 8) is \( \Delta_r G > 63 \) kJ/mol. By comparison, at pH=7.5 and an ionic strength of 0.5 M the hydrolysis of ATP (ATP + H\(_2\)O \( \rightarrow \) ADP + P\(_i\)) has a free energy of reaction of \( \Delta_r G = -44.8 \) kJ/mol if all reactants have a 1 mM concentration [27], which is thus insufficient to allow any significant phosphate uptake in the oligotrophic oceans using the assumptions above. In the same scenario, accumulating phosphate to only 0.01 mM instead of 1 mM lowers the total free energy cost of transport from to \( \Delta_r G > 52 \) kJ/mol, while increasing the ATP/ADP ratio to 10 (in addition to lowering \( [P_i] \) to 0.01 mM) lowers the free energy of ATP hydrolysis to \( \Delta_r G = -61.9 \) kJ/mol [27], together allowing phosphate transport to proceed in the net forward direction under conditions like those in the oligotrophic oceans.

Further, applying an energetic driving force solely on the uptake reaction itself may not always be the most efficient solution. For example, under N-limitation the uptake and assimilation of ammonia (NH\(_4^+\)) can lead to a futile cycle, as its conjugate base (NH\(_3\)) diffuses passively out of the cell, and this loss rate increases with the concentration gradient [28]. It is therefore thought that N-limited cells will pool the internal NH\(_4^+\) concentration to the minimal viable value to limit futile cycling [29], which is helped by increasing the energetic driving of downstream reactions. This can be understood by examining the Glutamine Synthetase reaction (Glutamate + NH\(_3\) + ATP \( \rightarrow \) Glutamine + ADP + P\(_i\)), the central highway for nitrogen into metabolism [25]. At pH=7.5 and an ionic strength of 0.5 M and again using commonly assumed physiological concentrations of 1 mM for all reactants, glutamine synthesis has a \( \Delta_r G \) of \(-15.4 \) kJ/mol [27], making it essentially irreversible. However, an internal NH\(_4^+\) concentration of 1 mM would lead to high levels of futile transport cycling as can be seen by comparing the fluxes for nitrogen assimilation (Eq. 2 in the main text) and diffusive NH\(_3\) loss [29]:

\[ v_N = \mu Q_N \]  

(12)

\[ v_{N,\text{loss}} = 3P_{NH_3}[\text{NH}_3]_{\text{in}} - [\text{NH}_3]_{\text{ex}}V_{\text{cyt}}/r_{\text{cell}} \]  

(13)

where \( P_{NH_3} \) is the membrane permeability for NH\(_3\) (\( \mu \text{mole time}^{-1} \)), \( [\text{NH}_3]_{\text{in,\text{out}}} \) are the internal and external concentrations of NH\(_3\) (moles L\(^{-1}\)), \( V_{\text{cyt}} \) is the mass-normalized cytoplasmic volume (mL (g of dry weight\(^{-1}\))), \( r_{\text{cell}} \) is the cell radius (\( \mu \text{m} \)), \( Q_N \) is the mass-normalized cell quota of nitrogen (moles of \( N \) (g of dry weight\(^{-1}\))) and \( \mu \) is the specific growth rate (time\(^{-1}\)). For consistency we have modified the equations of Ref. [29] to give fluxes in units of (g of dry weight\(^{-1}\)) rather than L\(^{-1}\). If we assume ammonia is actively taken up as NH\(_4^+\) to an internal concentration \( [\text{NH}_4^+]_{\text{in}} = 1 \) mM, and we further assume an internal pH of 7.5, then \( [\text{NH}_3]_{\text{in}} = 18 \) \( \mu \text{M} \) (the pKa of ammonia is 9.25 at room temperature). Further assuming a Prochlorococcus cell radius of \( r_{\text{cell}} = 0.3 \) \( \mu \text{m} \), a typical bacterial value for \( V_{\text{cyt}} \) of 2 mL (g of dry weight\(^{-1}\)) [29], and a relatively conservative \( P_{NH_3} \) of 10 \( \mu \text{mole s}^{-1} \) (values up to 2000 \( \mu \text{mole s}^{-1} \) have been measured [29]), then if \( [\text{NH}_3]_{\text{ex}} \) is near-vanishing [26] the diffusive NH\(_3\) loss would be 3.6 \( \mu \text{mole} \) (g of dry weight\(^{-1}\)) \( \text{s}^{-1} \). Meanwhile, a Prochlorococcus cell with a \( Q_N \) of \( \sim 7 \) mmole \( N \) (g of dry weight\(^{-1}\)) (based on an N content of \( \sim 0.1 \) g (g of dry weight\(^{-1}\)) [30] and a molar mass of \( N \) of 14 g mole\(^{-1}\)), growing near its maximal growth rate of \( \sim 0.7 \) day\(^{-1}\) would have a N assimilation flux of 57 mmole (g of dry weight\(^{-1}\)) \( \text{s}^{-1} \). The loss flux of ammonia through passive diffusion would thus be two orders of magnitude higher than the assimilation flux, leading to very high levels of futile transport cycling.

Keeping all other reactant concentrations fixed and lowering \( [\text{NH}_4^+]_{\text{in}} \) to 1 \( \mu \text{M} \) and thereby \( [\text{NH}_3]_{\text{in}} \) to 0.018 \( \mu \text{M} \) would lower the passive diffusive loss rate of NH\(_3\) to 3.6 nmole (g of dry weight\(^{-1}\)) \( \text{s}^{-1} \), well below the assimilation flux calculated above. However, the net flux of the glutamine synthetase reaction would run in
reverse with a $\Delta_rG$ of +1.7 kJ/mol [27], creating a kinetic bottleneck in the assimilation of ammonia. As before, increasing the $ATP/ADP$ ratio from 1 to 10 while lowering $[P_i]$ to 0.01 mM would lower the $\Delta_rG$ below 0 and drive glutamine synthesis, and thereby ammonia assimilation, forward in the oligotrophic oceans.

We note that cells using nitrate ($NO_3^-$) as their nitrogen source face an intermediary scenario compared to the examples of phosphate or ammonia uptake. Nitrate has a charge $Z = -1$ at pH = 8, resulting in a free energy cost of uptake $\Delta_rG > 51.5$ kJ/mol when accumulating it to an internal concentration of 1 mM from external levels of ≤0.1 mM [26]. This is lower than the free-energy cost of uptake for phosphate (which has a higher cost due to its charge of $Z = -2$), thus requiring less energetic driving than the latter. However, nitrate uptake again faces the costs of futile transport cycling, since it is first reduced to $NH_4^+$ before being assimilated via glutamine synthetase. Passive loss of $NH_3$ would thus also be significant during growth on nitrate in the oligotrophic oceans, unless $[NH_4^+]_{in}$ is poised at the minimal viable value [29], again requiring greater energetic driving (through an increased $ATP/(ADP \times P)$ ratio) of glutamine synthesis.

The examples above illustrate the importance of increasing the $ATP/(ADP \times P)$ ratio to drive forward nutrient assimilation in the oligotrophic oceans. However, increasing the $ATP/(ADP \times P)$ ratio also makes ATP synthesis more thermodynamically uphill, which therefore requires a greater proton motive force, and thus for photosynthetic cells a greater photosynthetic electron flux [20].

Increasing metabolic rate and the benefits of excreting organic carbon

The preceding sets the stage for understanding why excreting organic carbon may be beneficial to cells with a plentiful energy supply. Cells can drive up their $ATP/ADP$ ratio by increasing the ATP supply rate, but to maintain steady-state balance must then also increase their ATP consumption rate [11,20]. A major sink of ATP in a photosynthetic cell is CO$_2$-fixation. But, cells are limited in the carbon-flux density then can accommodate, and this upper limit is proportional to both the maximal cellular carbon density (i.e. the maximal $Q_C$) and the growth rate $\mu$. This upper limit on the carbon flux density is thus lower under strong nutrient-limitation, which favors slow growth (Eq. 3), but in principle exists for any $\mu$. We thus argue that excretion of organic carbon allows cells to increase their metabolic rate above the point where the carbon flux exceeds the growth requirement. This can be further understood by expanding Eq. 3 using several assumptions. We assume that under nutrient limitations cells take up limiting nutrient $n$ with maximal efficiency, but can be inefficient with all non-limiting resources. In other words, if we further assume that biomass has an elemental stoichiometry $Q_C/Q_n$, then cells must fix CO$_2$ at a rate $at$ least equivalent to $v_n(Q_C/Q_n)$, but can in principle fix carbon at higher rates and excrete the excess carbon flux. We further assume that the fraction of fixed carbon that is excreted or respired is $\beta$, which falls in a range (0<$\beta$<1), and that the efficiency of CO$_2$-fixation is $#C/#e$. The carbon/electron flux ratio $#C/#e$ depends on the oxidation state of the carbon source, biomass and excreted carbon. For photosynthetic cells fixing CO$_2$ and producing biomass and excreted carbon at an average oxidation state of CH$_2$O, $#C/#e$ is 1/4 (i.e. CO$_2$ + 4 H$^+$ + 4 e$^-$$\Rightarrow$ CH$_2$O + H$_2$O). Photoprotective mechanisms such as PTOX (Fig. S1) that direct electron flux out of the electron transport chain prior to carbon-fixation act to lower $#C/#e$. Under these assumptions, the nutrient, carbon and electron flux densities of nutrient-limited cells are related through:

$$v_n = \frac{Q_n}{Q_C} v_e (#C/#e)(1 - \beta)$$

Finally, we can combine Equations 3, 9 and 14 to give:

$$[n] = \frac{K_{M,n}}{[E]k^+} \times \frac{v_n}{1 - e^{\Delta_rG/RT}} = \frac{K_{M,n}}{[E]k^+} \times \frac{Q_n}{Q_C} v_e (#C/#e)(1 - \beta)$$

Eq. 15 describes how selection to lower $[n]^*$ shapes cellular features. As discussed earlier, selection to lower $[n]^*$ favors decreases in $v_n$ and $\Delta_rG$ (Equations 3, 9 and 10), and as we just argued decreases in $\Delta_rG$ are achieved through an increased electron flux. But an increased electron flux also leads to an increased carbon flux, which in turn acts to drive up $[n]^*$. Therefore, selection to lower $[n]^*$ should simultaneously favor an increasing $v_e/v_n$ and an increasing excretion of organic carbon, exactly as we observe for the evolution of Prochlorococcus. See main text for further discussion.
References

27. Calculations were done using the Equilibrator (equilibrator.weizmann.ac.il); Noor, Elad et al. "An integrated open framework for thermodynamics of reactions that combines accuracy and coverage." Bioinformatics 28.15 (2012): 2037-2044.
Figure S1. Details of the changes in the photosynthetic machinery (left panel) and metabolic core (right panel) of Prochlorococcus during its evolutionary divergence (Fig. 3 of main text) from a universal cyanobacterium (black). The photosystems of Prochlorococcus (purple) have an altered set of accessory pigments [1-4] and the PSII:Cyt b$_6$f:PSI ratio increases along its phylogeny [5,6]. The photosynthetic electron transport chain of the LL1 and HL clades further contains the Plastoquinol Terminal Oxidase (PTOX) (Table S1) as indicated by the orange box. Within core carbohydrate metabolism, key genes are gained and lost as indicated by blue and yellow lines (Table S1). Iron-sulfur cluster enzymes are highlighted in brown. Red circles indicate putative redox safety valves, including those excreting pyruvate, glycolate and citrate or isocitrate, while the blue circle indicates a putative uptake pathway for malate. We suspect citrate rather than isocitrate is excreted because under iron-limitation or oxidative stress aconitase (two-step reaction 17 converting citrate to isocitrate) is post-translationally converted to the iron regulatory protein (IRP) [7], which may lead to accumulation of citrate. Enzyme names of numbered reactions are shown in Table S1 and Fig. S2.

Figure S2. Genomic evolution and expression of transport pathways in Prochlorococcus. Chromosomal location of key metabolic genes and ABC transporters are shown along the Prochlorococcus phylogeny (left panel). In some plots only Cyanobium gracile and marine picocyanobacteria are shown because the corresponding metabolic genes are absent in other cyanobacteria (Table S1). Color-coded names of genes are shown on the right, and gene numbers are as in Fig. S1. Changes in the relative transcript levels of genes (log-fold changes in mRNA number) in Prochlorococcus MED4 (HLI clade) over a diel L:D cycle [redrawn from Ref. 1] are shown on the right.

Figure S3. Kinetics of nutrient uptake (based on Ref. [1]). The Y-axis is the nutrient uptake rate $\nu_n$ and the X-axis is nutrient concentration $[n]$. Dotted lines indicate the specific affinity $a_0^n$ and the upper (diffusion) limit on the specific affinity $a_{\text{max}}^n$. Dashed lines indicate the maximal saturated uptake rate $V_{\text{max}}$. The Michaelis constant $K_{M,n}$ is the nutrient concentration at which $\nu_n = 1/2 V_{\text{max}}$.

Equations show the relationships between $\nu_n, a_0^n, V_{\text{max}}$ and $K_{M,n}$, and between $\nu_n, \mu$ and $Q_n$. For a fixed nutrient affinity, cells can lower their minimal subsistence nutrient concentration by lowering their required nutrient flux $\nu_n$ (blue arrows). Cells can increase $\nu_n$ at any nutrient concentration by increasing $a_0^n$ (red arrows), which is achieved by increasing $V_{\text{max}}$ and/or decreasing $K_{M,n}$ (green arrows), providing a second route to lowering the minimal subsistence nutrient concentration.

Figure S4. Details of the changes in the core metabolism of SAR11 over the course of its evolution. Key genes are gained and lost as indicated by the blue and yellow lines (Table S2). Red dots indicate putative redox outlets that excrete malate and citrate or isocitrate, and blue dots indicate putative uptake pathways for glycolate and pyruvate. As in Prochlorococcus we suspect that citrate rather than isocitrate is excreted because aconitase doubles as the iron-regulatory protein [1] (Fig. S1). Enzyme names of numbered reactions are shown in Table S2 and Figure S5.

Figure S5. Genomic evolution of transport pathways in SAR11. Chromosomal location of key metabolic genes and ABC transporters are shown along the Prochlorococcus phylogeny. In some plots SAR11 group V was excluded because it lacks the relevant metabolic genes (Table S2). Color-coded names of genes are shown below each plot, and gene numbers are as in Fig. S2. Here a variety of transporters are involved. The drug/metabolite (DME) transporter family (pyruvate pathway) is implicated in both export [1] and uptake [2], and in shuttling electrons from the cytoplasm to the periplasm under oxidative stress [3,4]. Major Facilitator Superfamily (MFS) uptake transporters and Tripartite ATP-independent periplasmic (TRAP) uptake transporters (glycolate pathway) both depend on proton motive force rather than ATP-hydrolysis to drive transport [5,6]. The SecA transporter is part of the Sec protein translocation system [7], which in E. coli contains a plug that prevents transfer of small metabolites [8], but in eukaryotes it is permeable to them [9-11]. We also observe an ABC transporter exclusive to the IA clade (malate pathway).

### SI Tables

**Table S1.** Distribution of core metabolic genes across cyanobacteria highlights the remodeling of the metabolic core (Fig. S1 & Fig. 3 in the main text) during the evolution of marine picocyanobacteria as they diverged from their ancestors

<table>
<thead>
<tr>
<th>Clade</th>
<th>Number of Genomes</th>
<th>Reaction/enzyme* (genomic abundance)</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10/15</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyanobacteria</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All other groups</td>
<td>53</td>
<td></td>
<td>12</td>
<td>44</td>
<td>53</td>
<td>53</td>
<td>0</td>
<td>52^g</td>
<td>43</td>
<td>51</td>
<td>52</td>
<td>48</td>
<td>34</td>
<td>53</td>
<td>23</td>
<td>17</td>
</tr>
<tr>
<td><em>Synechococcus</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>elongatus</em></td>
<td>2</td>
<td></td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td><em>Cyanobium</em></td>
<td></td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Marine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Synechococcus</em></td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-cluster 5.2</td>
<td>3</td>
<td></td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1^d</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Sub-cluster 5.3</td>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sub-cluster 5.1B</td>
<td>5</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Sub-cluster 5.1A</td>
<td>6</td>
<td></td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>4^d</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Prochlorococcus</td>
<td>41</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLIV</td>
<td>5</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5^d</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5^c</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LLI-III</td>
<td>10</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>6^d</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>10^c</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LLI</td>
<td>4</td>
<td></td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4^c</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HLI</td>
<td>3</td>
<td></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>HLII</td>
<td>19</td>
<td></td>
<td>19</td>
<td>0</td>
<td>0</td>
<td>19</td>
<td>1^d</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

*Enzyme names (see Fig. S1):* 1 – Plastoquinol Terminal Oxidase (PTOX), 2 – 6-phosphofructokinase, 4 – phosphoenolpyruvate synthase, 5 – malic enzyme, 6 – malate:quinone oxidoreductase (MQO), 7 – succinate dehydrogenase (A,B,C subunits), 8 – succinyl-CoA synthetase, 9 – succinate-semialdehyde dehydrogenase, 10 – 2-oxoglutarate decarboxylase, 11 – acetate kinase, 12 – acetylphosphatase, 13 – 2-glycolate oxidase (D,E,F subunits), 14 – tartronate semialdehyde reductase, 15 – glyoxylate carboligase, 18 – catalase (KatG)

**Notes:** ^a During our analysis we realized that previous experimental studies on the TCA cycle and photorepiration in cyanobacteria implied the same gene as catalyzing reactions 10 & 15 [1,2], suggesting that both functions are performed by a promiscuous enzyme. ^b It has previously been noted that in marine picocyanobacteria carrying this 3-subunit enzyme it is a homologue acquired from proteobacteria, that potentially catalyzes the reaction in the opposite direction [3,4]. ^c The presence of this gene in many strains lacking the gene for reaction 15, as well as its absence in freshwater picocyanobacteria, suggests loss prior to reacquisition and involvement in a different function in marine picocyanobacteria.

2. Eisenhut M (2008) The photorepiratory glycolate metabolism is essential for cyanobacteria and might have been conveyed endosymbiontically to plants. Proc Natl Acad Sci USA 105(44):17199-17204.
Table S2. Distribution of core metabolic genes across SAR11 sub-groups highlights the remodeling of
the metabolic core (Fig. S4 & Fig. 5 in the main text) during the evolution of the groups dominating the
oligotrophic surface oceans as they diverged from their ancestors

<table>
<thead>
<tr>
<th>SAR11 Sub-group</th>
<th>Number of Genomes</th>
<th>Reaction/enzyme* (genomic abundance)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>18</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td>2</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td>1</td>
<td></td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>IA.1/IA.2</td>
<td>8</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>IA.3</td>
<td>4</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

*Enzyme names (see Fig. S4): 1 – 6-phosphofructokinase, 2 – pyruvate kinase, 3 – pyruvate carboxylase, 4 –
phosphoenolpyruvate carboxylase, 5 – phosphoenolpyruvate synthase, 6 – malic enzyme, 8 – malate synthase, 9 – glycolate
oxidase, 10 – isocitrate lyase, 11 – glucose dehydrogenase, 12 – gluconolactonase, 13 – repressor, ORF, kinase (ORK), 14 – 6-
phosphogluconate dehydratase, 15 – glucose/ribitol dehydrogenase, 16 – fumarylacetacetate hydrolase, 18 – catalase
(KatG).

Notes: * Genes postulated to be involved in an Entner-Doudoroff glycolytic variant pathway [1].

SAR11 genomes is positively correlated with ocean productivity. Environ Microbiol 12(2):490-500.