Global niche partitioning of purine and pyrimidine cross-feeding among ocean microbes

R. Braakman[1,*], B. Satinsky[2,3], T.J. O'Keefe[2,4], K. Longnecker[5], S.L Hogle[2,6], J.W. Becker[2,7], K. Dooley[2,8], A. Arellano[2,9], M.C. Kido Soule[5], E.B. Kujawinski[5], S.W. Chisholm[2,10]

[1] Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA
[2] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA
[3] current address: Dyno Therapeutics, Watertown, MA, USA
[4] current address: Institute of Marine Sciences, University of North Carolina at Chapel Hill, Morehead City, NC, USA
[5] Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA
[6] current address: Department of Biology, University of Turku, Turku, Finland
[7] current address: Department of Science, Alvernia University, Reading, PA, USA
[8] current address: Renewable Resources and Enabling Sciences Center, National Renewable Energy Laboratory, Golden, CO, USA
[9] current address: Department of Bacteriology, University of Wisconsin-Madison, Madison, WI, USA
[10] Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

*Corresponding author: braakman@mit.edu

Abstract

Characterizing the diversity of cross-feeding pathways in ocean microbes illuminates forces shaping co-evolution, ecosystem self-assembly and carbon cycling. Here we uncover a purine and pyrimidine cross-feeding network in globally abundant groups. The cyanobacterium *Prochlorococcus* exudes both compound classes, which metabolic reconstructions suggest follows synchronous daily genome replication. Co-occurring heterotrophs differentiate into purine or pyrimidine specialists, or generalists, and use compounds for different purposes. The most abundant heterotroph, SAR11, is a specialist that uses purines as sources of energy, carbon and/or nitrogen, with subgroups differentiating along ocean-scale gradients in the supply of energy and nitrogen, in turn leading to putative cryptic nitrogen cycles that link many microbes. Finally, in a SAR11 subgroup that dominates where *Prochlorococcus* is abundant, adenine additions to cultures inhibit DNA synthesis, poising cells for replication. We argue this subgroup uses inferred daily pulses of adenine from *Prochlorococcus* to metabolically synchronize to the daily supply of photosynthate from surrounding phytoplankton.

Introduction

Microbial processes and interactions lie at the heart of the oceanic carbon cycle (1-3). Carbon enters the marine biosphere through photosynthesis (4,5), is transformed through metabolism (6) and viral infections and lysis (7,8), is transferred to higher trophic levels through grazing (2,9), and is ultimately broken back down and released as CO_2 through bacterial catabolism and respiration (1-3,6). An increasingly recognized link within this global cycle of organic carbon is cross-feeding, in which cells excrete specific compounds that are consumed by surrounding cells (10-13). Evidence that cross-

feeding is widespread in the ocean comes from the highly streamlined genomes and complex nutrient dependencies of many highly abundant oceanic bacteria (14-16). However, partly due to the difficulty of measuring extracellular metabolites in seawater (3), the diversity and distribution of cross-feeding pathways has not been systematically characterized, obscuring both the ecological roles they play and the forces shaping their evolution.

To explore this using a bounded model system, we have begun characterizing the production of organic carbon by the highly abundant oceanic cyanobacterium *Prochlorococcus* (17) and assessing its fate. By performing ~10-15% of oceanic CO_2-fixation (18,19), *Prochlorococcus* is a major source of organic carbon to ocean ecosystems, and over geologic timescales is thought to have evolved metabolic interdependencies with co-occurring heterotrophs (20). In recent experiments, we noticed that pyrimidines and purines are both abundant exudates in *Prochlorococcus* (21). Here we further examine the production and consumption of these compounds and uncover a globally distributed cross-feeding network involving many abundant oceanic microbes. We begin to characterize the forces driving large-scale niche partitioning within this network.

Results and Discussion

Pyrimidine and purine exudation by Prochlorococcus

To better understand organic carbon production by *Prochlorococcus* we have measured the extracellular abundance of a large suite of organic compounds in exponentially growing cultures of several strains under a range of light and nutrient conditions (21). A signal standing out is that thymidine is among the most abundant extracellular compounds across strains and culture conditions (Fig. 1, Table S1). The exudation of high levels of this nitrogen-rich pyrimidine is unexpected, since it is generally thought that the genome and proteome of *Prochlorococcus* are highly streamlined to minimize the nitrogen requirements of growth in oligotrophic waters (22,23). We calculated (Methods) that over the course of the experiment production of extracellular thymidine amounted to almost half of what was incorporated into DNA (Fig. 1, Table S1). We also detected adenine and guanine in the exudates of *Prochlorococcus*, and their levels were higher under phosphorus-limitation than in nutrient replete conditions (Fig. 1, Table S1). While we could not precisely quantify these compounds due to an extraction efficiency <1% (21), their uncorrected levels reached ~2-3% of that of thymidine, suggesting their exudation was at a similar scale (Fig. 1, Table S1).

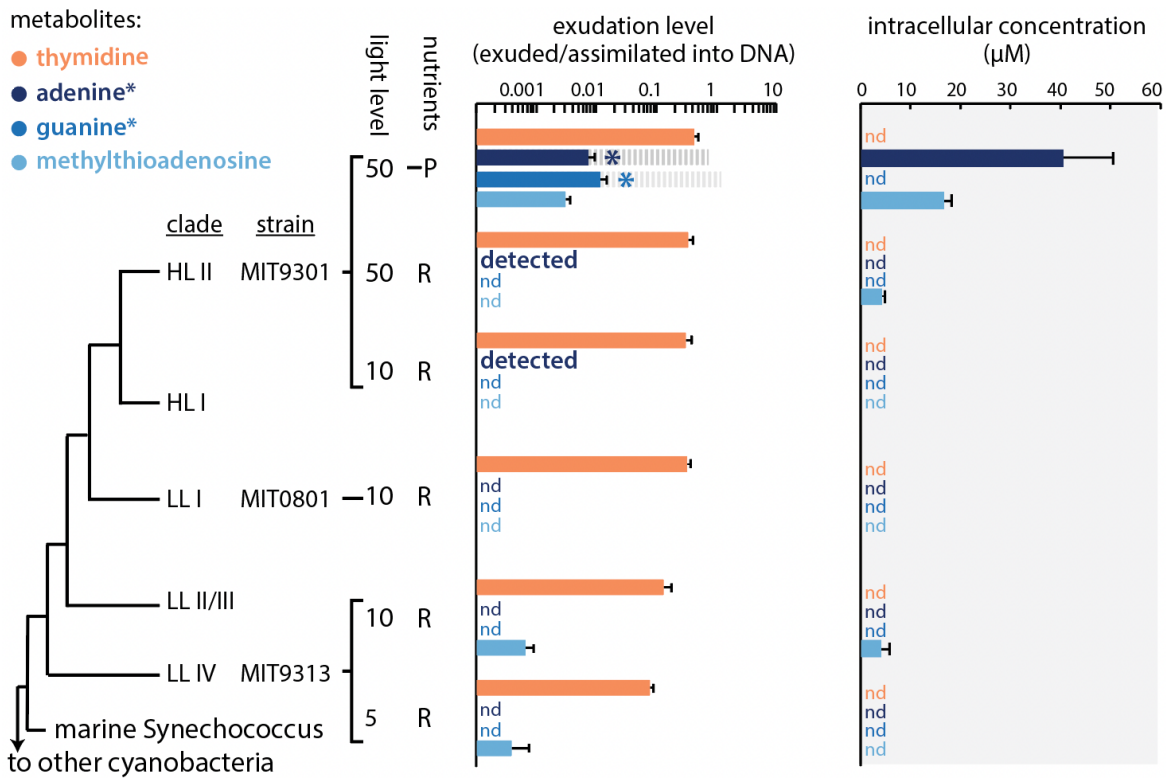
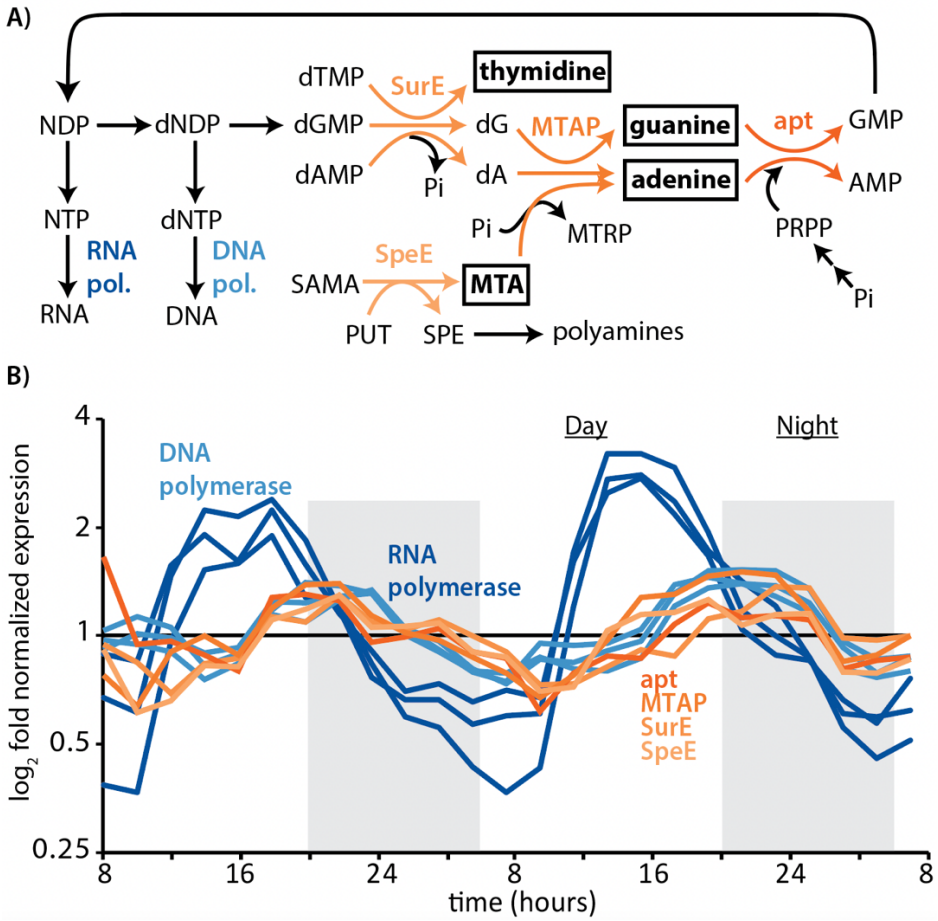


Fig. 1 Extra- and intracellular pyrimidines and purines in cultures of *Prochlorococcus*. Extracellular levels of the pyrimidine thymidine (orange), and the purines adenine, guanine and 5-methylthioadenosine (shades of blue are normalized to the total amounts of purines or pyrimidines incorporated into DNA and shown alongside cytosolic concentrations (gray shading) of these compounds, for three different strains of *Prochlorococcus* grown under different conditions. Error bars show the standard deviation between biological replicates. Light levels are in units of μmol photons m^{-2} s^{-1}, 'R' stands for batch culture growth in nutrient replete media and '–P' stands for growth under semi-continuous phosphorus limitation. Exudation levels of adenine and guanine (marked with asterisks) are uncorrected from the measured concentrations due to having an extraction efficiency of less than 1%. For a sense of scale relative to thymidine exudation levels, gray dashed bars are included at 100x the measured adenine and guanine levels (i.e. the theoretical correction factor for an extraction efficiency of 1%). Incidences in which metabolites were only detected in single replicates are denoted with 'detected', while incidences in which metabolites were not detected in any replicate are denoted with 'nd'.

To understand the excretion of thymidine, adenine and guanine, we searched for possible pathways producing and/or consuming them in *Prochlorococcus*. The only pathway we identified involving these compounds (Methods, Text S1) is a putative pathway for recycling deoxyribonucleotides, the direct precursors to DNA (Fig. 2A). There thymidine, deoxyadenosine and deoxyguanosine are generated from dTMP, dAMP and dGMP, respectively, through the action of 5'-nucleotidase (SurE) (24). Deoxyadenosine and deoxyguanosine are then further processed to adenine and guanine through the action of methylthioadenosine phosphorylase (MTAP), and then to AMP and GMP through the action of adenine phosphoribosyltransferase (apt), together forming a salvage pathway also observed in other systems (25). We could not identify other genes involved in processing thymidine, suggesting it forms an endpoint and is not further processed after the first (phosphate-salvaging) step of the pathway, perhaps because it is only involved in DNA synthesis, while deoxyadenosine and deoxyguanosine can be repurposed for ATP/GTP and RNA synthesis. This is consistent with the fact that we never measured intracellular thymidine, suggesting it is rapidly excreted after the phosphate is removed from dTMP. Original genome annotations assigned a narrower functionality to MTAP (i.e.

107 acting only on methylthioadenosine) than we hypothesize here, but this enzyme is known to act on
108 multiple substrates in other microbial systems, including as part of the pathway we identify here
109 (25,26). Enzyme multi-functionality is generally thought to be an important strategy in bacteria with
110 streamlined genomes like *Prochlorococcus* (27,28). A lack of other identifiable pathways thus leads us
111 to tentatively conclude that thymidine, adenine and guanine are all intermediates or endpoints of a
112 shared, broad specificity deoxyribonucleotide recycling pathway (Fig. 2A).
113



114
115 **Fig. 2 Putative deoxyribonucleotide recycling pathway in *Prochlorococcus*. A)** Structure of the deoxyribonucleotide
116 recycling pathway in relation to pathways for DNA and RNA synthesis and polyamine metabolism. Metabolites highlighted in
117 boxes are exuded by *Prochlorococcus* (Fig. 1). **B)** transcriptional dynamics of genes involved in the deoxyribonucleotide
118 recycling pathway, as well as DNA and RNA polymerase genes, in *Prochlorococcus* MED4 cells synchronized to a diurnal
119 light:dark cycle (29). In both panels deoxyribonucleotide recycling pathway reactions/genes are highlighted using the orange
120 spectrum, and DNA and RNA polymerase are highlighted in shades of blue. Abbreviations: NDP = nucleoside diphosphate,
121 dNDP = deoxynucleoside diphosphate, NTP = nucleoside triphosphate, dNTP = deoxynucleoside triphosphate, dTMP =
122 deoxythymidine monophosphate, dGMP = deoxyguanosine monophosphate, dAMP = deoxyadenosine monophosphate, dA =
123 deoxyadenosine, MTA = 5-methyl-thioadenosine, SAMA = S-adenosylmethioninamine, PUT = putrescine, SPE = spermidine,
124 MTRP = methyl-5-thioribose, PRPP = phosphoribose diphosphate, Pi = orthophosphate, SurE = survival protein E (5'-
125 nucleotidase), MTAP = methylthioadenosine phosphorylase, apt = adenine phosphoribosyltransferase, SpeE = spermidine
126 synthase.
127
128 To further explore the possible existence of a deoxyribonucleotide recycling pathway in
129 *Prochlorococcus* we examined existing data on the expression dynamics of the relevant genes in

transcriptional data obtained from a strain whose growth was synchronized to a diurnal light:dark cycle (29). Under such conditions, metabolic processes are segregated in time, allowing us to examine the potential role of the putative deoxyribonucleotide recycling pathway in the broader processes of growth and cellular replication. The expression of all three genes (SurE, MTAP, apt) in the putative pathway are synchronized and maximally expressed just after dusk, when genome replication is complete and cells transition to RNA synthesis (Fig. 2B), consistent with their collective functioning in deoxyribonucleotide recycling.

Properties of the deoxyribonucleotide recycling pathway are also consistent with observed metabolite dynamics in *Prochlorococcus* cells under phosphorus-limited growth conditions. That is, regenerating AMP and GMP from adenine and guanine requires phosphate (Fig. 2A), and adenine accumulates intracellularly in phosphate-limited relative to nutrient-replete cells (Fig. 1). This suggests that exudation of adenine and guanine results from the emergence of a bottleneck in the deoxyribonucleotide recycling pathway due to a depletion of intracellular phosphate. This is consistent with the observation that methylthioadenosine, a byproduct of polyamine metabolism that is recycled via the same pathway (Fig. 2A), also accumulates and is excreted under phosphate limited relative to replete conditions (Fig. 1). A recent study showed that some strains of the purple sulfur bacterium *Rhodopseudomonas palustris* exude adenine when a bottleneck arises in its intracellular salvage due to low levels of the enzyme *apt* (30), highlighting a similar principle. In contrast, the pyrimidine thymidine, which unlike purines is not internally recycled (Fig. 2), never accumulates, and its extracellular levels are not significantly different under phosphorus stress (Fig. 1). Taken together, metabolite, genomic and transcriptional data thus suggest that *Prochlorococcus* cells generate excess dNTP during genome replication and excrete whatever they do not reuse or recycle when transitioning to ribosome synthesis. This in turn implies that *Prochlorococcus* provides daily pulses of pyrimidines and purines to the surrounding ecosystem. Time-resolved quantification of metabolites in the field and in diel L:D synchronized cultures of *Prochlorococcus* are a key area of future study that will help constrain both the timing of release and the relative contributions of *Prochlorococcus* to rhythms of these substrates in the ocean ecosystem.

Niche partitioning of purine and pyrimidine usage among heterotrophic bacterioplankton

We next sought to understand the ecological consequences of purine and pyrimidine excretion by *Prochlorococcus*. To this end, we searched for purine and pyrimidine usage genes in 305 partial and complete genomes of SAR11, SAR86 and SAR116, each among the most abundant of *Prochlorococcus*' sympatric heterotrophs (16,31-33). In SAR11, the occurrence of purine usage genes reaches frequencies similar to the average genome completeness of ~72%, suggesting purine usage is a universal or nearly universal trait in this genus, whereas pyrimidine usage genes occur at frequencies of only ~3% (Fig. 3, Data S1). In contrast, in SAR86, thymidine usage genes occur at frequencies similar to the average genome completeness of ~71%, suggesting thymidine usage is a universal or nearly universal trait in the genus, whereas purine usage genes are absent (Fig. 3, Data S1). Finally, in SAR116, purine usage genes occur at frequencies of ~49%, whereas pyrimidine genes occur at frequencies of ~43% (Fig. 3, Data S1), both below the average genome completeness of ~75%. These observations suggest that resource niche partitioning within the oceanic bacterioplankton community has given rise to purine and pyrimidine specialists, as well as generalists

that use both classes of compounds, in turn driving differentiation of cross-feeding with *Prochlorococcus*.

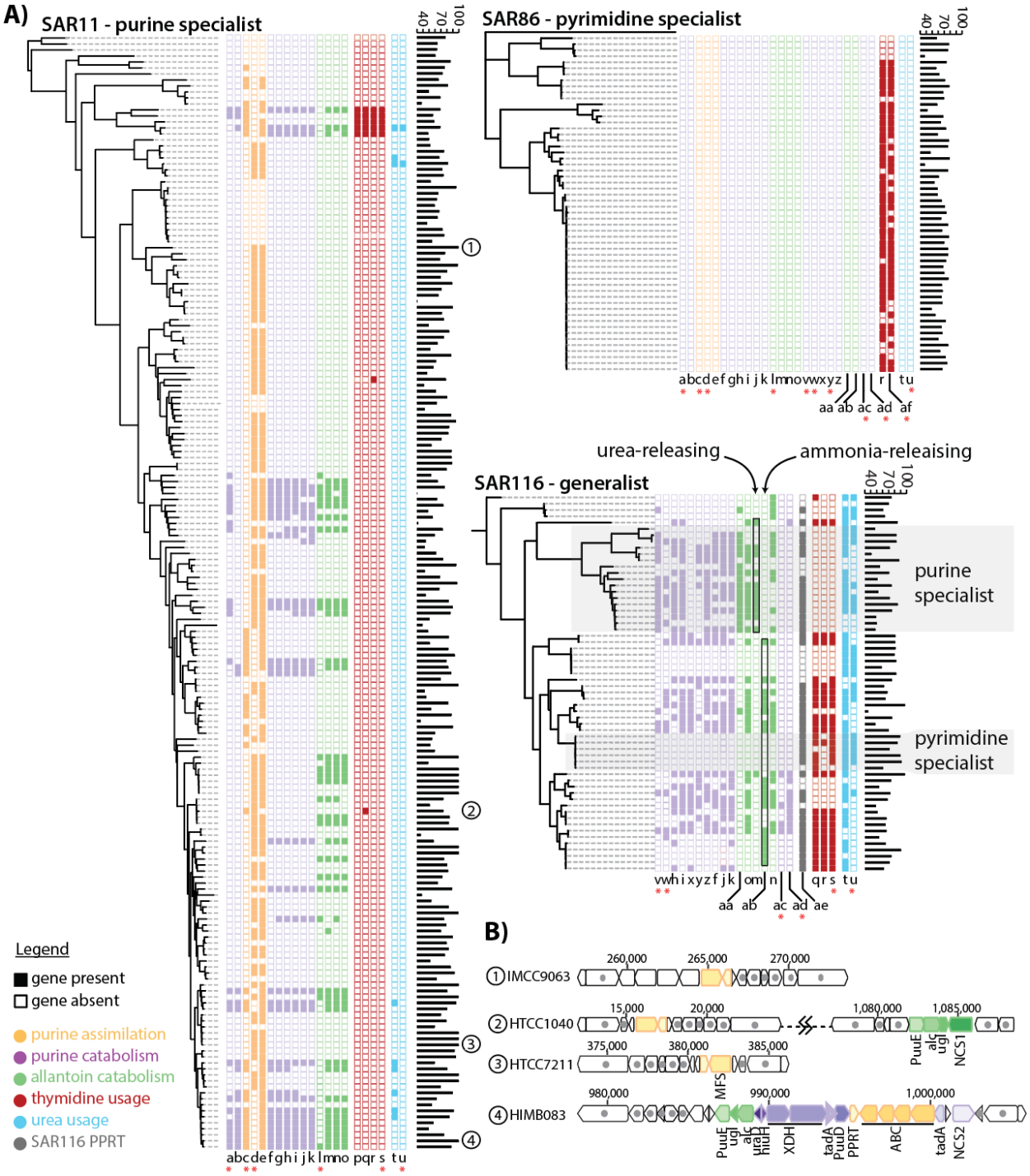


Fig. 3 Niche partitioning of purine and thymidine usage in oceanic bacterioplankton. A) Phylometabolic diagrams show presence/absence (filled/empty symbols) of genes involved in purine and pyrimidine usage in the abundant often co-occurring heterotrophic bacterioplankton groups SAR11, SAR86 and SAR116. Rows in the phylometabolic diagrams represent the metabolic gene profiles within individual genomes in the tree. Transporter genes are indicated with pink asterisks. Colors represent different functions of pathways: purine assimilation genes are in yellow, purine catabolism genes in purple, allantoin

catabolism genes in green, thymidine usage genes in dark red and urea assimilation genes in light blue. SAR116 purine phosphoribosyltransferase (PPRT), which occurs within both purine and thymidine usage gene operons (Fig. S2), is colored in dark gray. Genome completeness statistics (%) are shown as histogram next to gene profiles. **B)** Representative genome profiles of SAR11 sub-groups specializing in either purine assimilation (IMCC9063 and HTCC7211), purine assimilation and allantoin catabolism (HTCC1040), or full purine catabolism (HIMB083). Genes: a = purine NCS2 permease, b = purine deaminase (tadA), c = purine ABC transporter permease, d = purine MFS permease, e = purine phosphoribosyltransferase (PPRT), f = urate oxidase (PuuD), g = purine deaminase (tadA), h = Xanthine dehydrogenase Mo-subunit (XDH_Mo) , i = xanthine dehydrogenase FAD-subunit (XDH_FAD) , j = 5-hydroxyisourate lyase (hiuH), k = 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline decarboxylase (OHCU decarboxylase, uraD), l = allantoin NCS1 permease, m = allantoicase (alc), n = ureidoglycolate lyase (ugl) , o = allantoinase (PuuE), p = uracil phosphoribosyltransferase, q = cytidine deaminase, r = thymidine kinase, s = pyrimidine ABC transporter, t = urease, u = urea ABC transporter, v = purine ABC transporter, w = purine ABC transporter, x = XDH accessory protein, y = purine NCS2 permease, z = purine deaminase, aa = allantoin racemase, ab = ureidoglycine aminohydrolase, ac = purine TRAP transporter, ad = DMT transporter, ae = purine/pyrimidine phosphoribosyltransferase, af = CNT transporter.

Niche partitioning of purine and pyrimidine cross-feeding occurs across multiple taxonomic scales. For example, while at the genus-level SAR116 is a purine and pyrimidine generalist, it also contains subgroups that have genes for only purine or pyrimidine usage, but not both (Fig. 3, Fig. S2). In addition, different parts of the SAR116 tree contain variations in the pathway of purine catabolism that are characterized by differences in the presence or absence of allantoin racemase and the production of either ammonia or urea in the final steps of the pathway (Fig. 3, Fig. S3), as well as an association with primary, ATP-consuming transporters or with secondary, ion-gradient-driven transporters (Fig. 3).

Similarly, different SAR11 sub-groups have gene profiles suggesting they use purines for different purposes. For example, most genes making up a large operon for purine catabolism (34) occur in ~18-22% of genomes, but an uptake transporter and a purine phosphoribosyltransferase that together perform the initial assimilatory portion of the pathway occur at frequencies similar to the average genome completeness of ~72% (Fig. 3, Data S1), suggesting they are universal or nearly universal. This suggests that some SAR11 cells break purines down to their constituent components, while others use them intact. Further, an allantoin transporter occurs in ~12% of genomes (Data S1), including in at least 10 genomes that lack the full purine catabolism operon but retain genes involved in allantoin catabolism (Fig. 3), leading to an overall higher frequency of allantoin catabolism genes (~22%) than upstream purine catabolism genes (~18%) in SAR11 (Data S1). Allantoin is the product of the first sequence of reactions of purine catabolism in which one of the rings is cleaved and energy is released, while its further degradation releases carbon and nitrogen (Fig. 4). The occurrence of this allantoin-specialist SAR11 subgroup raises the possibility that in some environments other SAR11 cells primarily in search of energy incompletely degrade purines, releasing allantoin that is taken up and further degraded by surrounding cells (Fig. 4).

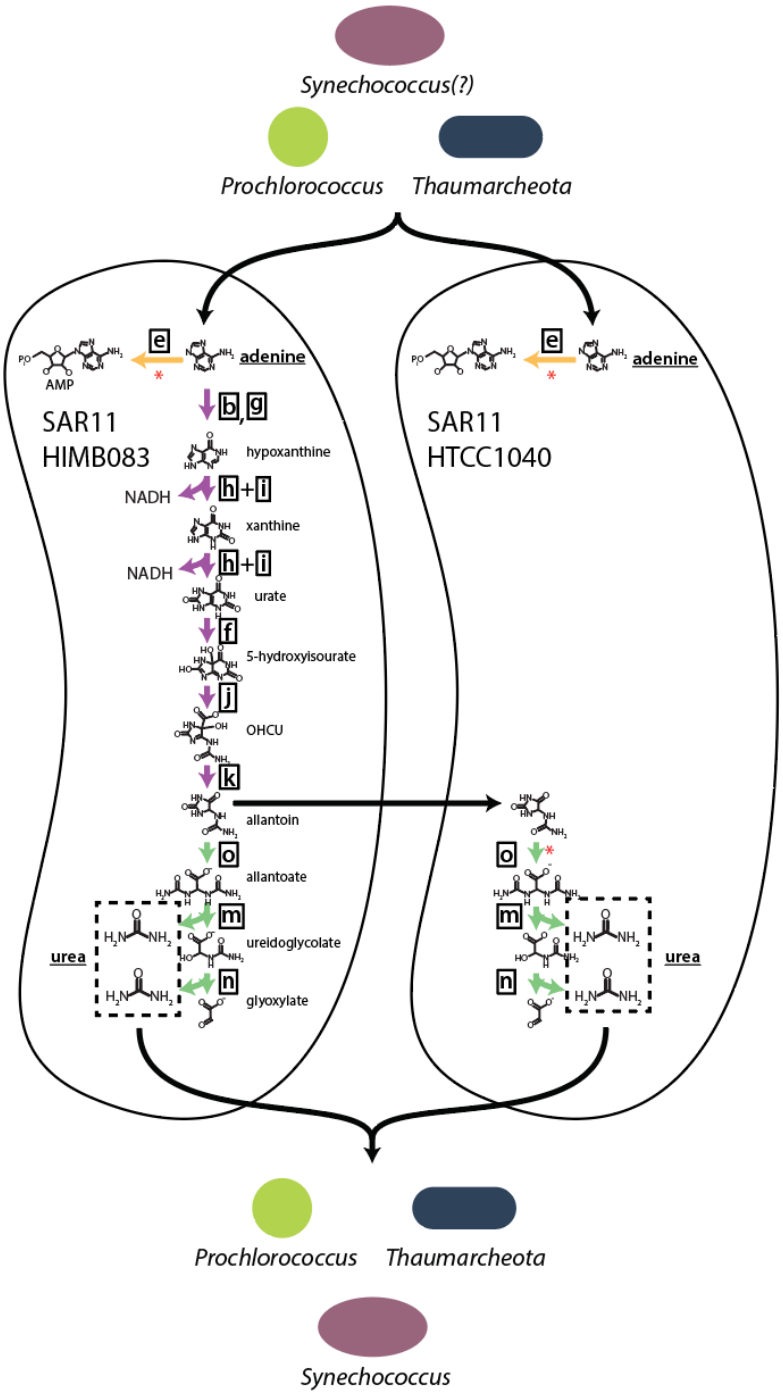


Fig. 4. Purine catabolism in SAR11 and associated putative purine:urea-mediated cryptic nitrogen cycles. *Prochlorococcus, Thaumarcheota* and potentially *Synechococcus* (see text) supply purines that are used by SAR11. Different SAR11 cell types, labeled by representative strains (Fig. 3), possess variants of purine catabolism as shown using sequences of reactions that perform different purine usage functions. Colors of purine usage functions are as defined in Fig. 3: purine assimilation genes in yellow, purine catabolism genes in purple and allantoin catabolism genes in green. Many SAR11 cells lack genomic capacity to use urea produced during purine catabolism, leading to its release and availability to surrounding groups, including *Prochlorococcus, Thaumarcheota* and *Synechococcus*. Some SAR11 cells lack purine catabolism but have genomic capacity to use allantoin, which is inferred to be released by other SAR11 cells that incompletely catabolize purines. OHCU = 2-oxo-4-hydroxy-4-carboxy-5-ureidoimidazoline. Reactions marked with a red asterisk represent enzymes that in subsets of SAR11 genomes are collocated with transporter genes (Fig. 3). Black arrows represent inferred cross-feeding pathways.

231 In contrast to the diversity in pyrimidine/purine usage functions observed in SAR116 and SAR11, nearly
232 all SAR86 genomes contain just two, often-collocated thymidine usage genes: thymidine kinase and a
233 thymidine transporter (Fig. 3, Fig. S2). SAR116 genomes, as well as some deep-branching SAR11
234 genomes, contain other genes for interconverting among pyrimidines that are collocated with
235 thymidine kinase and the transporter, including uridine ribohydrolase, uracil
236 phosphoribosyltransferase and cytidine deaminase, but these are absent in SAR86 (Fig. 3, Fig. S2). This
237 suggests that while SAR116 may have a more general pyrimidine usage strategy, SAR86 focuses
238 specifically on thymidine assimilation.
239
240 ***Niche partitioning of purine usage strategies in SAR11***
241 Since SAR11 is the most abundant of *Prochlorococcus*' sympatric heterotrophs (31,33), we decided to
242 examine its diversity of purine usage strategies more closely. As noted, purine catabolism in SAR11
243 leads to three products: energy in the form of NADH, carbon in the form of glyoxylate, and nitrogen in
244 the forms of ammonia and urea (Fig. 4). However, the bulk of nitrogen released during purine
245 catabolism is released as urea, whereas assimilation of nitrogen into biomolecules universally starts
246 from ammonia. To use nitrogen in urea, cells across the tree of life must first break it down to ammonia
247 using urease. We searched SAR11 genomes for urease genes, and found they occur at a much lower
248 frequency (~5%) than purine catabolism genes (~18-22%) (Fig. 3, Data S1). This suggests additional
249 niche partitioning of purine catabolism within SAR11, with some cells using it to obtain nitrogen, while
250 others use it primarily to obtain energy and/or carbon, releasing urea back to the environment.
251 Genomes in our sample that contain allantoin catabolism genes but lack the rest of the purine
252 catabolism operon also lack urease (Fig. 3), similarly suggesting that these cells use allantoin as a
253 carbon source and release urea back to the environment.
254
255 These observations suggest the existence of purine:urea-mediated cryptic nitrogen cycles involving
256 many abundant microbial groups (Fig. 4). That is, in addition to exuding purines (21) (Fig. 1),
257 *Prochlorococcus* is also a major urea user (35,36). This creates a putative cycle in which purines transfer
258 energy and carbon from *Prochlorococcus* to SAR11 and urea transfers nitrogen back to
259 *Prochlorococcus*. Incomplete purine catabolism by some SAR11 cells leading to potential transfer of
260 allantoin to other surrounding SAR11 cells (Fig. 3 and surrounding discussion) may add an additional
261 step to this cycle in some environments (Fig. 4). Similarly, marine *Synechococcus* also uses urea
262 (35,36), and while purine and thymidine exudation have not been documented in this genus, a marine
263 strain of *Synechococcus elongatus* was found to exude thymidine (37), suggesting the
264 deoxyribonucleotide recycling pathway we identified in *Prochlorococcus* (Fig. 2) could have a broader
265 distribution across other cyanobacteria. Further, ammonia-oxidizing archaea (Thaumarcheota), which
266 are highly abundant autotrophic (CO_2-fixing) microbes lying at the base of the food web below the
267 euphotic zone (38), also exude thymidine and purines (39) and use urea (40). Finally, SAR11 has a
268 broad geographic range that extends across all these groups, from coastal to open waters and from
269 the surface to the deep (41,42). Together these observations suggest that purine:urea-mediated
270 cryptic nitrogen cycles (Fig. 4) may operate throughout much of the world's oceans, linking many
271 abundant groups, warranting their further study in both laboratory and field settings.
272

To further disentangle the forces shaping SAR11 purine usage functions and associated putative nitrogen cycles, we next examined the abundance of relevant genes in metagenomes collected from diverse oceans. We estimated the fraction of SAR11 cells with the capacity for different purine usage strategies in different environments by normalizing the abundance of purine usage genes to a set of single-copy core genes (Methods). We found that genes involved in purine assimilation are always abundant and have a relatively uniform distribution geographically, with depth, and as a function of nutrient concentrations (Fig. 5, Fig. S4, Table S4), consistent with their near-universal presence in SAR11 genomes (Fig. 3).

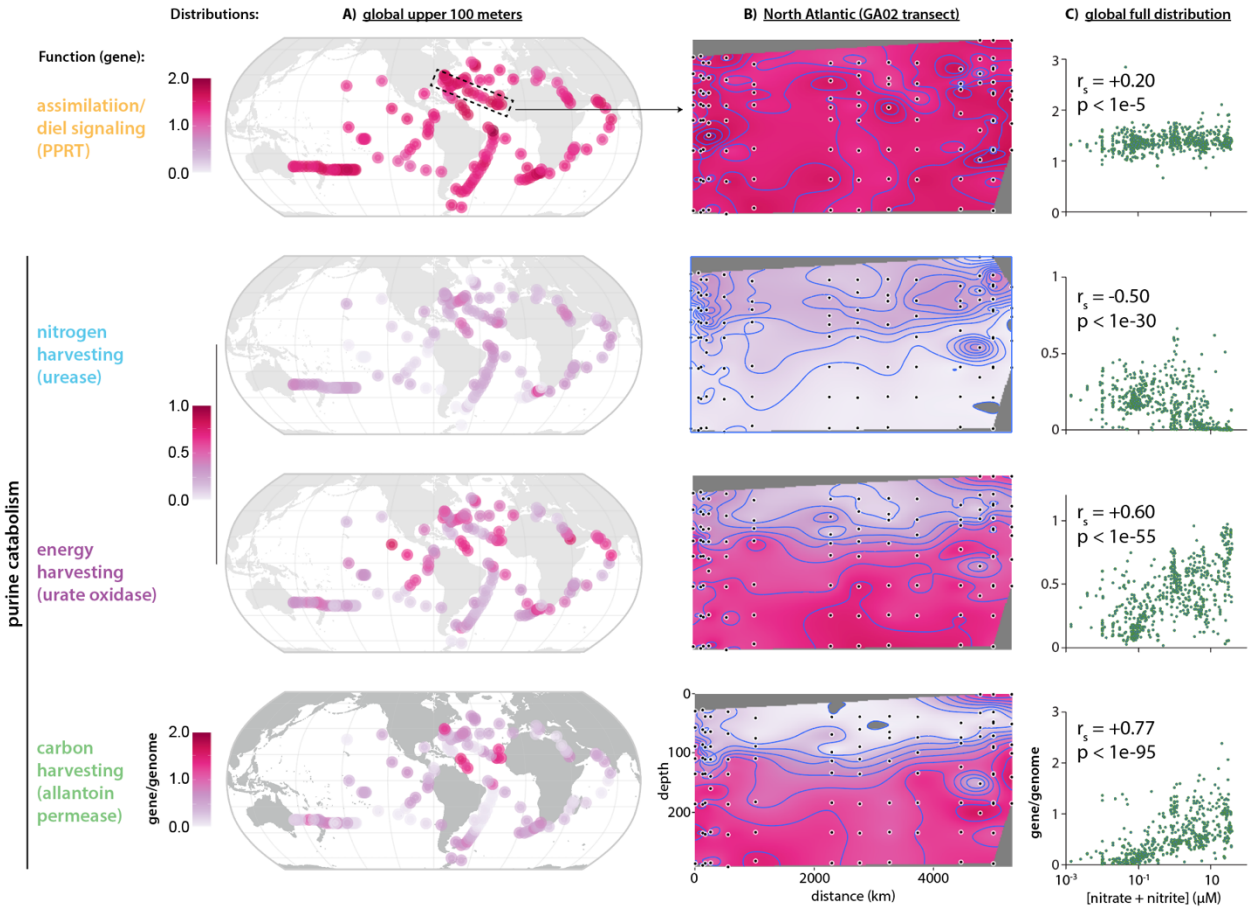


Figure 5. Distribution of SAR11 purine usage genes. Metagenomically-derived estimates for the gene/genome frequency of different purine usage genes in SAR11 are shown as a function of **A)** surface (i.e. upper 100 meters) biogeography, **B)** depth in the water column in the North Atlantic, and **C)** inorganic nitrogen concentration. Purine usage functions are color coded as defined in Fig. 3. For surface biogeography panels each data point represents the mean across samples within the upper 100 meters of the water column at a given station. Spearman rank correlation coefficients (r_S) of the correlation between gene/genome frequency and nitrogen concentration are shown as insets in panels on the right.

Genes mediating other purine usage functions in SAR11 have distinct, non-uniform distributions: genes of the carbon- and energy-harvesting portions of the pathway increase in frequency in environments where nutrients are elevated, including in surface regions in the Arabian Sea, the upwelling regions of the Eastern Equatorial Pacific and the Banguelas system off the tip of South Africa, and the Amazon River plume off the northern coast of South America (Fig. 5, Fig. S4). Genes involved

295 in harvesting nitrogen from purines follow the opposite pattern, reaching their highest abundances in
296 environments where nutrients are depleted (Fig. 5, Fig. S4). While energy and carbon harvesting can
297 proceed without the downstream process of nitrogen harvesting, nitrogen harvesting can only
298 proceed if the full pathway is present (Fig. 4). Hence, the similar abundance of genes for both functions
299 in the nutrient-poor open ocean – where the abundances of genes for carbon/energy-harvesting reach
300 their minimum – suggests that in those environments SAR11 primarily uses purine catabolism to obtain
301 nitrogen.
302
303 Further insight into the forces underlying niche partitioning come from the depth distribution of purine
304 catabolism functions in SAR11, as seen most clearly in the North Atlantic (Fig. 5, Fig. S4). That is, both
305 energy- and carbon-harvesting genes are most abundant at depth, where inorganic nutrient
306 concentrations increase but the energy supply to heterotrophs from primary producers decreases. In
307 contrast, nitrogen-harvesting genes are most abundant near the surface where energy is more
308 available, but nutrients are depleted. Taken together these observations suggest that niche
309 partitioning of SAR11 purine usage functions occurs along a central axis defined by the relative
310 availability of nitrogen and energy. This is consistent with the different transporters associated with the
311 different functions: nitrogen-harvesting genes are associated with an ABC transporter, a high-affinity
312 but energetically costly system requiring ATP hydrolysis (43), while both energy- and carbon-
313 harvesting genes are associated with more energetically efficient nucleotide:cation symporters (44)
314 (Fig. S4, Table S4).
315
316 While genes for the allantoin transporter characteristic of the allantoin-using SAR11 subgroup (Figs
317 3,4) follow a broadly similar trend as genes involved in energy-harvesting, they are nearly absent at the
318 surface where energy-harvesting genes remain present (Fig. 5, Fig. S4). The frequency of carbon-
319 harvesting genes also has a stronger positive correlation with nutrient concentrations than energy-
320 harvesting genes (Fig. 5, Fig. S4, Table S4). These observations are consistent with a decoupling of
321 energy- and carbon-harvesting portions of the purine catabolism pathway in some SAR11 cells as the
322 total flux of purine catabolism increases, thereby making allantoin available to other surrounding cells
323 (Fig. 4). This is further consistent with the use of symporter systems to acquire purines and allantoin at
324 depth, as it suggests that levels of both are higher in deeper waters than at the surface.
325
326 Some of the patterns of niche partitioning of SAR11 purine usage functions appear to be linked to the
327 niche partitioning of clades, adding further support to our conclusions. For example, clades IC and IIB,
328 which dominate in the mesopelagic (41,42) and upper mesopelagic (41,45), respectively, have higher
329 frequencies of genes involved in harvesting energy (~71% in clade IC, ~47% in clade IIB) and carbon
330 (~57% in clade IC, ~44% in clade IIB) from purines than other clades, while also having a high
331 frequency of allantoin transporter genes (~57% in clade IC, ~50% in clade IIB), but lacking urease
332 genes (Fig. S6, Data S1). This is consistent with our conclusions that the harvesting of energy and
333 carbon from purines and the cross-feeding of allantoin become more prominent at depth, whereas
334 nitrogen harvesting from purines is restricted to surface environments (Fig. 5 and surrounding
335 discussion). Further, genomes of clade IA.1, which is abundant in cold surface waters that are typically
336 richer in nutrients (46) have high frequencies of genes involved in allantoin transport (~57%) and
337 catabolism (~71%), but lack genes involved in harvesting energy from purines. This is consistent with

338 our conclusion that in some environments other SAR11 cells more focused on obtaining energy from
339 purines make allantoin available to surrounding cells (Figs 3 and 4 and surrounding discussion).
340 Finally, clades IV, IIIA and IIA, all of which are abundant in surface waters in the Northwestern Atlantic
341 from spring through early fall when waters are relatively more stratified (41), lack all genes for purine
342 catabolism but retain genes for purine assimilation (Fig. S6, Data S1), consistent with our conclusion
343 that this function is key in oligotrophic surface waters. Together these patterns suggest that
344 differentiation of purine usage strategies played a role in the broader metabolic differentiation of
345 SAR11 clades.
346
347 ***Inferred role of purines in the synchronization of ocean-surface populations of SAR11 to daily***
348 ***rhythms in the supply of carbon from photosynthesis***
349 Phylogenomic and metagenomic analyses of purine usage in SAR11 pose an interesting question:
350 what is the function of purine assimilation (yellow pathway in Fig. 4) in cells that lack the downstream
351 purine catabolism pathway? In surface waters of the open ocean, where *Prochlorococcus* dominates,
352 the frequency of purine catabolism genes in SAR11 genomes generally remains below ~30-35% (Fig.
353 5), suggesting that assimilation of intact purines is the major form of purine usage in these populations.
354 To examine possible benefits of purine assimilation, we cultured SAR11 HTCC7211, a strain
355 possessing assimilation genes but lacking purine catabolism genes (Fig. 3), in media amended with
356 adenine at a range of concentrations. Since purines are nitrogen-rich and our metagenomic analyses
357 suggest that SAR11 cells near the surface are often in search of nitrogen (Fig. 5), we performed these
358 experiments in both nutrient-balanced (pyruvate:glycine:methionine = 1:1:0.2) and glycine-depleted
359 (pyruvate:glycine:methionine = 50:1:10) media (see Methods). We expected that by lowering the need
360 for *de novo* purine synthesis, SAR11 cultures might receive a modest growth boost from the adenine
361 additions, reflected either as an increased growth rate and/or a higher final culture density. However,
362 we did not observe growth stimulation, but rather a decrease in growth rate at higher adenine
363 concentrations (Fig. 6). The decrease in growth rate occurred at lower adenine concentrations in
364 glycine-depleted than in nutrient-balanced cultures (Fig. 6).
365

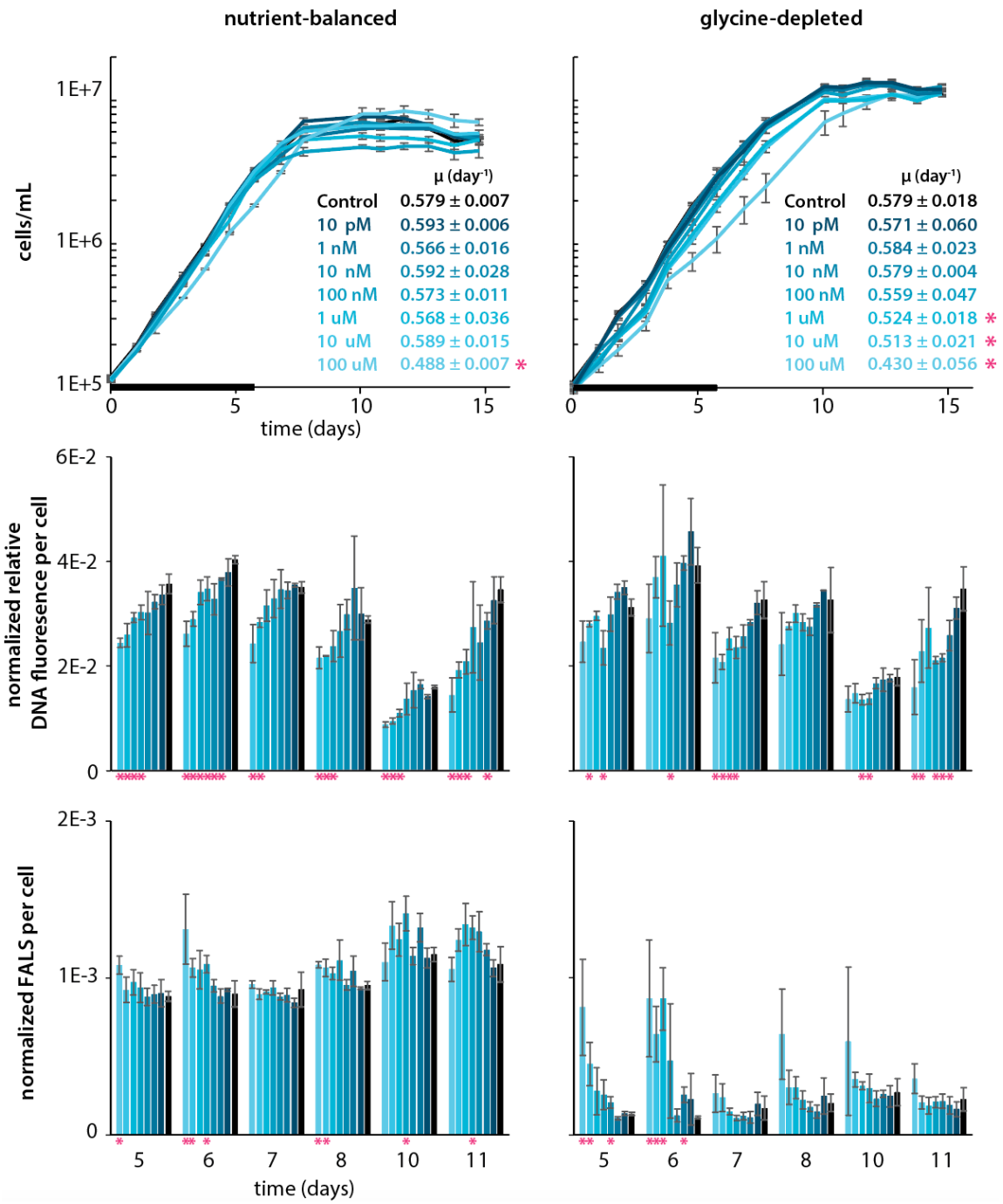


Figure 6. Growth and cell physiology of SAR11 HTCC7211 grown with and without adenine. Top panels: growth curves of SAR11 cells in nutrient-balanced (pyruvate:glycine:methionine = 1:1:0.2) and glycine-depleted (pyruvate:glycine:methionine = 50:1:10) media amended with different concentrations of adenine (shades of blue). Black bars at bottom reflect the time frame used to determine growth rates, which are shown in the insets. Middle and bottom panels: histograms of the averages of bead-normalized relative DNA fluorescence per cell and bead-normalized forward angle light scatter per cell (FALS), respectively, at several time points between mid-exponential growth and stationary phase. For all panels, error bars reflect the standard deviation across three biological replicates, while values significantly different ($p<0.05$, 2-tailed Student's t-test) between adenine-treated and unamended controls are highlighted with a pink asterisk.

376 Puzzled by these results, we wondered if adenine could provide an indirect benefit to cells lacking
377 purine catabolism. A potentially relevant clue comes from the observation that in warm oligotrophic
378 waters where *Prochlorococcus* dominates, SAR11 metabolism is synchronized to the diurnal light:dark
379 cycle (47) despite not being a photoautotroph, and this coupling dissipates in nutrient-enhanced
380 upwelling regions with higher nutrients and fewer *Prochlorococcus* (48). Further, negative feedbacks
381 are often central in the synchronization of biological systems (49,50), and in eukaryotic systems purine
382 and pyrimidine additions have been used to synchronize cell cultures (51-53). These purine and
383 pyrimidine "block" techniques rely on the fact that the specificity of ribonucleotide reductase, which
384 generates all dNTP precursors to DNA (Fig. 2), is controlled through allosteric binding of dNTP
385 molecules (54). Consequently, imbalances in intracellular dNTP pools due to pyrimidine/purine
386 amendments can drive depletion in one or more dNTPs, thereby slowing down or fully inhibiting DNA
387 synthesis and genome replication (51). Subsequent removal of the pyrimidine/purine block
388 amendments by washing cultures results in cells undergoing genome replication and cell division in
389 synchrony (51,52). If similar mechanisms play out in bacteria, this raises the possibility that inferred
390 daily pulses of purines from *Prochlorococcus* (Fig. 2 and surrounding discussion) could act as a signal
391 that helps SAR11 synchronize its metabolism to the daily rhythms of primary production.
392
393 To explore this hypothesis, we first examined flow cytometric data from SAR11 cultures amended with
394 adenine to better understand how adenine inhibits SAR11 growth. Cells in SAR11 cultures amended
395 with adenine have a lower average DNA content, and DNA content decreases as adenine
396 concentrations increase (Fig. 6, Fig. S7). Further, cells from adenine-amended cultures are also larger
397 at the higher concentrations of adenine that inhibit growth (Fig. 6). This increase in size is particularly
398 noticeable in glycine-depleted cultures where growth inhibition sets in at lower adenine
399 concentrations (Fig. 6). Together these results suggest that adenine additions inhibit DNA synthesis in
400 SAR11, which, as the inhibition strengthens, leads to inhibition of cell division. This is similar to how
401 purine and pyrimidine "blocks" function in eukaryotic systems (51-53).
402
403 Next, we examined whether SAR11 cells whose DNA synthesis has been inhibited by adenine are
404 poised to replicate. We grew SAR11 cultures in media with and without adenine, washed cells in mid-
405 exponential growth, and compared their growth after resuspending in adenine-free media. Control
406 cultures grown without adenine often experienced significant growth retardation after washing, with
407 >60% of replicates experiencing major lags and/or never fully recovering (Fig. 7), suggesting that the
408 ultra-centrifugation-based washing protocol induces stress in SAR11 cells. The reasons for the
409 significant divergence of growth trajectories after washing among replicates are not entirely clear, but
410 it was a highly repeatable hallmark of cultures not exposed to adenine. In contrast, cultures treated
411 with adenine display fewer signs of stress after washing, with <20% of replicates experiencing lags or
412 failing to resume growth under most conditions tested (Fig. 7). In nutrient-balanced cultures the
413 benefit of adenine additions is clearest at higher adenine concentrations, while in glycine-depleted
414 cultures the benefit of adenine additions was clear across the full range of concentrations we
415 examined (Fig. 7). Higher adenine concentrations, as well as adenine amendments in glycine-depleted
416 relative to nutrient-balanced cultures, are both associated with greater inhibition of DNA synthesis and
417 cell division (Fig. 6). More consistent reignition of growth after washing of SAR11 cultures experiencing
418 greater inhibition of DNA synthesis is consistent with the inhibition poising cells for replication.

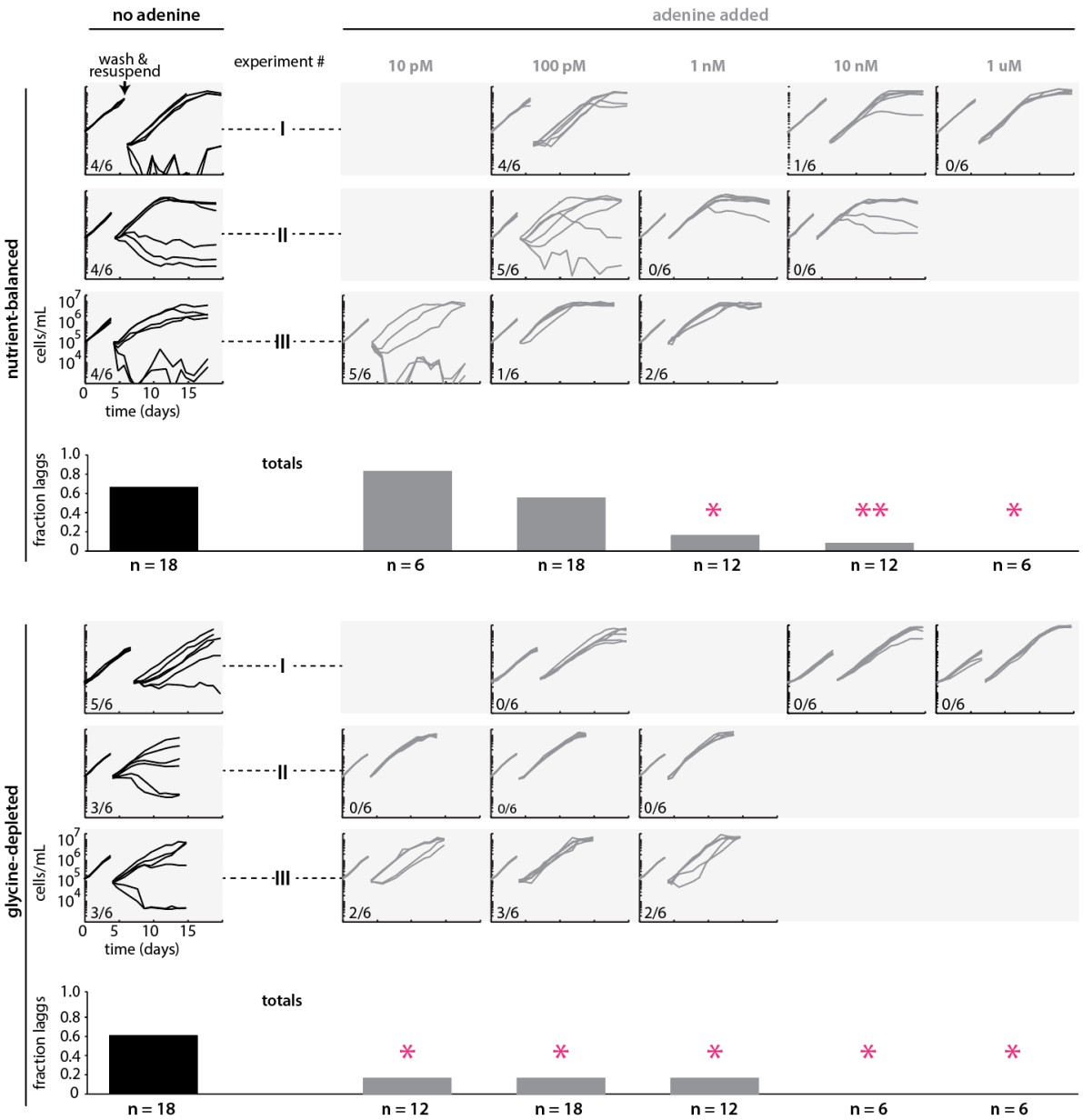


Figure 7. Post-washing growth of SAR11 HTCC7211 grown with and without adenine. Growth curves of SAR11 cultures grown with and without adenine, before and after washing and resuspension in adenine-free media. Three identical experiments were performed in both nutrient-balanced and glycine-depleted cultures, which are labeled using roman numerals. Each experiment included cultures grown with adenine at three different concentrations as well as an adenine-free control, with 6 biological replicates for each condition. Empty entries at given concentrations in different rows represent adenine concentrations not included in different experiments. Fractions of replicates experiencing a lag of 2 or more days after washing at a given adenine concentration and culture condition are shown as inset within each experimental panel, and as fractions of totals across experiments at bottom. Statistically significant differences between adenine amended and adenine-free cultures according to a 2-tailed Fisher's exact test are shown as single pink asterisks for $p<0.05$ and double pink asterisks for $p<0.01$.

From these observations, a conceptual model emerges in which SAR11 populations in surface waters of the open ocean use purines as a signal to achieve synchrony to daily rhythms in the supply of energy and carbon from photosynthesis (Fig. 8). In this model, inferred daily pulses of purines

originating from *Prochlorococcus* (and possibly others) at night (Fig. 2) inhibit SAR11 metabolism, while the daily supply of photosynthate during daytime hours positively stimulates SAR11 metabolism. Over long timescales these alternating signals act in concert to drive synchronization to Earth's rotations, as reflected in *in situ* gene expression (47), allowing SAR11 to achieve greater coupling to the daily supply of solar energy, thereby providing the selective force that maintains purine assimilation genes despite loss of other purine catabolism genes (Fig. 3). In this context, it is worth reconsidering SAR86's genomic profile for using thymidine, which our metabolic reconstruction suggests is also released by *Prochlorococcus* in nightly pulses (Fig. 2). That is, SAR86 has a genomic profile that suggests it focuses specifically on assimilating – and not interconverting, or breaking down – thymidine (Fig. 2), and in that sense is like SAR11 populations that possess purine assimilation genes but lack purine catabolism genes. Further, as in SAR11, *in situ* gene expression indicates that SAR86 metabolism is synchronized to the diurnal light:dark cycle in oligotrophic waters (47), but not in regions of upwelling where *Prochlorococcus* is less abundant (48). This raises the possibility that thymidine could play a similar metabolic role in SAR86 that adenine plays in SAR11, further highlighting the potential value of field studies of the dynamics of dissolved purines and pyrimidines, as well as of the coupled dynamics of *Prochlorococcus* and sympatric heterotrophs.

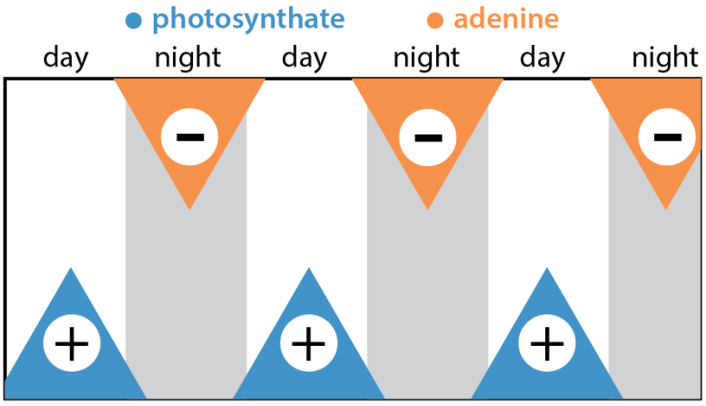


Fig. 7. Hypothesized role of purines in driving synchronization of SAR11 metabolism to the diurnal light:dark cycle. Organic carbon produced by phytoplankton during the day (photosynthate, blue) stimulates SAR11 metabolism, while purines (orange) produced by *Prochlorococcus* at night after genome replication (Fig. 2) inhibit DNA synthesis in SAR11 (Fig. 6). Daily alternation of these positive and negative signals over long timescales drives synchronization to the 24-hour period (47).

Methods

Determining the scale of pyrimidine and purine production and exudation by Prochlorococcus

Data on the intra- and extracellular abundance of metabolites in *Prochlorococcus* were obtained in Ref (21) (also available at MetaboLights (https://www.ebi.ac.uk/metabolights/MTBLS567). Briefly, in our previous study (21), spent media was obtained in mid-exponential growth from nutrient replete batch cultures of *Prochlorococcus* strains MIT9301, MIT0801 and MIT9313 grown at multiple light levels, as well as from semi-continuous P-limited cultures of *Prochlorococcus* MIT9301. Intracellular metabolites were directly extracted from cells using organic solvents and water. Metabolites from the spent media were first extracted using solid phase extraction. Extracts were then analyzed using UPLC-MS/MS to

469 determine concentrations. For extracellular metabolites, concentrations in the extracts were corrected
470 for the extraction efficiency using the formula: $[m]_{media} = 100 \times [m]_{extract} / EF\%$, where $[m]_{media}$ is the
471 concentration of metabolite *m* in the spent media, $[m]_{extract}$ is the measured concentration in the
472 extract, and EF% (range = 0 – 100) is the efficiency with which the solid phase extraction resinextracts
473 metabolite *m* from seawater (55). This correction is only performed for metabolites with EF% > 1 due
474 to associated uncertainties in precisely determining extraction efficiencies when they are low (21).
475 While thymidine and methylthioadenosine have an EF%>1, adenine and guanine have an EF%<1, and
476 so their concentrations are left uncorrected.
477
478 Next, extracellular concentrations of purines and thymidine were normalized to the total amount of
479 these compounds incorporated into DNA over the course of the experiment using experimentally
480 determined cell counts and known GC% of the genomes for each strain, and assuming one
481 chromosome per *Prochlorococcus* cell (29). Thymidine, adenine and methylthioadenosine were
482 normalized to the AT% of genomes, while guanine was normalized to the GC% of genomes.
483
484 For intracellular purines and pyrimidines, concentrations were measured in units of ng/mL in the cell
485 extract and converted to fg/cell using the extract volume and the total number of cells captured on the
486 filter. Total cell numbers are obtained by multiplying culture volume and culture density as determined
487 by flow cytometry. Concentrations in units of fg/cell are then divided by the molar mass of metabolites
488 to obtain concentrations in units of moles/cell. Finally, concentrations in units of moles/cell are divided
489 by cellular volume to obtain molar concentrations. Cellular volumes for *Prochlorococcus* were
490 determined as follows. The volumes of *Prochlorococcus* strains MIT9313 (0.44 μm^3) and MED4 (0.2
491 μm^3) were measured in Ref (56), while the cellular dry masses of strains MIT9301 (60 femtograms),
492 MED4 (66 femtograms), NATL2A (91 femtograms) and MIT9313 (158 grams) were measured in Ref
493 (57). The volume of MIT9301 was estimated from that of MED4 by scaling by the relative dry weight
494 masses of the two strains (i.e. 60 fg/cell for MIT9301 and 66 fg/cell for MED4, giving a ratio of 60/66),
495 leading to 0.18 μm^3 for MIT9301. While there are no volume or mass measurements of MIT0801 it
496 belongs to the same "low light adapted I" (LLI) ecotype as NATL2A and so we estimated its volume
497 from that of MED4 by scaling by the relative dry weight masses of NATL2A and MED4 (i.e. 91 fg/cell for
498 NATL2A and 66 fg/cell for MED4, giving a ratio of 91/66), leading to 0.28 μm^3 for MIT0801.
499
500 *Reconstructing putative deoxyribonucleotide recycling pathway*
501
502 To identify possible pathways involved in the production and/or consumption of thymidine, adenine
503 and guanine, we searched the annotated genomes of all *Prochlorococcus* strains available on the
504 KEGG database (58) for all possible genes involved in the metabolism of these compounds. All strains
505 used in this study (MIT9301, MIT0801 and MIT9313) are included on KEGG. To place findings in a
506 broader context we also searched all *Prochlorococcus* genomes for all genes involved in the
507 metabolism of nucleobases, nucleosides, deoxynucleosides, ribonucleotide monophosphates and
508 deoxyribonucleotide monophosphates, all intermediates of a larger nucleotide metabolic network that
509 thymidine, adenine and guanine could be a part of. To account for the possibility of gene
510 misannotations or incorrect functional assignments in some *Prochlorococcus* strains, homologs
511 identified in some *Prochlorococcus* strains were used to query all *Prochlorococcus* genomes using the

512 BLASTp algorithm within the KEGG BLAST feature. Finally, to identify potentially relevant genes that
513 are present but not functionally assigned in any *Prochlorococcus* genome, we also blasted reference
514 gene sequences from *E. coli* (which has an extensive network of nucleotide metabolism) against
515 *Prochlorococcus* genomes. We included in our search FAMIN, a recently characterized enzyme that is
516 highly multi-functional and catalyzes multiple reactions within nucleotide metabolism (59). Together
517 this led to a list of 9 genes, 5 of which are universal in all strains used in this study (Table S2). Data on
518 the temporal transcriptional profiles of all 9 genes were obtained from a study in which the growth of
519 *Prochlorococcus* MED4 was synchronized to diurnal L:D conditions (29). To evaluate transcriptional
520 dynamics of these genes relative to DNA and RNA synthesis, the transcriptional profile of each gene
521 was plotted alongside the transcriptional profile of DNA polymerase and RNA polymerase (Fig. S1).
522 Pearson correlation coefficients of the expression of each gene relative to DNA and RNA polymerase
523 genes were calculated in Excel (Table S3).
524
525 *Phylometabolic analysis of purine and pyrimidine strategies in heterotrophs*
526
527 Genomes, curated genome phylogenies and associated metadata for SAR11, SAR86 and SAR116 were
528 obtained from the MARMICRODB database (60). The MARMICRODB database includes genomes from
529 a variety of sources, including isolate genomes, single-cell amplified genomes (SAGs) and
530 metagenome assembled genomes (MAGs), but genomes with %completeness – 5x %contamination <
531 30 were excluded in the final dataset, resulting in 186 SAR11 genomes, 57 SAR86 genomes and 59
532 SAR116 genomes. These genomes were first searched for the presence of purine and pyrimidine
533 usage genes using reference sequences in other organisms obtained from the KEGG database (58).
534 This led to the identification of homologs in SAR11, SAR86 and SAR116 that typically co-occur in
535 operons. Homologs from those operons were subsequently used for local searches using the BLASTp
536 algorithm obtained from NCBI (61), with final counting of gene frequencies relying on a combination
537 of sequence similarity and genomic context. For example, many SAR11 and SAR116 genomes that
538 contain a complete purine catabolism operon contain a second homolog also annotated as 5-
539 hydroxyisourate lyase, but which occurs in a different genomic location and has low sequence
540 similarity to the homolog occurring in the operon. Since this gene further occurs in many genomes that
541 do not contain any other purine catabolism genes, we conclude it has a different function and leave it
542 out of our final tally. All purine and pyrimidine usage genes thus identified are listed in Data S1, and
543 their presence/absence data for individual genomes was mapped onto the leaves of the SAR11,
544 SAR86 and SAR116 genome phylogenies to generate phylometabolic trees.
545
546 *Metagenomic analyses*
547
548 675 metagenomes were obtained from TARA and BioGeoTRACES (62,63). These metagenomes were
549 searched for SAR11 purine usage genes and SAR11 ribosomal proteins using custom Hidden Markov
550 Models (HMM's) that were created as follows. First, a set of ~20 representative sequences for the
551 genes of each individual step in the purine catabolism pathway (Figs. 3,4), as well as for a set of 14
552 single-copy ribosomal proteins (64) were obtained at Uniprot (65) and assembled into HMM's. For
553 phosphoribosyltransferase, which has many homologs operating on different nucleobases, the initial
554 set was very narrowly chosen to consist of primarily SAR11 variants occurring with purine catabolism

555 operons. Resulting HMM's were searched against the Uniprot database using the PHMMER algorithm
556 within the HMMER web portal (66), and results were manually inspected to determine e-value cut offs
557 (Data S2) that captured maximal diversity while eliminating homologous genes associated with other
558 functions. The original 20-gene HMM's were then searched against the MARMICRODB database
559 (33,60) and all sequences falling below the e-value cutoffs determined from HMMER results were
560 included our final HMM's to ensure they have broad coverage of diversity within the extant microbial
561 oceans. Next, our custom HMM's were searched against our full set of metagenomes using the GraftM
562 algorithm (64). The results of this search were again manually inspected to determine GraftM cutoff
563 scores (Data S2) that mostly eliminate homologous genes associated with other functions while mostly
564 capturing sequences of interest. To estimate gene/genome frequencies of individual purine usage
565 genes in SAR11, the counts for each purine usage gene were then normalized to the HMM amino acid
566 length and to the median count of the 14 single-copy ribosomal protein genes. Chemical and
567 hydrographic data associated with the BioGEOTRACES metagenomes (63) were obtained from the
568 GEOTRACES Intermediate Data Product IDP2017 version 2 (accessed January 2019), specifically from
569 sections GA02 (67,68) GA03, GA10 (69), and GP13. Environmental data from the Tara Oceans project
570 was obtained from https://doi.pangaea.de/10.1594/PANGAEA.875579. Metagenome sequence data
571 were quality controlled, environmental data matched to metagenomic samples, and missing
572 environmental data were imputed as described in detail in the supplementary material of Hogle et al.
573 2022 (70). Spearman-rank correlations between nutrient concentrations and gene frequencies were
574 determined in R.
575
576 *SAR11 adenine amendment experiments*
577
578 Cells of SAR11 strain HTCC7211 were cultured under constant light at 22 °C in ProMS medium (33).
579 ProMS is a seawater-based modified version of the artificial AMS1 medium (15), in which the
580 concentrations of the main sources of carbon (pyruvate), nitrogen (glycine) and sulfur (methionine)
581 have been lower 50-fold to promote detection of the impact of added organic substrates or
582 phytoplankton exudates (33). In addition, HTCC7211 cells were cultured in a glycine-depleted version
583 of ProMS in which the concentration of all nutrients other than glycine was increased 50-fold to bring
584 them back to the levels of the AMS1 medium. These cultures were then amended with adenine at
585 concentrations ranging seven orders of magnitude, from 10 pM to 100 mM, and their growth was
586 tracked alongside that of adenine-free control cultures for an entire growth curve. Culture density and
587 purity and cell size and DNA content were measured using flow cytometry. Because sub-populations in
588 different phases of the cell cycle are difficult to distinguish in SAR11 (15), we instead quantified the
589 averages and standard deviations over the total bulk fluorescence distributions (Fig. S7). To examine
590 whether adenine exposure poises SAR11 cells for replication, cultures grown with and without adenine
591 were harvested in mid-exponential growth and washed three times using adenine-free culture media
592 and ultracentrifugation at ~25,000 x g (15,000 rpm) for 45 minutes at 22 °C. After the final washing
593 step cell pellets were resuspended in adenine-free media and subsequent growth was monitored via
594 flow cytometry. Fisher's exact tests of the significance in the differences of the number of cultures
595 experiencing lags were determined in R.
596
597

Acknowledgements

We thank Eli Salcedo for help in generating relative DNA fluorescence/cell histograms from raw flow cytometry data for SAR11 adenine amendment experiments, Gretchen Swarr for help in the processing of samples for metabolomics analysis, and Steve Biller for discussions regarding genome replication in bacteria. This research was supported by grants from the Simons Foundation (Award ID 509034SCFY20 to R.B. and S.W.C., Award ID 509034FY20 to E.B.K., and SCOPE Award ID 329108 to M.J. Follows) and from the National Science Foundation (OCE-2019589 to R.B.). This is the NSF Center for Chemical Currencies of a Microbial Planet (C-CoMP) publication #33.

References

1. Pomeroy, L.R., 1974. The ocean's food web, a changing paradigm. *Bioscience*, *24*(9), pp.499-504.
2. Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A. and Thingstad, F., 1983. The ecological role of water-column microbes in the sea. *Marine ecology progress series. Oldendorf*, *10*(3), pp.257-263.
3. Moran, M.A., Kujawinski, E.B., Schroer, W.F., Amin, S.A., Bates, N.R., Bertrand, E.M., Braakman, R., Brown, C.T., Covert, M.W., Doney, S.C, Dyhrman, S.T., A.S. Edison, A.M. Eren, N.M. Levine, L. Li, A.C. Ross, M.A. Saito, A.E. Santoro, D. Segrè, A. Shade, M.B. Sullivan and A. Vardi, 2022. Microbial metabolites in the marine carbon cycle. *Nature microbiology*, *7*(4), pp.508-523.
4. Field, C.B., Behrenfeld, M.J., Randerson, J.T. and Falkowski, P., 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. *science*, *281*(5374), pp.237-240.
5. Falkowski, P.G. and Raven, J.A., 2013. *Aquatic photosynthesis*. Princeton University Press.
6. Braakman, R. and Smith, E., 2012. The compositional and evolutionary logic of metabolism. *Physical biology*, *10*(1), p.011001.
7. Fuhrman, J.A., 1999. Marine viruses and their biogeochemical and ecological effects. *Nature*, *399*(6736), pp.541-548.
8. Wilhelm, S.W. and Suttle, C.A., 1999. Viruses and nutrient cycles in the sea: viruses play critical roles in the structure and function of aquatic food webs. *Bioscience*, *49*(10), pp.781-788.
9. Ward, B.A., Dutkiewicz, S., Jahn, O. and Follows, M.J., 2012. A size-structured food-web model for the global ocean. *Limnology and Oceanography*, *57*(6), pp.1877-1891.
10. Durham, B.P., Sharma, S., Luo, H., Smith, C.B., Amin, S.A., Bender, S.J., Dearth, S.P., Van Mooy, B.A., Campagna, S.R., Kujawinski, E.B. and Armbrust, E.V., 2015. Cryptic carbon and sulfur cycling between surface ocean plankton. *Proceedings of the National Academy of Sciences*, *112*(2), pp.453-457.
11. Segev, E., Wyche, T.P., Kim, K.H., Petersen, J., Ellebrandt, C., Vlamakis, H., Barteneva, N., Paulson, J.N., Chai, L., Clardy, J. and Kolter, R., 2016. Dynamic metabolic exchange governs a marine algal-bacterial interaction. *elife*, *5*, p.e17473.
12. Fritts, R.K., McCully, A.L. and McKinlay, J.B., 2021. Extracellular metabolism sets the table for microbial cross-feeding. *Microbiology and Molecular Biology Reviews*, *85*(1), pp.e00135-20.
13. Giri, S., Oña, L., Waschina, S., Shitut, S., Yousif, G., Kaleta, C. and Kost, C., 2021. Metabolic dissimilarity determines the establishment of cross-feeding interactions in bacteria. *Current Biology*, *31*(24), pp.5547-5557.
14. Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M. and Rappé, M.S., 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, *309*(5738), pp.1242-1245.

15. Carini, P., Steindler, L., Beszteri, S. and Giovannoni, S.J., 2013. Nutrient requirements for growth of the extreme oligotroph 'Candidatus Pelagibacter ubique'HTCC1062 on a defined medium. *The ISME journal*, *7*(3), pp.592-602.
16. Swan, B.K., Tupper, B., Sczyrba, A., Lauro, F.M., Martinez-Garcia, M., González, J.M., Luo, H., Wright, J.J., Landry, Z.C., Hanson, N.W. and Thompson, B.P., 2013. Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proceedings of the National Academy of Sciences*, *110*(28), pp.11463-11468.
17. Chisholm, S.W., Olson, R.J., Zettler, E.R., Goericke, R., Waterbury, J.B. and Welschmeyer, N.A., 1988. A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature*, *334*(6180), pp.340-343.
18. Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincón, J., Zabala, L.L., Jiao, N., Karl, D.M., Li, W.K., Lomas, M.W., Veneziano, D. and Vera, C.S., 2013. Present and future global distributions of the marine Cyanobacteria Prochlorococcus and Synechococcus. *Proceedings of the National Academy of Sciences*, *110*(24), pp.9824-9829.
19. Letscher, R.T., Moore, J.K., Martiny, A.C. and Lomas, M.W., 2023. Biodiversity and Stoichiometric Plasticity Increase Pico-Phytoplankton Contributions to Marine Net Primary Productivity and the Biological Pump. *Global Biogeochemical Cycles*, *37*(8), p.e2023GB007756.
20. Braakman, R., Follows, M.J. and Chisholm, S.W., 2017. Metabolic evolution and the self-organization of ecosystems. *Proceedings of the National Academy of Sciences*, *114*(15), pp.E3091-E3100.
21. Kujawinski, E.B., Braakman, R., Longnecker, K., Becker, J.W., Chisholm, S.W., Dooley, K., Kido Soule, M.C., Swarr, G.J. and Halloran, K., 2023. Metabolite diversity among Prochlorococcus strains belonging to divergent ecotypes. *mSystems*, e01261-22.
22. Dufresne, A., Garczarek, L. and Partensky, F., 2005. Accelerated evolution associated with genome reduction in a free-living prokaryote. *Genome biology*, *6*, pp.1-10
23. Grzymski, J.J. and Dussaq, A.M., 2012. The significance of nitrogen cost minimization in proteomes of marine microorganisms. *The ISME journal*, *6*(1), pp.71-80.
24. Proudfoot, M., Kuznetsova, E., Brown, G., Rao, N.N., Kitagawa, M., Mori, H., Savchenko, A. and Yakunin, A.F., 2004. General enzymatic screens identify three new nucleotidases in Escherichia coli: biochemical characterization of SurE, YfbR, and YjjG. *Journal of Biological Chemistry*, *279*(52), pp.54687-54694.
25. Vodnala, M., Ranjbarian, F., Pavlova, A., de Koning, H.P. and Hofer, A., 2016. Trypanosoma brucei methylthioadenosine phosphorylase protects the parasite from the antitrypanosomal effect of deoxyadenosine: implications for the pharmacology of adenosine antimetabolites. *Journal of Biological Chemistry*, *291*(22), pp.11717-11726.
26. Cacciapuoti, G., Bertoldo, C., Brio, A., Zappia, V. and Porcelli, M., 2003. Purification and characterization of 5′-methylthioadenosine phosphorylase from the hyperthermophilic archaeon Pyrococcus furiosus: substrate specificity and primary structure analysis. *Extremophiles*, *7*, pp.159-168.
27. Noell, S.E. and Giovannoni, S.J., 2019. SAR11 bacteria have a high affinity and multifunctional glycine betaine transporter. *Environmental Microbiology*, *21*(7), pp.2559-2575.
28. Noell, S.E., Barrell, G.E., Suffridge, C., Morré, J., Gable, K.P., Graff, J.R., VerWey, B.J., Hellweger, F.L. and Giovannoni, S.J., 2021. Sar11 cells rely on enzyme multifunctionality to metabolize a range of polyamine compounds. *MBio*, *12*(4), pp.e01091-21.
29. Zinser, E.R., Lindell, D., Johnson, Z.I., Futschik, M.E., Steglich, C., Coleman, M.L., Wright, M.A., Rector, T., Steen, R., McNulty, N. and Thompson, L.R., 2009. Choreography of the transcriptome, photophysiology, and cell cycle of a minimal photoautotroph, Prochlorococcus. *PloS one*, *4*(4), p.e5135.
30. Chuang, Y.-C. , Haas, N.W. , Pepin, R., Behringer, M., Oda, Y., LaSarre, B., Harwood, C.S., McKinlay, J.B., 2023. A purine salvage bottleneck leads to bacterial adenine cross-feeding, bioRxiv 2023.10.17.562681
31. Morris, R.M., Rappé, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A. and Giovannoni, S.J., 2002. SAR11 clade dominates ocean surface bacterioplankton communities. *Nature*, *420*(6917), pp.806-810.

32. Treusch, A.H., Vergin, K.L., Finlay, L.A., Donatz, M.G., Burton, R.M., Carlson, C.A. and Giovannoni, S.J., 2009. Seasonality and vertical structure of microbial communities in an ocean gyre. *The ISME journal*, *3*(10), pp.1148-1163.
33. Becker, J.W., Hogle, S.L., Rosendo, K. and Chisholm, S.W., 2019. Co-culture and biogeography of Prochlorococcus and SAR11. *The ISME journal*, *13*(6), pp.1506-1519.
34. Haro-Moreno, J.M., Rodriguez-Valera, F., Rosselli, R., Martinez-Hernandez, F., Roda-Garcia, J.J., Gomez, M.L., Fornas, O., Martinez-Garcia, M. and López-Pérez, M., 2020. Ecogenomics of the SAR11 clade. *Environmental Microbiology*, *22*(5), pp.1748-1763.
35. Moore, L.R., Post, A.F., Rocap, G. and Chisholm, S.W., 2002. Utilization of different nitrogen sources by the marine cyanobacteria Prochlorococcus and Synechococcus. *Limnology and oceanography*, *47*(4), pp.989-996.
36. Berthelot, H., Duhamel, S., L'helguen, S., Maguer, J.F., Wang, S., Cetinić, I. and Cassar, N., 2019. NanoSIMS single cell analyses reveal the contrasting nitrogen sources for small phytoplankton. *The ISME Journal*, *13*(3), pp.651-662.
37. Fiore, C.L., Longnecker, K., Kido Soule, M.C. and Kujawinski, E.B., 2015. Release of ecologically relevant metabolites by the cyanobacterium *Synechococcus elongatus* CCMP 1631. *Environmental microbiology*, *17*(10), pp.3949-3963.
38. Karner, M.B., DeLong, E.F. and Karl, D.M., 2001. Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature*, *409*(6819), pp.507-510.
39. Bayer, B., Hansman, R.L., Bittner, M.J., Noriega-Ortega, B.E., Niggemann, J., Dittmar, T. and Herndl, G.J., 2019. Ammonia-oxidizing archaea release a suite of organic compounds potentially fueling prokaryotic heterotrophy in the ocean. *Environmental Microbiology*, *21*(11), pp.4062-4075.
40. Kitzinger, K., Padilla, C.C., Marchant, H.K., Hach, P.F., Herbold, C.W., Kidane, A.T., Könneke, M., Littmann, S., Mooshammer, M., Niggemann, J. and Petrov, S., 2019. Cyanate and urea are substrates for nitrification by Thaumarchaeota in the marine environment. *Nature microbiology*, *4*(2), pp.234-243.
41. Giovannoni, S.J., 2017. SAR11 bacteria: the most abundant plankton in the oceans. *Annual review of marine science*, *9*, pp.231-255.
42. Cameron Thrash, J., Temperton, B., Swan, B.K., Landry, Z.C., Woyke, T., DeLong, E.F., Stepanauskas, R. and Giovannoni, S.J., 2014. Single-cell enabled comparative genomics of a deep ocean SAR11 bathytype. *The ISME journal*, *8*(7), pp.1440-1451.
43. Rees, D.C., Johnson, E. and Lewinson, O., 2009. ABC transporters: the power to change. *Nature reviews Molecular cell biology*, *10*(3), pp.218-227
44. Weyand, S., Shimamura, T., Yajima, S., Suzuki, S.I., Mirza, O., Krusong, K., Carpenter, E.P., Rutherford, N.G., Hadden, J.M., O'Reilly, J. and Ma, P., 2008. Structure and molecular mechanism of a nucleobase–cation–symport-1 family transporter. *Science*, *322*(5902), pp.709-713
45. Carlson, C.A., Morris, R., Parsons, R., Treusch, A.H., Giovannoni, S.J. and Vergin, K., 2009. Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *The ISME journal*, *3*(3), pp.283-295.
46. Brown, M.V., Lauro, F.M., DeMaere, M.Z., Muir, L., Wilkins, D., Thomas, T., Riddle, M.J., Fuhrman, J.A., Andrews-Pfannkoch, C., Hoffman, J.M. and McQuaid, J.B., 2012. Global biogeography of SAR11 marine bacteria. *Molecular systems biology*, *8*(1), p.595.
47. Ottesen, E.A., Young, C.R., Gifford, S.M., Eppley, J.M., Marin III, R., Schuster, S.C., Scholin, C.A. and DeLong, E.F., 2014. Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science*, *345*(6193), pp.207-212.
48. Ottesen, E.A., Young, C.R., Eppley, J.M., Ryan, J.P., Chavez, F.P., Scholin, C.A. and DeLong, E.F., 2013. Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proceedings of the National Academy of Sciences*, *110*(6), pp.E488-E497.

49. Gonze, D., Halloy, J. and Goldbeter, A., 2002. Robustness of circadian rhythms with respect to molecular noise. *Proceedings of the National Academy of Sciences*, *99*(2), pp.673-678.
50. Mondragón-Palomino, O., Danino, T., Selimkhanov, J., Tsimring, L. and Hasty, J., 2011. Entrainment of a population of synthetic genetic oscillators. *Science*, *333*(6047), pp.1315-1319.
51. Xeros, N., 1962. Deoxyriboside control and synchronization of mitosis. *Nature*, *194*(4829), pp.682-683.
52. Mitchison, J.M. and Creanor, J., 1971. Induction synchrony in the fission yeast Schizosaccharomyces pombe. *Experimental cell research*, *67*(2), pp.368-374.
53. Bjursell, G. and Reichard, P., 1973. Effects of thymidine on deoxyribonucleoside triphosphate pools and deoxyribonucleic acid synthesis in Chinese hamster ovary cells. *Journal of Biological Chemistry*, *248*(11), pp.3904-3909.
54. Jordan, A. and Reichard, P., 1998. Ribonucleotide reductases. *Annual review of biochemistry*, *67*(1), pp.71-98.
55. Johnson, W.M., Kido Soule, M.C. and Kujawinski, E.B., 2017. Extraction efficiency and quantification of dissolved metabolites in targeted marine metabolomics. *Limnology and Oceanography: Methods*, *15*(4), pp.417-428.
56. Ting, C.S., Hsieh, C., Sundararaman, S., Mannella, C. and Marko, M., 2007. Cryo-electron tomography reveals the comparative three-dimensional architecture of Prochlorococcus, a globally important marine cyanobacterium. *Journal of bacteriology*, *189*(12), pp.4485-4493.
57. Cermak, N., Becker, J.W., Knudsen, S.M., Chisholm, S.W., Manalis, S.R. and Polz, M.F., 2017. Direct single-cell biomass estimates for marine bacteria via Archimedes' principle. *The ISME journal*, *11*(3), pp.825-828.
58. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. and Morishima, K., 2017. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic acids research*, *45*(D1), pp.D353-D361.
59. Cader, M.Z., de Almeida Rodrigues, R.P., West, J.A., Sewell, G.W., Muhammad, N., Reikine, S., Sirago, G., Unger, L.W., Iglesias-Romero, A.B., Ramshorn, K. and Haag, L.M., 2020. FAMIN is a multifunctional purine enzyme enabling the purine nucleotide cycle. Cell, 180(2), pp.278-295.
60. Hogle, S. L. (2019). MARMICRODB database for taxonomic classification of (marine) metagenomes (1.0.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.3520509
61. Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. and Madden, T.L., 2009. BLAST+: architecture and applications. *BMC bioinformatics*, *10*, pp.1-9.
62. Sunagawa, S., et al., 2015. Structure and function of the global ocean microbiome. *Science*, *348*(6237), p.1261359.
63. Biller, S.J., Berube, P.M., Dooley, K., Williams, M., Satinsky, B.M., Hackl, T., Hogle, S.L., Coe, A., Bergauer, K., Bouman, H.A. and Browning, T.J., 2018. Marine microbial metagenomes sampled across space and time. *Scientific data*, *5*(1), pp.1-7.
64. Boyd, J.A., Woodcroft, B.J. and Tyson, G.W., 2018. GraftM: a tool for scalable, phylogenetically informed classification of genes within metagenomes. *Nucleic Acids Research*, *46*(10), pp.e59-e59.
65. "UniProt: the universal protein knowledgebase in 2023." *Nucleic Acids Research* 51, no. D1 (2023): D523-D531.
66. Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R. and Finn, R.D., 2018. HMMER web server: 2018 update. *Nucleic acids research*, *46*(W1), pp.W200-W204.
67. Salt, L.A., van Heuven, S.M., Claus, M.E., Jones, E.M. and De Baar, H.J.W., 2015. Rapid acidification of mode and intermediate waters in the southwestern Atlantic Ocean. *Biogeosciences*, *12*(5), pp.1387-1401.
68. Rijkenberg, M.J., Middag, R., Laan, P., Gerringa, L.J., van Aken, H.M., Schoemann, V., de Jong, J.T. and De Baar, H.J., 2014. The distribution of dissolved iron in the West Atlantic Ocean. *PloS one*, *9*(6), p.e101323.

69. Wyatt, N.J., Milne, A., Woodward, E.M.S., Rees, A.P., Browning, T.J., Bouman, H.A., Worsfold, P.J. and Lohan, M.C., 2014. Biogeochemical cycling of dissolved zinc along the GEOTRACES South Atlantic transect GA10 at 40 S. *Global Biogeochemical Cycles*, *28*(1), pp.44-56.
70. Hogle, S.L., Hackl, T., Bundy, R.M., Park, J., Satinsky, B., Hiltunen, T., Biller, S., Berube, P.M. and Chisholm, S.W., 2022. Siderophores as an iron source for picocyanobacteria in deep chlorophyll maximum layers of the oligotrophic ocean. *The ISME Journal*, *16*(6), pp.1636-1646.