



Circadian clock helps cyanobacteria manage energy in coastal and high latitude ocean

Ferdi L. Hellweger¹ · Maria Luísa Jabbur² · Carl Hirschie Johnson² · Erik van Sebille³ · Hideharu Sasaki⁴

Received: 17 May 2019 / Revised: 9 October 2019 / Accepted: 24 October 2019 / Published online: 4 November 2019
© The Author(s), under exclusive licence to International Society for Microbial Ecology 2019

Abstract

The circadian clock coordinates cellular functions over the diel cycle in many organisms. The molecular mechanisms of the cyanobacterial clock are well characterized, but its ecological role remains a mystery. We present an agent-based model of *Synechococcus* (harboring a self-sustained, bona fide circadian clock) that explicitly represents genes (e.g., *kaiABC*), transcripts, proteins, and metabolites. The model is calibrated to data from laboratory experiments with wild type and no-clock mutant strains, and it successfully reproduces the main observed patterns of glycogen metabolism. Comparison of wild type and no-clock mutant strains suggests a main benefit of the clock is due to energy management. For example, it inhibits glycogen synthesis early in the day when it is not needed and energy is better used for making the photosynthesis apparatus. To explore the ecological role of the clock, we integrate the model into a dynamic, three-dimensional global circulation model that includes light variability due to seasonal and diel incident radiation and vertical extinction. Model output is compared with field data, including in situ gene transcript levels. We simulate cyanobacteria with and without a circadian clock, which allows us to quantify the fitness benefit of the clock. Interestingly, the benefit is weakest in the low latitude open ocean, where *Prochlorococcus* (lacking a self-sustained clock) dominates. However, our attempt to experimentally validate this testable prediction failed. Our study provides insights into the role of the clock and an example for how models can be used to integrate across multiple levels of biological organization.

Introduction

The circadian clock, a timekeeping mechanism with an ~24 h period, helps organisms coordinate functions over the diel light cycle. Many of the mechanisms of the cyanobacterial clock are now characterized at the molecular level [1, 2] and

the fitness benefit has been demonstrated experimentally [3]. Despite this, the mechanism(s) by which the circadian clock increases fitness is unknown and fundamental questions in ecology and evolution, e.g., why some cyanobacteria have a circadian clock and some do not, remain unanswered. Connecting across these scales of biological organization, from molecular biology to ecology, is a challenge common to many areas of the biological sciences.

In *Synechococcus*, the circadian clock controls gene expression at the genome scale [4–6], so its role in the biology and fitness is expected to be multifaceted and complex. Recent evidence suggests the clock plays a role in the metabolism of glycogen (GLG) [7–9]. In general, GLG is synthesized during the day as an energy reserve, and then broken down at night to support cellular functions. Several of the genes involved in this mechanism are under the control of the circadian clock. The (relatively) well-understood intracellular function of GLG and clock control make the GLG metabolism mechanism a good candidate to begin to investigate the role of the clock on fitness.

The circadian clock may play a role in the biogeography of the two most abundant marine cyanobacteria;

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-019-0547-0>) contains supplementary material, which is available to authorized users.

✉ Ferdi L. Hellweger
ferdi.hellweger@tu-berlin.de

¹ Water Quality Engineering, Technical University of Berlin, Berlin, Germany

² Department of Biological Sciences, Vanderbilt University, Nashville, TN, USA

³ Institute for Marine and Atmospheric Research, Utrecht University, Utrecht, The Netherlands

⁴ Application Laboratory, Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokohama, Japan

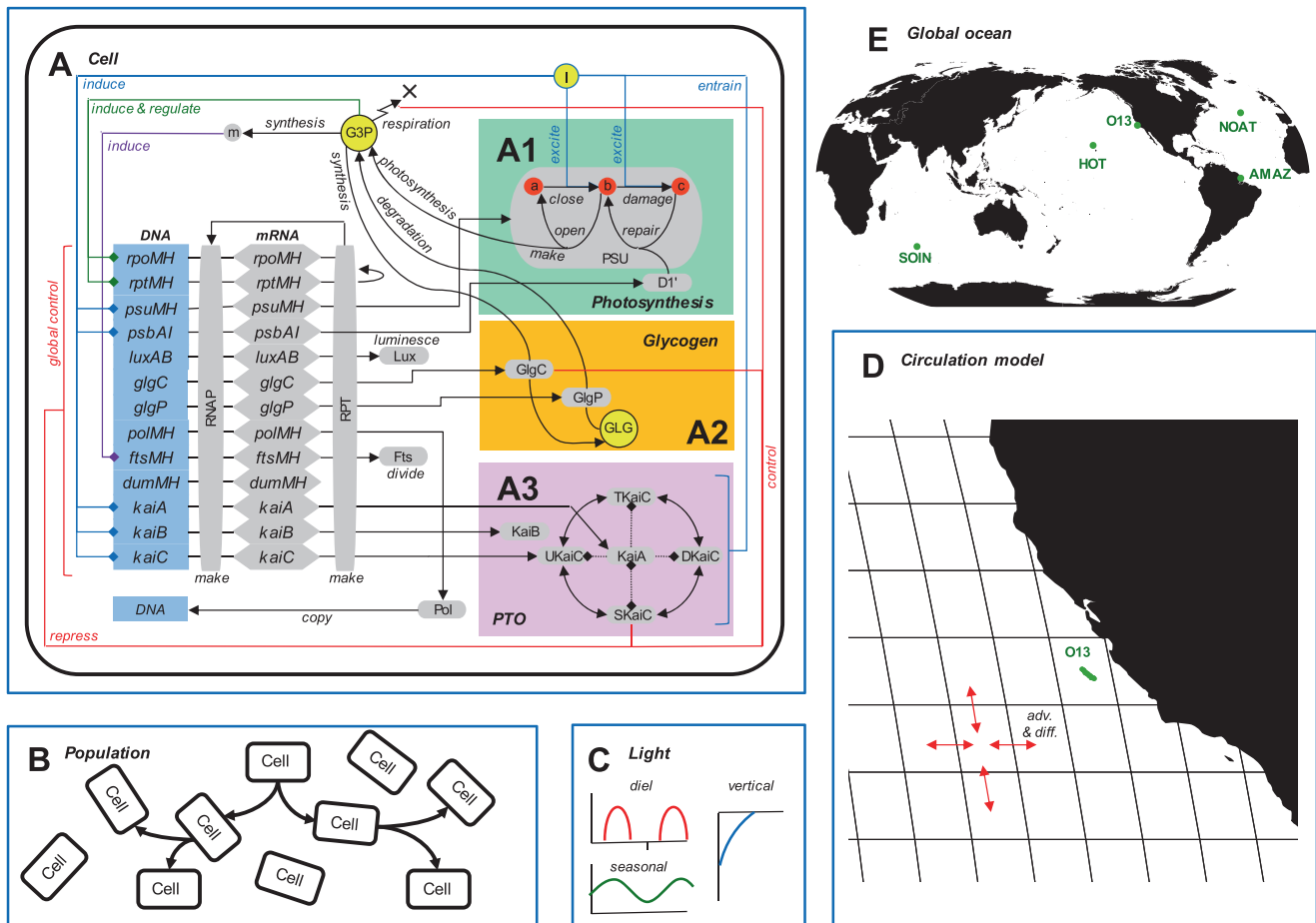


Fig. 1 Model schematic. From genes to ecosystems. **a** Cell with basic intracellular mechanisms including (A1) photosynthesis system, (A2) glycogen metabolism, and (A3) Posttranslational oscillator (PTO). I irradiance, PSU photosynthetic unit (a = open, b = closed, c = damaged), D1' free photosystem II reaction center protein, G3P internal stored energy, m = cell size/mass, and GLG glycogen. Gene/protein: *rpoMH*/RNAP RNA polymerase, *rptMH*/RPT ribosome, *psuMH*/PSU photosynthetic unit, *psbA*/D1 PSII reaction center, *luxAB*/Lux luciferase, *polMH*/Pol DNA polymerase, *ftsMH*/Fts cell division, *dumMH*—, dummy (accounts for genes not explicitly considered), *kaiA*/KaiA, *kaiB*/KaiB, *kaiC*/KaiC (UKaiC unphosphorylated, SKaiC phosphorylated only on serine 431, TKaiC phosphorylated only on threonine 432,

DKaiC phosphorylated on both sites), circadian clock, and *glgC*/GlgC glycogen synthesis. *glgP*/GlgP glycogen degradation. Schematic is simplified and not all processes are shown for clarity (see Section S1, Figs. S1–S14, Tables S1–24 for model details). **b** Population. An agent-based approach is used to scale up from individual cells to the population. **c** Light regime. Diel light pattern based on sunset/sunrise calculations, seasonal, and horizontal based on monthly MODIS photosynthetically available radiation (PAR) and vertical attenuation based on MODIS Chlorophyll *a*. **d** Circulation model grid. $2^\circ \times 2^\circ$ resolution in the horizontal, 22 layers in the vertical, transport based on OFES model [20]. Green points marked O13 is Lagrangian track of transcriptomics survey off Californian coast from ref. [32], see Fig. 3c. **e** Global scale

Prochlorococcus, which dominates in the low-latitude open ocean, and *Synechococcus*, which dominates in high-latitude and coastal waters [10, 11]. The geographic distribution of these cyanobacteria depends on many factors, including their response to temperature and light [11], but they also have different timekeeping mechanisms. *Synechococcus* has a bona fide circadian clock that exhibits free-running rhythms. *Prochlorococcus* does not have a circadian clock, rather an hourglass timing mechanism that does not free-run [2, 12–14].

Here we explore the ecological role of the cyanobacterial circadian clock using mechanistic modeling. We develop a model of *Synechococcus*, which includes a molecular-level

representation of the circadian clock, including genes, transcripts, proteins (incl. phosphorylation state), and metabolites, as well as GLG metabolism. We calibrate the model to laboratory observations, which shows it reproduces the main observed pattern for wild type and no-clock mutant strains. Then we integrate it into a global ocean model and compare with field observations, including clock gene (*kaiABC*) and photosynthesis (*psbA*) transcript levels. We simulate wild type and no-clock mutants and compute the spatial pattern of fitness benefit (selection coefficient) conferred by the circadian clock, which is generally consistent with the observed biogeography of marine cyanobacteria *Synechococcus* and *Prochlorococcus*.

Model overview

The dynamic gene-level model of *Synechococcus* is based on a previous version [15] that was expanded to include glycogen metabolism and integrated into a global ocean circulation model. A complete model description, following the ODD (Overview, Design concepts, and Details) model documentation standard [16] is provided in Section S1 (Figs. S1–14, Tables S1–24). The remainder of this section presents an overview of the model (following the Overview part of the ODD protocol).

The purpose of the model is to explore how the circadian clock acts and interacts with other intracellular mechanisms (i.e., glycogen metabolism) to produce cellular behavior, and how it ultimately affects the ecological fitness of cyanobacteria. The basic entities in the model are individual cells, which are simulated using an agent-based modeling approach to resolve population heterogeneity (i.e., cell cycle phasing [17]) and allow for comparison to single-cell observations [15]. Each cell has a number of state variables, including genes, transcripts, proteins (incl. phosphorylation level), and metabolites. The model simulates a select number of representative genes using a coarse-grained modeling approach [18]. For example, *glgP* is representative of all genes involved in the degradation of GLG to G3P. The model resolves the diel light cycle and is applied at a number of temporal and spatial scales, ranging from a few days in a

zero-dimensional laboratory beaker to a year in the three-dimensional global ocean.

The model includes a number of intracellular processes that are resolved at the gene or molecular level (Fig. 1a). Genes are transcribed by the RNA polymerase (RNAP) to produce transcripts (mRNA), which are translated by the ribosome (RPT) to yield proteins that perform various functions. For example, the photosynthetic unit (PSU, Fig. 1A1) produces energy (G3P), which is then converted to biomass (m).

The model incorporates a GLG pool, genes responsible for synthesis (*glgC*) and degradation (*glgP*), and the associated metabolic cost, based on recent evidence demonstrating the role of the circadian clock in GLG metabolism [7, 8, 19] (Fig. 1A2). The energy pool (G3P) serves as substrate and breakdown product for GLG. In the dark, G3P is removed by respiration and supplied by GLG breakdown, and when the cell runs out of G3P it dies (referred to here as “dark death”).

The circadian clock is encoded by the *kaiABC* genes and consists of two parts, including posttranslational oscillator (PTO, Fig. 1A3) and transcriptional/translational feedback loop. The model can reproduce the major observed features of the clock in laboratory experiments, including entrainment to light/dark (LD) cycles, free-running rhythms, phase shifts by dark pulses, and KaiC protein dynamics. The model proceeds in discrete time steps using an explicit finite difference solution to the intracellular differential equations (e.g., protein mass balances).

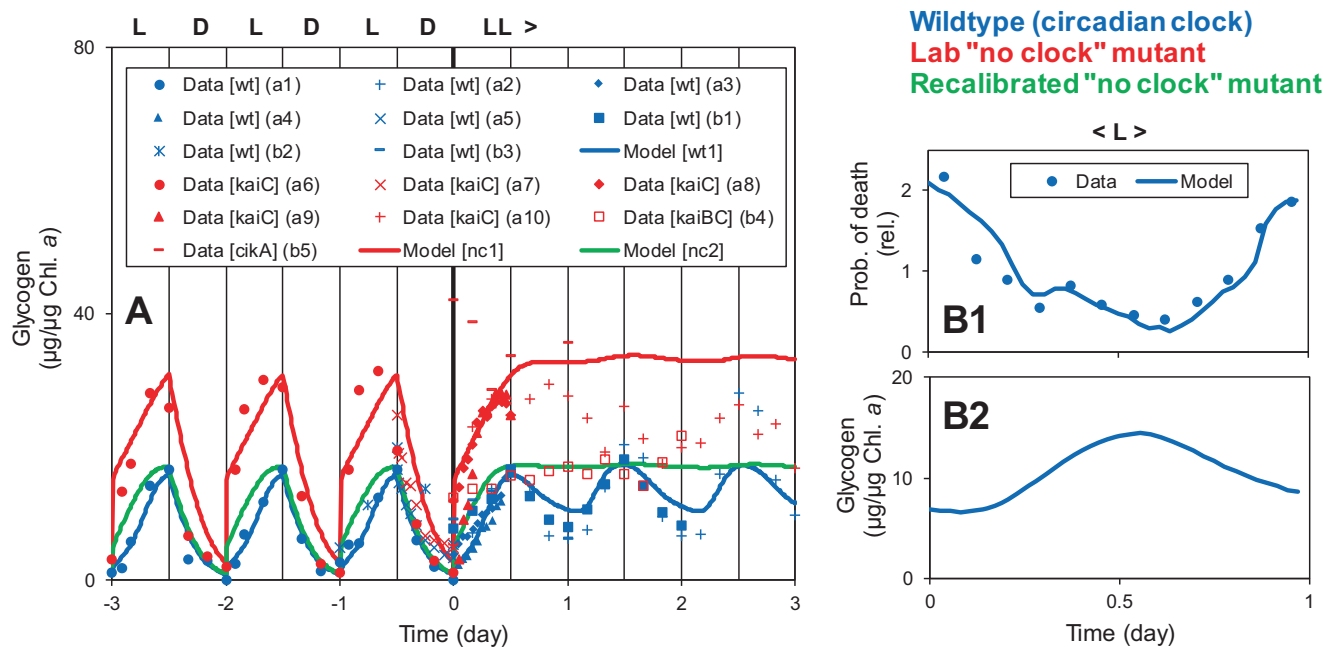


Fig. 2 Model-data comparison for glycogen metabolism in *Synechococcus elongatus* strain PCC 7942. Blue: Wildtype (wt1), Red and green: No-clock mutants (nc1, nc2). **a** GLG level in cells grown in L:D 12:12 and transferred to LL at 0 day. Data from **a** ref. [7] and **b** ref. [8].

b Probability of death by an 18 h dark pulse applied at different points in the circadian cycle and GLG level. Data from ref. [19]. See ref. [15] and Section S2 for additional model-data comparisons and discussion

The *Synechococcus* model is integrated into a global ocean circulation model [20], which was aggregated in the horizontal onto a $2^\circ \times 2^\circ$ grid [21]. Incident light intensity (photosynthetically available radiation, PAR) is based on MODIS satellite observations, which was spread out over the diel cycle and attenuated vertically based on MODIS Chlorophyll *a* [22]. Individual cells are transported based on output from the circulation model (e.g., velocities). The population grows as a result of individual cells dividing (cell growth is based on the net effect of photosynthesis and respiration) and dying (intrinsic death when the cell runs out of G3P, i.e., dark death). Nutrients are not included and to constrain the population, an extrinsic death process that accounts for grazing and viral lysis is additionally included. Specifically, the extrinsic death rate is adjusted spatially and dynamically to produce total population patterns consistent with observed horizontal Chlorophyll *a* concentrations from MODIS, which is based on formulations used in previous ocean models [23, 24]. The total population size is therefore imposed and not an emergent output from the model, which is acceptable in this case because the focus is on the relative fitness of various *Synechococcus* strains. Additional discussion of imposed vs. emergent patterns, including a classification of all model output, is presented in Section S1 and Table S1.

Glycogen metabolism of the wild type in the lab

In the wild type, the GLG synthesis rate is light- and clock-controlled and the GLG level fluctuates in L:D and this fluctuation continues when transferred to LL (Fig. 2a, the mutants are discussed in the next section). The sustained rhythm in continuous light is clear evidence for the control of the circadian clock on GLG synthesis, but this is of little ecological relevance since cells do not live in this light regime in the wild. In L:D cycles, the circadian clock functions to suppress GLG synthesis early in the light period, which is evident in the time course of GLG level, especially when compared with the no-clock mutant. This feature has an important effect on the fitness difference between the wild type and mutant strains, as discussed in detail later.

The model was designed and calibrated to reproduce these data, i.e., by making the GLG synthesis rate light- and clock-controlled and calibrating various parameters (e.g., the GlgC rate constant, see Section S1), and this is therefore not an emergent pattern or independent prediction. Nonetheless, the model-data comparison serves as a check and supports the mechanistic realism of the model.

Glycogen serves as a substrate for nighttime respiration [25] and mutants of *Synechocystis* defective in GLG synthesis are not viable in L:D light regimes [26]. We included this effect by killing cells that run out of G3P, which happens in cells without a GLG reserve in the dark.

This is consistent with recent observations that show cells are more susceptible to killing by a dark pulse (referred to here as “dark death”) at dawn, because they haven’t synthesized enough GLG to survive the dark period [19] (Fig. 2b, dawn corresponds to times 0 and 1 days). The probability of dark death is generally lower at dusk when the GLG level is highest (Fig. 2b, dusk corresponds to time 0.5 days). Experiments [8] also show that energy (ATP/(ATP + ADP)) is maintained at a higher level when a dark pulse is administered at dusk (time of higher GLG). This general pattern is consistent with the model, although the model uses G3P as an energy currency.

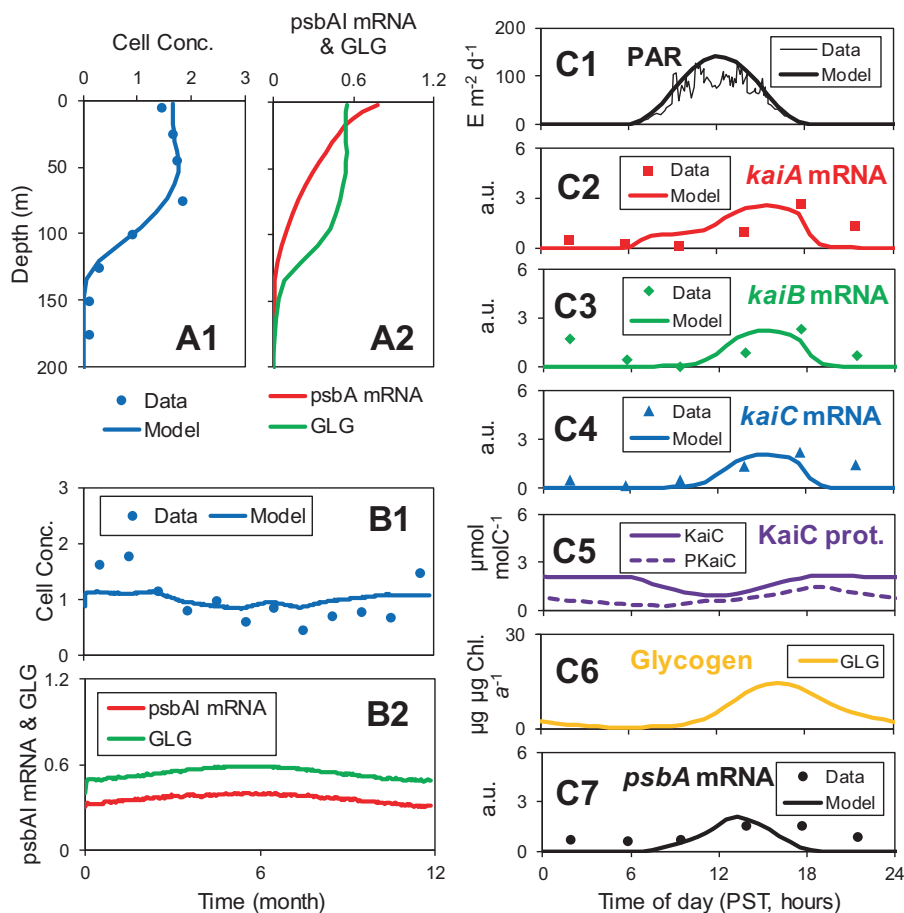
In addition to the primary pattern, which shows lower probability of death at dusk, there is a secondary minimum at ~ 0.3 days. The model reproduces this pattern, despite the lower GLG level at that time, because the clock-controlled respiration (which consumes G3P) rate is also lower at this time (illustrated in Fig. S15). The susceptibility to dark death is not only a function of the GLG level, but also how fast it is used up.

The circadian clock controls the timing of respiration and that affects the G3P level and susceptibility to dark death. Model predictions suggest that this constitutes an important fitness benefit of the circadian clock in the wild type over the no-clock mutant at the global scale, as discussed in more detail later.

The model was designed and calibrated to reproduce the observed pattern of dark death probability. First, cells that run out of G3P are killed, which results in a probability inverse to GLG (which supports G3P in the dark), i.e., the primary pattern in Fig. 2b. Second, the respiration rate was made clock-controlled, which produces the secondary minimum, as discussed above. As for the diel pattern in GLG level, this model-data comparison serves as a check and shows the model accurately represents the effect of GLG on the susceptibility to dark death.

Additional model-data comparisons are presented in the SI (see Section S2, Figs. S15–26 for additional model results and discussion). This includes GlgC and GlgP protein levels, which are relatively constant in L:D cycles [27]. Observations for the corresponding transcript levels are inconsistent [4, 28]. The clock control of GLG synthesis (Fig. 2a) is included in the model by making the GlgC enzyme activity clock-controlled. Allosteric activation of GlgC has been demonstrated [7]. The model is also compared with growth and photosynthesis rates of wild type and *glgA* and *glgC* mutants [25], which shows that photosynthesis is enhanced in the presence of GLG. This effect was suggested to be related to electron transfer efficiency [25] and is included in the photosynthesis component of the model. Overall, the model reproduces the main observed patterns of GLG metabolism in *Synechococcus*.

Fig. 3 Model-data comparison. **a, b** *Synechococcus* cell concentration, *psbA* mRNA, and GLG at Hawaii Ocean Time-series (HOT) Station ALOHA (haha.ana.soest.hawaii.edu/hot, see Fig. 1e for location). Cell concentration normalized to mean, *psbA* mRNA: 10^{-8} mol molC⁻¹, GLG: 10^{-1} molC molC⁻¹. Data for vertical profile and time series (0–200 m) are average for all available years (1990–2014). **c** *Synechococcus* transcripts, total and phosphorylated KaiC protein, and GLG levels off the coast of northern California (see Fig. 1d, e for location). Transcript levels are percent of *Synechococcus* transcripts, scaled to arbitrary units (a.u.). Model at 24 m. Data at 23 m from ref. [32]. All data except PAR are averages over two days in mid-September. See Fig. S21 for comparison for 2-day time series, and Fig. S22 for results at different location and depth. See also Movie S1 for an animation of various model outputs



Glycogen metabolism and fitness of the no-clock mutant in the lab

To investigate the fitness advantage of clock-controlled GLG synthesis we construct a model strain that has clock output functions fixed at a constant value (no clock, nc1). Driven only by light, this strain produces GLG accumulation that is rhythmic in LD and relatively constant in continuous light (LL), a pattern consistent with observations [7, 8] (Fig. 2a). The GLG level is unnecessarily high and the growth rate in L:D 12:12 is substantially lower than the wild type. The selection coefficient ($s = \text{growth rate of wild type} / \text{growth rate of no-clock mutant} - 1$) is 30%, which is of similar magnitude as the 20% difference quantified for the long period mutant (P28 [15]). The larger difference here suggests that having no clock is worse than having a slow clock.

The observations for the no-clock strains in Fig. 2a are from lab-generated mutants and the high GLG synthesis rate is unlikely representative of wild-type no-clock strains. In other words, we would expect that such a mutant (Fig. 2a, nc1) would probably acquire some secondary/compensatory mutation(s) to reduce the GLG synthesis rate if it is to survive in the wild. The laboratory knockout mutant strain

may not have acquired such mutations. To allow for a fairer comparison of the clock-controlled and light-controlled GLG synthesis strategies, we construct a no-clock strain with lower GlgC rate, which produces GLG levels comparable to the wild type (nc2) (Fig. 2a). Despite the similar GLG levels, the mutant synthesizes more GLG early in the light phase compared with the wild type. In the wild type, the clock suppresses GLG synthesis early in the light phase. In the mutant, GLG synthesis is controlled only by light and there is no mechanism to slow it down early in the light phase. This timing difference in GLG synthesis affects the relative fitness of the strains. Specifically, rapid GLG synthesis early in the light phase lowers energy, which reduces synthesis of the photosynthesis apparatus and that lowers photosynthesis overall (illustrated in Fig. S18). The net effect is that the no-clock mutant strain grows a little slower than the wild type ($s = 4.0\%$).

As for the wild type, the temporal pattern of GLG level in the mutant, i.e., synthesis in light and degradation in dark, and no rhythm in continuous light, is a direct and obvious consequence of the model design and calibration, and therefore not an emergent pattern. However, the absolute and relative growth rates (selection coefficient) are independent predictions and emergent model output.

Marine wild type and no-clock mutants

The model was developed based on observations from a freshwater strain and for the ocean application several photosynthesis parameters were recalibrated. Additional model-data comparisons for the marine strain in the laboratory, including growth rate at various light intensities [29] and GLG content vs. time in LD cycle [30] are presented in the SI (Figs. S19, S20). In addition, parameters were calibrated against observed *Synechococcus* levels at Hawaii (see next section). To allow for a fair comparison between the clock and no-clock strategies, the GlgC rate for the ocean wild type (wt3) and no-clock mutant (nc3) were optimized to yield maximum growth at Hawaii (Fig. S4 presents results of optimization).

Growth and gene expression of the wild type in the ocean

Comparison to field data shows the model reproduces many of the observed features (Fig. 3). The observed *Synechococcus* concentration at Hawaii is higher near the surface and shows a weak subsurface maximum at 75 m (Fig. 3A1). The model reproduces this feature as a result of calibration. Therefore, although the model accurately reproduces the vertical distribution of cells, this is not an emergent output. Observations show a relatively weak seasonal pattern of *Synechococcus* concentration, which is related to increased vertical mixing and nitrate availability in the winter [31]. The model population follows the satellite Chlorophyll *a*, which shows a weaker seasonal signal (Fig. S6).

Comparison to in situ observations of gene expression illustrate that the model captures the main observed pattern in *Synechococcus* clock and photosynthesis genes [32] (Fig. 3c). The observed transcript levels for the clock genes (*kaiABC*) peak around 18:00. The model predicts a slightly earlier peak and a faster decline following the peak. Observed photosynthesis gene (*psbA*) transcripts peak earlier, around 16:00. The model also predicts an earlier peak, but predicts no expression at night. Overall, the observed gene expression appears more damped than the model, which may be a consequence of the higher genetic diversity in the field compared with the laboratory experiments used to support the model development. Further, the circadian clock model was developed based on observations from experiments with freshwater *Synechococcus* and exclusively for continuous light conditions.

For the global simulation there was some recalibration of photosynthesis and glycogen metabolism components in the model (see previous section), but the circadian clock part was not changed. Although the pattern can be considered imposed, because the model was designed and calibrated to reproduce the general observed diel patterns of *psbA* and *kaiABC* gene expression, the application to a different environment and comparison to new data can be considered a validation. Overall, the model captures the main observed patterns and can reproduce the functioning of the circadian clock in the ocean.

The model also computes other parameters that were not measured in the field study [32], such as total and phosphorylated KaiC protein and GLG levels (Fig. 3C5 and 3C6), and it generates output at different locations, depths, and

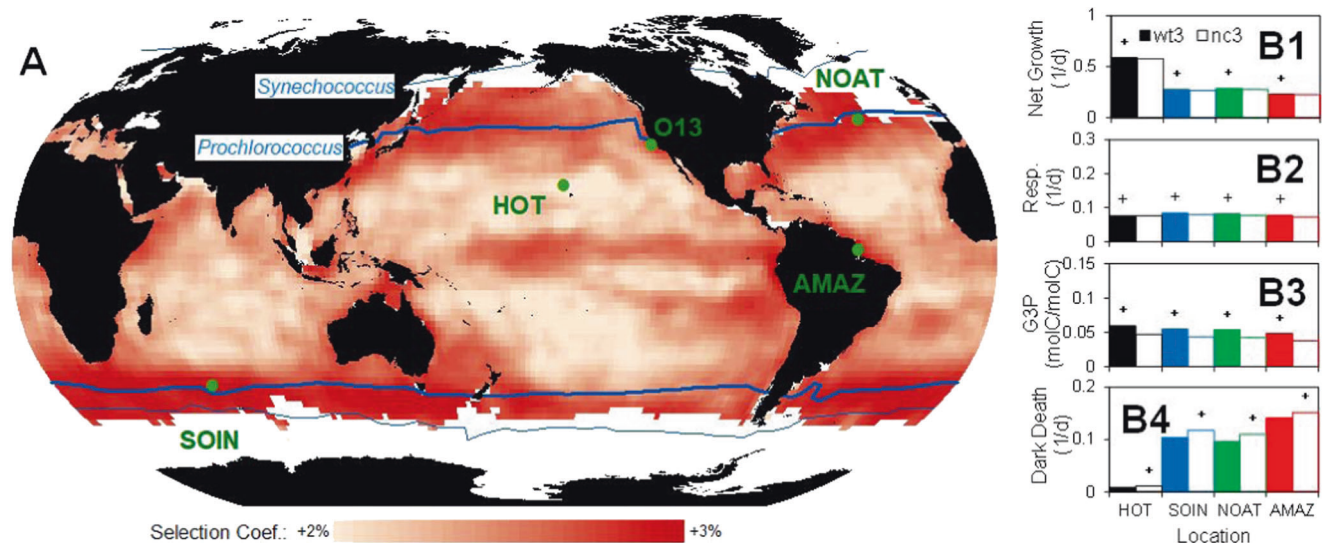


Fig. 4 Fitness benefit of circadian clock in the global ocean. **a** Global pattern of selection coefficient ($s = \text{net growth rate of wildtype [wt3]/no-clock mutant [nc3]} - 1$). Blue lines are mean annual latitudinal extent of *Synechococcus* (circadian clock) and *Prochlorococcus* (only a reduced clock) from ref. [11]. **b** Comparison of select variables at

HOT, south Indian (SOIN), north Atlantic (NOAT), and Amazon River (AMAZ) locations (see Fig. S24 for time series). Filled boxes are for the wild-type, while white boxes are for the no-clock mutant. All values are averaged over model water column and 1 year (2010) simulation period

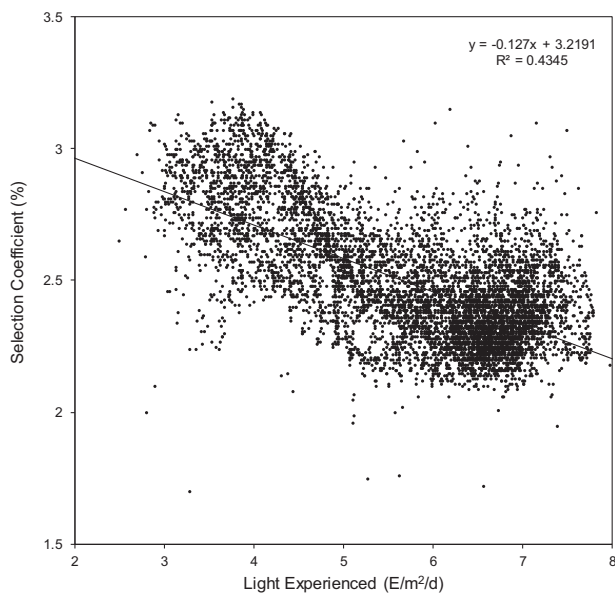


Fig. 5 Selection coefficient vs. light experienced by the population. Each point corresponds to one model grid box. Note axis scales. Some outliers (<0.1%) fall outside of plot range

times (see Fig. S22), illustrating the capability of the model to provide information and fill data gaps in various dimensions.

Spatial patterns of clock fitness benefit in the ocean

To understand the fitness benefit conferred by the circadian clock, we simulate wild type and “no-clock” strains in the global ocean model. Both strains grow well with higher growth rates at the sunnier low latitude and less turbid open ocean (Fig. S23). The wild type grows faster everywhere, which is mostly due to the diel rhythm in respiration and effect on G3P and dark death (the same mechanism responsible for the secondary minimum in Fig. 2b), as follows. In the model, respiration constitutes a loss of energy and the beneficial function of respiration is not explicitly considered (a common approach in phytoplankton ecosystem models). This means respiration is a fitness cost. The overall respiration rates are similar, but slightly lower for the mutant (Fig. 4B2). Therefore, the fitness benefit for the wild type does not arise from simply respiring less across the board. Rather, it is due to the timing of respiration, which is repressed early in the dark period by the clock in the wild type. This shifts respiration to later in the dark period and allows the cell to maintain a higher G3P level and consequently experience less dark death (Fig. 4b). The fitness benefit (selection coefficient) is generally larger at darker locations (Fig. 5), because dark death is a more significant factor there, which amplifies the fitness advantage of the wild type.

This is a testable prediction coming out of the model, and we attempted to validate this pattern experimentally, using

laboratory experiments with wild type and laboratory no-clock mutant strains. These experiments were unsuccessful, in that they showed a stronger fitness benefit at higher light intensities (Section S3, Figs. S27, 28 which presents methods and results of this experiment). These results suggest the model is incomplete, but the discrepancy may also be due to differences between the mutant strains used for the model development (Fig. 2, *kaiC*, *kaiBC*, and *cikA* mutants) and experiment (*kaiABC* mutant), optimization of GlgG rate in the ocean environment, different light regimes in the ocean and lab, and neither strain may be a good representative of the wild type no-clock strain.

Interestingly, the fitness map (Fig. 4a) shows more structure than a simple low-to-high latitude gradient. That is because a number of factors affect the light experienced by the population, including (in order of significance), incoming radiation ($R^2 = 0.78$, also lower at the cloudier equator, Fig. S11), vertical mixing ($R^2 = 0.37$, highest in Southern Ocean, Fig. S8), light extinction ($R^2 = 0.096$, higher in turbid coastal areas and to lesser extent equator and northern latitudes), and vertical advection ($R^2 = 0.046$, positive at Equator and coastal upwelling areas, Fig. S9).

The model predicts that the fitness benefit of the circadian clock is less in the low latitude open ocean, where the light level is higher. These patterns constitute a true prediction in the sense that no parameters were calibrated to achieve these patterns. Further, they are an emergent property of the model. In other words, they result from the complex interaction of the intracellular mechanisms and are not imposed or directly defined by the model design or input. The model is based entirely on laboratory experiments with simple LD cycles (i.e., L:D 12:12) and constant light intensity, which is quite different from the light pattern in the model, which varies gradually over the diel cycle and year, and horizontally and vertically. It is entirely possible, for example, that the model would have predicted a weaker fitness benefit of the clock at higher latitudes, due to summertime increased photoperiod.

Interestingly, the global pattern is generally consistent with the observed biogeography of *Prochlorococcus* and *Synechococcus*, including latitudinal extent (blue lines in Fig. 4a) and dominance in coastal vs. open ocean domains [10], as well as laboratory experiments that show that *Synechococcus* is better able to survive darkness than *Prochlorococcus* [33]. Comparing the model no-clock *Synechococcus* strain to *Prochlorococcus* is a simplification, because although the latter does not have a bona fide circadian clock, it has an hourglass timer [2, 12–14]. The simplification is necessary however, because the present model does not include the *Prochlorococcus* timekeeping mechanism (which also has not been characterized at the molecular level). Nonetheless, we believe the comparison is meaningful, because it goes in the right direction along the dimension of timekeeping

functionality, from light-induced behavior (*Synechococcus* no-clock mutant, model no-clock mutant) to hourglass timer (real *Prochlorococcus*) to bona fide circadian clock (*Synechococcus* wild type, model wild type).

Summary and outlook

The circadian clock manages energy over the diel cycle, and our study shows that this results in an important fitness benefit and that this benefit is weaker in the low latitude open ocean, which is consistent with observed marine microbe circadian ecology.

Several past models have combined gene-level and ecosystem processes [34–36], but our approach to dynamically simulate gene, transcript, protein, and metabolite levels for an entire environmental system is novel.

We investigated the role of a molecular mechanism, the circadian clock, in the fitness of cyanobacteria, and then scaled up to the global biogeography, the distribution of *Synechococcus* and *Prochlorococcus*. The scope of our study is wide and complete in the dimension of biological organization, from the molecular to the globe, but also narrow and limited in the mechanistic dimension, the role of the clock in energy management. Since the clock controls genome-wide gene expression [4, 5], we expect there to be many additional effects on cellular fitness. Even for GLG metabolism there are multiple effects, because GLG is not only an energy store, but its synthesis also serves as an electron sink during daytime [9]. In addition, there are many other factors controlling the distribution of *Synechococcus* and *Prochlorococcus*, like temperature [11]. Our choice to keep the model simple is in part based on necessity, because most other mechanisms are not as well understood as the circadian clock and its role in energy management. It was also a strategic choice, because excluding other mechanisms allows us to isolate and illustrate the effect of the included mechanism. The downside of the focus on a single mechanism is that it limits the model's ability to comprehensively represent the fitness benefits of the clock and biogeography of marine cyanobacteria. Nonetheless, our study is proof-of-concept that a molecular mechanism can be integrated into a full-scale ecosystem model and produce meaningful and intuitive results, and the same general approach can be used to explore additional mechanisms, maybe by extending the present model.

The present model only resolves a handful of mechanisms and corresponding representative genes, but the structure permits expanding it to the genome scale, for which observations are available [4, 27, 28, 32]. Genome-scale transcription and metabolism models for cyanobacteria have been developed [37–39] and some of those concepts can be adopted for this effort. The model can be extended to include other cyanobacteria and heterotrophs, which also exhibit diel

rhythms [32], and then integrated into biogeochemical models [40, 41]. Eventually, this methodology can be extended to include the dark ocean and closed biogeochemical cycles, and used to make climate change predictions. One major advantage of such a model would be that it can be informed by molecular-level observations, including gene, transcript, protein and metabolite levels, and single-cell observations [32, 42–45]. This would constitute a big step toward closing the growing gap between our models and observations and knowledge [46, 47].

Acknowledgements Jeremy Guest helped with the metabolic modeling. Erik Zinser provided helpful suggestions. Supported by NIH (MLJ and CHJ, GM107434 and GM067152). The OFES model was conducted using the Earth Simulator under the support of JAMSTEC. Two anonymous reviewers provided constructive criticism.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

1. Cohen SE, Golden SS. Circadian rhythms in cyanobacteria. *Microbiol Mol Biol Rev.* 2015;79:373–85.
2. Johnson CH, Zhao C, Xu Y, Mori T. Timing the day: what makes bacterial clocks tick? *Nat Rev Microbiol.* 2017;15:232.
3. Ouyang Y, Andersson CR, Kondo T, Golden SS, Johnson CH. Resonating circadian clocks enhance fitness in cyanobacteria. *Proc Natl Acad Sci USA.* 1998;95:8660–4.
4. Ito H, Mutsuda M, Murayama Y, Tomita J, Hosokawa N, Terauchi K, et al. Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*. *Proc Natl Acad Sci USA.* 2009;106:14168–73.
5. Hosokawa N, Hatakeyama TS, Kojima T, Kikuchi Y, Ito H, Iwasaki H. Circadian transcriptional regulation by the post-translational oscillator without de novo clock gene expression in *Synechococcus*. *Proc Natl Acad Sci USA.* 2011;108:15396–401.
6. Liu Y, Tsinoremas NF, Johnson CH, Lebedeva NV, Golden SS, Ishiura M, et al. Circadian orchestration of gene expression in cyanobacteria. *Genes Dev.* 1995;9:1469–78.
7. Diamond S, Jun D, Rubin BE, Golden SS. The circadian oscillator in *Synechococcus elongatus* controls metabolite partitioning during diurnal growth. *Proc Natl Acad Sci USA.* 2015;112:E1916–E25.
8. Pattanayak Gopal K, Phong C, Rust Michael J. Rhythms in energy storage control the ability of the cyanobacterial circadian clock to reset. *Curr Biol.* 2014;24:1934–8.
9. Welkie DG, Rubin BE, Diamond S, Hood RD, Savage DF, Golden SS. A hard day's night: cyanobacteria in diel cycles. *Trends Microbiol.* 2019;27:231–42.
10. Zwirgmaier K, Jardillier L, Ostrowski M, Mazard S, Garczarek L, Vaulot D, et al. Global phylogeography of marine *Synechococcus* and *Prochlorococcus* reveals a distinct partitioning of lineages among oceanic biomes. *Environ Microbiol.* 2008;10:147–61.

11. Flombaum P, Gallegos JL, Gordillo RA, Rincón J, Zabala LL, Jiao N, et al. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc Natl Acad Sci USA*. 2013;110:9824–9.
12. Chew J, Leypunskiy E, Lin J, Murugan A, Rust MJ. High protein copy number is required to suppress stochasticity in the cyanobacterial circadian clock. *Nat Commun*. 2018;9:3004.
13. Zinser ER, Lindell D, Johnson ZI, Futschik ME, Steglich C, Coleman ML, et al. Choreography of the transcriptome, photo-physiology, and cell cycle of a minimal photoautotroph, *Prochlorococcus*. *PLOS ONE*. 2009;4:e5135.
14. Holtzendorff J, Partensky F, Mella D, Lennon J-F, Hess WR, Garczarek L. Genome streamlining results in loss of robustness of the circadian clock in the marine cyanobacterium *prochlorococcus marinus* PCC 9511. *J Biol Rhythms*. 2008;23:187–99.
15. Hellweger FL. Resonating circadian clocks enhance fitness in cyanobacteria in silico. *Ecol Model*. 2010;221:1620–9.
16. Grimm V, Berger U, DeAngelis DL, Polhill JG, Giske J, Railsback SF. The ODD protocol: a review and first update. *Ecol Model*. 2010;221:2760–8.
17. Hellweger FL. The role of inter-generation memory in diel phytoplankton division patterns. *Ecol Model*. 2008;212:382–96.
18. Castellanos M, Wilson DB, Shuler ML. A modular minimal cell model: purine and pyrimidine transport and metabolism. *Proc Natl Acad Sci USA*. 2004;101:6681–6.
19. Lambert G, Chew J, Rust Michael J. Costs of clock-environment misalignment in individual cyanobacterial cells. *Biophys J*. 2016;111:883–91.
20. Masumoto Y, Sasaki H, Kagimoto T, Komori N, Ishida A, Sasai Y, et al. A fifty-year eddy-resolving simulation of the world ocean: preliminary outcomes of OFES (OGCM for the Earth Simulator). *J Earth Simulator*. 2004;1:35–56.
21. Hellweger FL, van Sebille E, Calfee BC, Chandler JW, Zinser ER, Swan BK, et al. The role of ocean currents in the temperature selection of plankton: insights from an individual-based model. *PLOS ONE*. 2016;11:e0167010.
22. Manizza M, Le Quére C, Watson AJ, Buitenhuis ET. Bio-optical feedbacks among phytoplankton, upper ocean physics and sea-ice in a global model. *Geophys Res Lett*. 2005;32:n/a-n/a.
23. Shirani S, Hellweger FL. Neutral evolution and dispersal limitation produce biogeographic patterns in *microcystis aeruginosa* populations of lake systems. *Microb Ecol*. 2017;74:416–26.
24. Rabouille S, Edwards CA, Zehr JP. Modelling the vertical distribution of *Prochlorococcus* and *Synechococcus* in the North Pacific Subtropical Ocean. *Environ Microbiol*. 2007;9:2588–602.
25. Suzuki E, Ohkawa H, Moriya K, Matsubara T, Nagaïke Y, Iwasaki I, et al. Carbohydrate metabolism in mutants of the cyanobacterium *Synechococcus elongatus* PCC 7942 defective in glycogen synthesis. *Appl Environ Microbiol*. 2010;76:3153–9.
26. Gründel M, Scheunemann R, Lockau W, Zilliges Y. Impaired glycogen synthesis causes metabolic overflow reactions and affects stress responses in the cyanobacterium *Synechocystis* sp. PCC 6803. *Microbiology*. 2012;158:3032–43.
27. Guerreiro ACL, Benevento M, Lehmann R, van Breukelen B, Post H, Giansanti P, et al. Daily rhythms in the cyanobacterium *Synechococcus elongatus* probed by high-resolution mass spectrometry-based proteomics reveals a small defined set of cyclic proteins. *Mol Cell Proteom*. 2014;13:2042–55.
28. Vijayan V, Zuzow R, O’Shea EK. Oscillations in supercoiling drive circadian gene expression in cyanobacteria. *Proc Natl Acad Sci USA*. 2009;106:22564–8.
29. Moore LR, Goericke R, Chisholm SW. Comparative physiology of *Synechococcus* and *Prochlorococcus*: influence of light and temperature on growth, pigments, fluorescence and absorptive properties. *Mar Ecol Prog Ser*. 1995;116:259–75.
30. Wyman M, Thom C. Temporal orchestration of glycogen synthase (GlgA) gene expression and glycogen accumulation in the oceanic picoplanktonic cyanobacterium *Synechococcus* sp. strain WH8103. *Appl Environ Microbiol*. 2012;78:4744–7.
31. Campbell L, Liu H, Nolla HA, Vault D. Annual variability of phytoplankton and bacteria in the subtropical North Pacific Ocean at Station ALOHA during the 1991–1994 ENSO event. *Deep Sea Res Part I Oceanogr Res Pap*. 1997;44:167–92.
32. Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, et al. Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc Natl Acad Sci USA*. 2013;110:E488–E97.
33. Coe A, Ghizzoni J, LeGault K, Biller S, Roggensack SE, Chisholm SW. Survival of *Prochlorococcus* in extended darkness. *Limnol Oceanogr*. 2016;61:1375–88.
34. Scheibe TD, Mahadevan R, Fang Y, Garg S, Long PE, Lovley DR. Coupling a genome-scale metabolic model with a reactive transport model to describe in situ uranium bioremediation. *Microb Biotechnol*. 2009;2:274–86.
35. Reed DC, Algar CK, Huber JA, Dick GJ. Gene-centric approach to integrating environmental genomics and biogeochemical models. *Proc Natl Acad Sci USA*. 2014;111:1879–84.
36. Hellweger FL, Fredrick ND, McCarthy MJ, Gardner WS, Wilhelm SW, Paerl HW. Dynamic, mechanistic, molecular-level modelling of cyanobacteria: *Anabaena* and nitrogen interaction. *Environ Microbiol*. 2016;18:2721–31.
37. McDermott JE, Oehmen CS, McCue LA, Hill E, Choi DM, Stockel J, et al. A model of cyclic transcriptomic behavior in the cyanobacterium *Cyanothece* sp. ATCC 51142. *Mol Biosyst*. 2011;7:2407–18.
38. Nogales J, Gudmundsson S, Knight EM, Pálsson BO, Thiele I. Detailing the optimality of photosynthesis in cyanobacteria through systems biology analysis. *Proc Natl Acad Sci USA*. 2012;109:2678–83.
39. Reimers A-M, Knoop H, Bockmayr A, Steuer R. Cellular trade-offs and optimal resource allocation during cyanobacterial diurnal growth. *Proc Natl Acad Sci USA*. 2017;114:E6457–E65.
40. Follows MJ, Dutkiewicz S, Grant S, Chisholm SW. Emergent biogeography of microbial communities in a model. *Ocean Sci*. 2007;315:1843–6.
41. Coles VJ, Stukel MR, Brooks MT, Burd A, Crump BC, Moran MA, et al. Ocean biogeochemistry modeled with emergent trait-based genomics. *Science*. 2017;358:1149–54.
42. Church MJ, Björkman KM, Karl DM, Saito MA, Zehr JP. Regional distributions of nitrogen-fixing bacteria in the Pacific Ocean. *Limnol Oceanogr*. 2008;53:63–77.
43. Hanson B, Hewson I, Madsen E. Metaproteomic survey of six aquatic habitats: discovering the identities of microbial populations active in biogeochemical cycling. *Micro Ecol*. 2014;67:520–39.
44. Ankrah NYD, May AL, Middleton JL, Jones DR, Hadden MK, Gooding JR, et al. Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J*. 2014;8:1089–100.
45. Eichner MJ, Klawonn I, Wilson ST, Littmann S, Whitehouse MJ, Church MJ, et al. Chemical microenvironments and single-cell carbon and nitrogen uptake in field-collected colonies of *Trichodesmium* under different pCO₂. *ISME J*. 2017;11:1305–17.
46. Fuhrman J, Follows M, Forde S. Applying “-omics” data in marine microbial oceanography. *Eos Trans Am Geophys Union*. 2013;94:241–.
47. Hellweger FL. 100 years since Streeter and Phelps: it is time to update the biology in our water quality models. *Environ Sci Technol*. 2015;49:6372–3.