AMERICAN SOCIETY FOR MICROBIOLOGY | Applied and Environmental Microbiology®

ENVIRONMENTAL MICROBIOLOGY

Evaluation of Genomic Sequence-Based Growth Rate Methods for Synchronized *Synechococcus* Cultures

Julia Carroll,[a] Nicolas Van Oostende,[a] Bess B. Ward[a]

[a]Department of Geosciences, Princeton University, Princeton, New Jersey, USA

ABSTRACT Standard methods for calculating microbial growth rates (μ) through the use of proxies, such as *in situ* fluorescence, cell cycle, or cell counts, are critical for determining the magnitude of the role bacteria play in marine carbon (C) and nitrogen (N) cycles. Taxon-specific growth rates in mixed assemblages would be useful for attributing biogeochemical processes to individual species and understanding niche differentiation among related clades, such as found in *Synechococcus* and *Prochlorococcus*. We tested three novel DNA sequencing-based methods (iRep, bPTR, and GRiD) for evaluating the growth of light-synchronized *Synechococcus* cultures under different light intensities and temperatures. *In vivo* fluorescence and cell cycle analysis were used to obtain standard estimates of growth rate for comparison with those of the sequence-based methods (SBM). None of the SBM values were correlated with growth rates calculated by standard techniques despite the fact that all three SBM were correlated with the percentage of cells in S phase (DNA replication) over the diel cycle. Inaccuracy in determining the time of maximum DNA replication is unlikely to account entirely for the absence of a relationship between SBM and growth rate, but the fact that most microbes in the surface ocean exhibit some degree of diel cyclicity is a caution for application of these methods. SBM correlate with DNA replication but cannot be interpreted quantitatively in terms of growth rate.

IMPORTANCE Small but abundant, cyanobacterial strains such as the photosynthetic *Synechococcus* spp. are important because they contribute significantly to primary productivity in the ocean. These bacteria generate oxygen and provide biologically available carbon, which is essential for organisms at higher trophic levels. The small size and diversity of natural microbial assemblages mean that taxon-specific activities (e.g., growth rate) are difficult to obtain in the field. It has been suggested that sequence-based methods (SBM) may be able to solve this problem. We find, however, that SBM can detect DNA replication and are correlated with phases of the cell cycle but cannot be interpreted in terms of absolute growth rate for *Synechococcus* cultures growing under a day-night cycle, like that experienced in the ocean.

KEYWORDS *Synechococcus*, iRep, bPTR, GRiD, sequence-based method, metagenomic growth rate estimator, cell cycle, growth rate

Small in size but abundant in number, picocyanobacteria, including the ubiquitous unicellular *Synechococcus* spp., are responsible for up to 25% of primary productivity in the ocean (1). These picocyanobacteria also produce oxygen and contribute fixed carbon to the microbial loop and to higher trophic-level organisms. Quantification of *Synechococcus* growth rate *in situ* is valuable for understanding its contribution to ocean productivity and as prey for micro-grazers and its response to changing sea climate (2). Quantifying the contributions of individual taxa to total productivity is critical for understanding the contribution of different species or clades to photosynthesis or trophic transfer and the environmental conditions that determine niche differentiation and the differential growth of closely related clades.

Citation Carroll J, Van Oostende N, Ward BB. 2022. Evaluation of genomic sequence-based growth rate methods for synchronized *Synechococcus* cultures. Appl Environ Microbiol 88:e01743-21. https://doi.org/10.1128/AEM.01743-21.

Editor Jennifer B. Glass, Georgia Institute of Technology

Address correspondence to Bess B. Ward, bbw@princeton.edu.

The authors declare no conflict of interest.

Received 30 August 2021
Accepted 13 October 2021
Accepted manuscript posted online 27 October 2021
Published 11 January 2022

Methods used to measure the growth rate of pure cultures in the lab, such as measurement of the rate of change of *in vivo* fluorescence or cell number, can also be employed to estimate *Synechococcus* growth rates in the field using the high resolution and specificity that is possible with flow cytometry (3). These methods require either incubations or resampling of the same water parcel over time, however, which can be very problematic (4). Time series observations with *in situ* flow cytometry have revolutionized the resolution of sampling for cell number and size in order to determine *in situ* growth rates for *Synechococcus* (5).

Incubation methods, such as the dilution grazing technique (6) or using isotopes to trace carbon or nitrogen uptake, can be made somewhat taxon specific by using size fractionation or analysis of diagnostic photosynthetic pigments (7). None of these approaches are truly species specific because closely related species can share similar pigments and flow cytometric fluorescence signatures. In addition, all incubation methods are susceptible to artifacts of enclosure and perturbation of nutrient fluxes, etc. Thus, for many reasons, it is not presently possible to determine species- or clade-specific activities (e.g., growth rate, nutrient uptake, etc.) in natural microbial populations.

Quantification of DNA replication and the cell cycle has the potential to provide a taxon-specific estimate of growth rate because information about both replication rate and species identification are contained within the same molecule, the chromosomal DNA. Using cell cycle analysis from flow cytometry, growth rate estimates were obtained for *Prochlorococcus* in the Arabian Sea by resampling the same parcel of water over the course of a day or more (4, 8). The method failed for *Synechococcus* due to difficulties with determining the maxima for S and G2 phases and ship movement, and inaccurate sampling depths resulted in inconsistent sampling of the same microbial populations. Sosik et al. (5) used a submersible flow cytometer to determine species-specific growth rate in the ocean and to limit the errors caused by advection of water masses. Combining the power of automated submersible flow cytometry with a matrix model made it possible to estimate daily division rates at the population level (2).

Direct measurement of DNA replication rates estimated from coverage of the chromosomal DNA sequence was introduced by Korem et al. (9) to calculate the growth rate of microbes in infant human guts using metagenomic sequencing. This first of the sequence-based growth rate methods (SBM) was called PTR (peak-to-trough ratio) and uses the ratio of read coverage at the origin of replication (*ori*) to read coverage at the terminus (*ter*) of replication in the circular bacterial genome to calculate growth rate. Brown et al. (10) expanded upon this technique by creating two indices, the Brown peak-to-trough ratio (bPTR) and index of replication (iRep), and were able to measure growth rates for approximately 50 different microbes in the human gut, many of which comprised less than one percent of the entire microbial community. For validation of the new method, absolute cell counts of different microbes were estimated using droplet digital PCR and compared to SBM values (10). The bPTR method calculates replication rates using complete genome sequences as opposed to iRep, which uses draft genomes. Furthermore, bPTR can also determine the origin and terminus sequences using GC skew instead of average read coverage at origin and terminus in a draft genome. Cumulative GC skew generally increases with distance from the origin, so gene order on the genome can be assigned independently of the actual sequence. These methods take advantage of the facts that bacteria possess circular chromosomes that replicate bi-directionally from the origin and that %GC generally decreases between the origin and terminus.

The growth rate index (GRiD) of Emiola and Oh (11) can estimate growth rate from genomes and metagenomic assembled genomes (MAGs) with extremely low coverage (0.2×), compared to the requirement of 5× coverage for bPTR and iRep. GRiD assumes that the origin and terminus can be identified by proximity to marker genes (*dnaA* near the origin and *dif* near the terminus) in complete genomes, which means that the gene order on the chromosome can influence the outcome. Unlike the other two SBM, GRiD generates a "species heterogeneity metric," which provides an estimate of variance in GRiD values resulting from

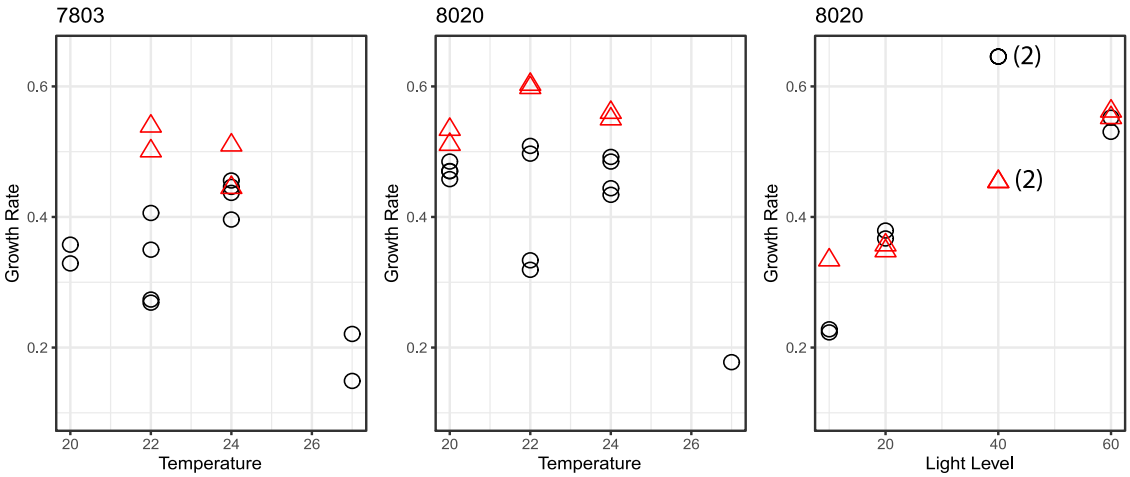

FIG 1 Temperature (°C)-controlled growth rate for strain WH7803 and strain WH8020 and light-controlled growth rate (light intensity in μmol photons m^{-2} s^{-1}) for strain 8020. (2) denotes two identical growth rate values from replicate incubations for both experiments 40 (μmol photons m^{-2} s^{-1}). Red triangles, μ_cc (d^{-1}) determined from cell cycle analysis; black circles, μ_fl (d^{-1}) determined from *in vivo* fluorescence.

the presence of closely related species. This metric may be essential for distinguishing between two ecologically important and highly related clades of *Synechococcus* in the natural environment.

SBM have the potential to quantify DNA replication at the clade level using complex metagenomic sequences, thus providing a clade-specific instantaneous growth rate estimate without incubation. The goal of the present study was to evaluate the SBM for growth rate determination in *Synechococcus* cultures that were synchronized to the diel light cycle as a preliminary to applying the methods in natural seawater assemblages. Two strains with published genomes were used in the experiments reported here, so there should be no uncertainty about the location of the origin and terminus, and it should be easy to obtain sufficient high-quality DNA and to sequence deeply enough for high coverage of the genome.

To test the SBM on *Synechococcus* cultures, the growth rates of two strains were determined using *in vivo* fluorescence and cell cycle analysis. Growth rate was controlled by temperature and light in order to investigate the relationship between growth rates determined by established methods and genome-based techniques. We hypothesized that SBM values should increase linearly with specific growth rate (μ) as measured by standard tools in synchronized *Synechococcus* cultures. Secondly, due to the circadian nature of cell replication, we hypothesized that SBM values should vary over the diel period and should be maximal in S phase, when DNA is being actively replicated. We did not attempt the SBM assays on asynchronous cultures because *Synechococcus* is always subject to some degree of diel cycle in nature, and if SBM is to be useful under environmental conditions, it must be applicable to light-synchronized cells.

RESULTS

Estimation of *Synechococcus* growth rates controlled by temperature or light. In order to obtain a range of growth rates, *Synechococcus* was grown at four temperatures under constant light and under four light levels at constant temperature. Cultures grew exponentially over the 4-day experiments (Fig. S1) and exhibited the typical "stair-step" pattern of growth during the light-dark cycle (Fig. S2). Growth rates as estimated from *in vivo* fluorescence (μ_fl) and cell cycle analysis (μ_cc) indicated optimal growth at 22–24°C for both strains (Fig. 1 left and middle panels), which is consistent with the temperatures of their environmental origin (12). Multiple experiments at 22°C produced a wide range of growth rates for both strains, such that no distinct growth rate maximum was clearly discernible. The optimum range for both strains was exceeded at 27°C. Growth rate estimates from μ_cc generally surpassed μ_fl.

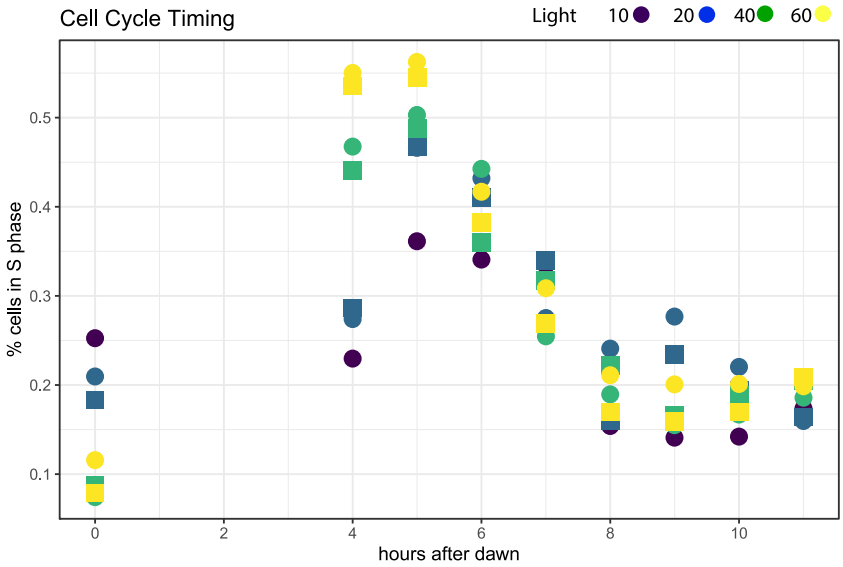

FIG 2 Variation in %S over the light period of the diel cycle in strain WH8020 grown under four light intensities on a 12:12 h light-dark cycle at 24°C. Light intensity is identified by color, and the two replicates at each intensity are indicated by shape.

Growth rate control by light was evaluated only for strain 8020 at the growth temperature of 24°C. In this experiment, μ_fl exceeded μ_cc only at a light intensity of 40 μmol photons m^{-2} s^{-1} (Fig. 1 right panel). With the exception of this high value for μ_fl at 40 μmol photons m^{-2} s^{-1}, both measures of growth rate increased with increasing light intensity. Note that in the right panel of Fig. 1, the replicates at 40 μmol photons m^{-2} s^{-1} were identical, i.e., each symbol actually represents two indistinguishable data points. Thus, both methods to estimate growth rate appear valid but we have no explanation for the discrepancy between methods. The range of growth rates obtained from a temperature range of 5°C was very similar to that resulting from a range of light intensities between 10 and 60 μmol photons m^{-2} s^{-1}.

Although μ_cc and μ_fl showed similar patterns in response to both light and temperature, the two estimates were not strongly correlated for strain 8020 and the relationship is mainly due to the greater range of μ_cc obtained in the light-controlled experiments (Fig. S4). The data set is limited because it was not possible to obtain robust μ_cc from the slowest growing cultures.

Diel cycle effects in synchronized *Synechococcus*. The synchronization of cell division to the diel cycle resulted in consistent and predictable timing of cell cycle features. Although the fraction of cells undergoing cell division varied among light levels (as expected from differences in growth rate), maximum DNA replication (maximum %S phase) occurred consistently at about 5 h after dawn (Fig. 2). Minimum %S (minimum percentage of cells in S phase) was usually observed at dawn and at the last time point assayed, 1 h before night.

The dynamic range of %S over the diel cycle (the difference between minimum and maximum %S) varies with light intensity, and thus with growth rate (Fig. S5). There is a significant (P = 0.012) correlation between maximum %S and μ_cc for all the WH8020 experiments (both temperature- and light-controlled) for which cell cycle data were available (Fig. 3). The relationship between maximum %S and μ_fl was not significant for the WH8020 experiments and maximum %S was not significantly correlated with either μ_cc and μ_fl for WH7803 alone or for the combined data set of both strains together.

The optimized DNA extraction protocol and sequencing yielded high quality DNA and metagenomic sequence data. SBM scripts generated graphical outputs from the sequence data that identified the origin and terminus of the circular chromosome to calculate the SBM values (Fig. S6). Because all the SBM depend on quantifying DNA replication in cell populations, we hypothesized that synchronization of the cell cycle to the light cycle would lead to a strong dependence of all SBM on timing of the cell

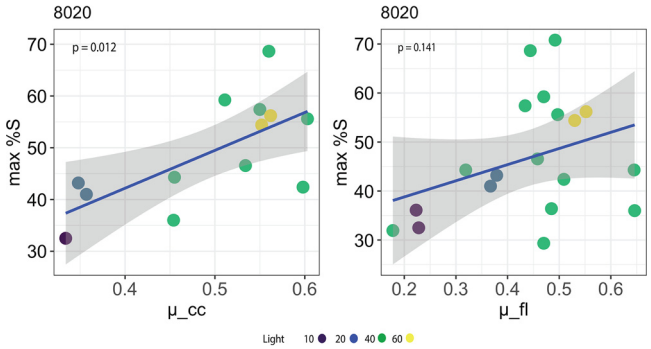

FIG 3 Relationship between measured maximum proportion of cells in S phase and the standard growth estimates for WH8020 only. The blue line and shaded area are linear regressions showing 95% confidence intervals. (See supplemental Table S1 for statistics on regressions.)

cycle. Indeed, all three SBM varied with the diel cycle and were correlated with %S, with maxima in all SBM occurring at around 5 h after dawn (Fig. 4, Fig. S7).

All three SBM were correlated with %S for light- and temperature-controlled experiments combined for strain WH8020 (Fig. 5), and iRep and bPTR were correlated with %S for both strains in the temperature-controlled experiments (Fig. S8) when samples throughout the diel cycle are considered. Despite some variability in the relationships between SBM and %S, these results validate the expected relationship between cell cycle (represented by %S) and SBM. Except for GRiD vs %S in the temperature-controlled experiments at 40 μmol photons m^{-2} s^{-1} (Fig. S8), the relationships are significant but the low *R*-squared value (Table S1) implies that the cell cycle does not explain all the variability in the SBM. The source of variability in the correlations is more likely due to the %S measurement than the SBM because of the difficulty in precisely timing the sampling of maximum DNA replication.

Relationship between SBM and cell cycle. The variation in the dynamic ranges of maximum %S (Fig. S5) and in the absolute maximum %S values observed at different growth rates (Fig. 4) yielded a significant correlation between maximum %S and μ_cc (P = 0.012) but not between maximum %S and μ_fl (P = 0.141) (Fig. 3), based on the time points at which maximum %S occurred.

The variation in maximum %S should be captured in the sequencing coverage data upon which the SBM are based. Only iRep was significantly correlated with maximum %S (P = 0.023) for WH8020 (Fig. 6). None of the SBM were significantly correlated with maximum %S for the combined data set for both strains (Fig. 6).

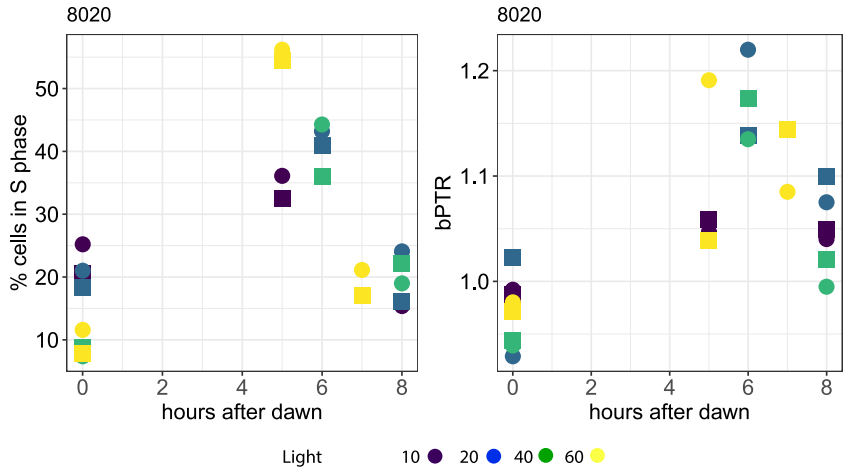

FIG 4 Variation in %S and bPTR, one of the SBM methods, with time of day for the light-controlled experiment with strain WH8020, for time points at which SBM values were measured. Light intensity is identified by color, and the two replicates at each intensity are indicated by shape.

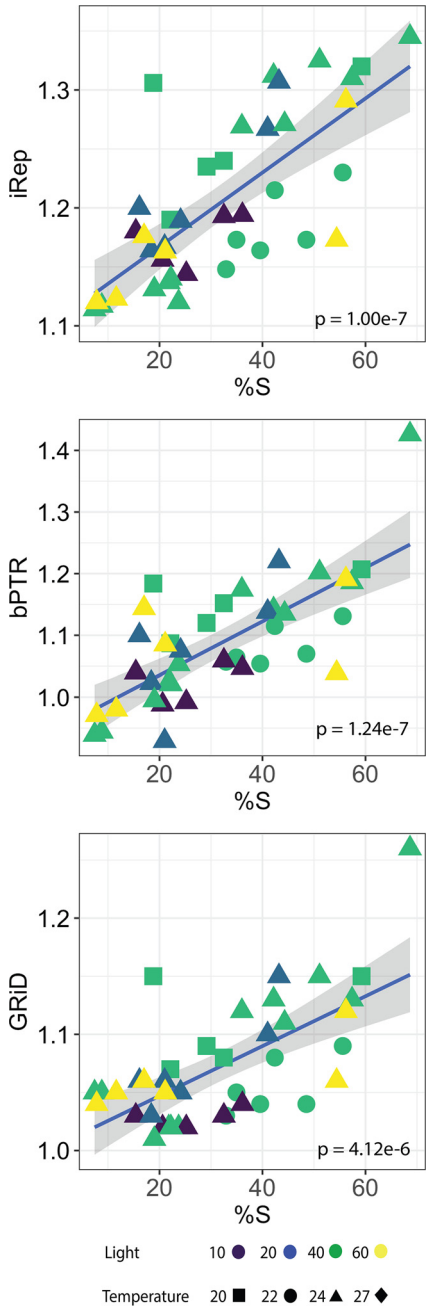

FIG 5 Relationship between individual SBM and the percentage of cells in S phase over the diel cycle for strain WH8020 at four light levels and three temperatures. The blue line is the linear regression for the statistics (supplemental Table S2), and the shaded area represents the 95% confidence interval around the line.

Relationships between SBM and standard estimates of growth rate based on *in vivo* fluorescence and cell cycle. The relationships between SBM and μ_fl and μ_cc were evaluated for temperature- and light-controlled experiments both separately and together. Due to the lack of μ_cc data from some of the temperature-controlled experiments, the comparisons between SBM and μ_fl are more robust, but neither μ_fl nor μ_cc is correlated with any of the SBM (Fig. 7 and Fig. S9). The only pattern is that all three SBM are relatively invariant at lower μ_fl. At higher μ_fl, the SBM are all more variable: the highest values of an SBM occur at higher growth rates, but the entire range of SBM values occurs at medium to high growth rates. This variability in SBM under apparently invariant conditions is clear when SBM are plotted against temperature or light intensity (Fig. S10). All three SBM show greatest variability and range under optimal conditions.

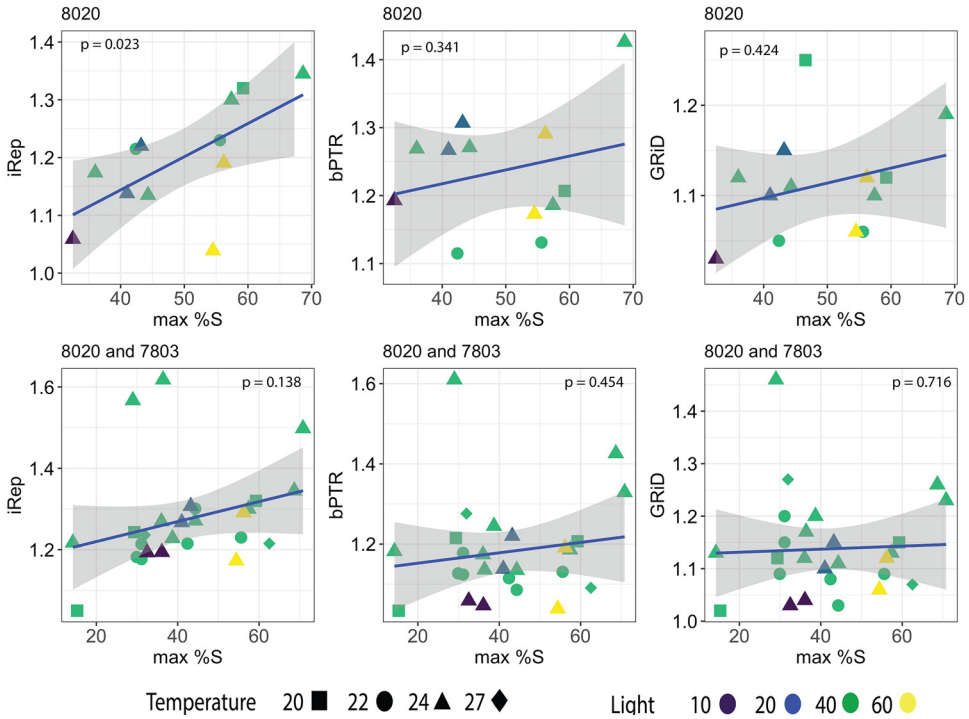

FIG 6 Top: Relationships between SBM and max %S values for 8020 light-controlled experiments. Bottom: Relationships between SBM and max %S values for 7803 temperature- and 8020 temperature- and light-controlled experiments. (See supplemental Table S3 for statistics on regressions.)

Thus, the last relationship—the hypothesis that SBM should be correlated with standard measures of growth rate—does not hold, and SBM are poor predictors of growth rates as estimated by either μ_fl or μ_cc.

Relationships among SBM. The three SBM were highly correlated with each other, especially for the temperature-controlled experiments, and the slopes varied slightly between the two strains (Fig. 8). The correlations were independent of temperature;

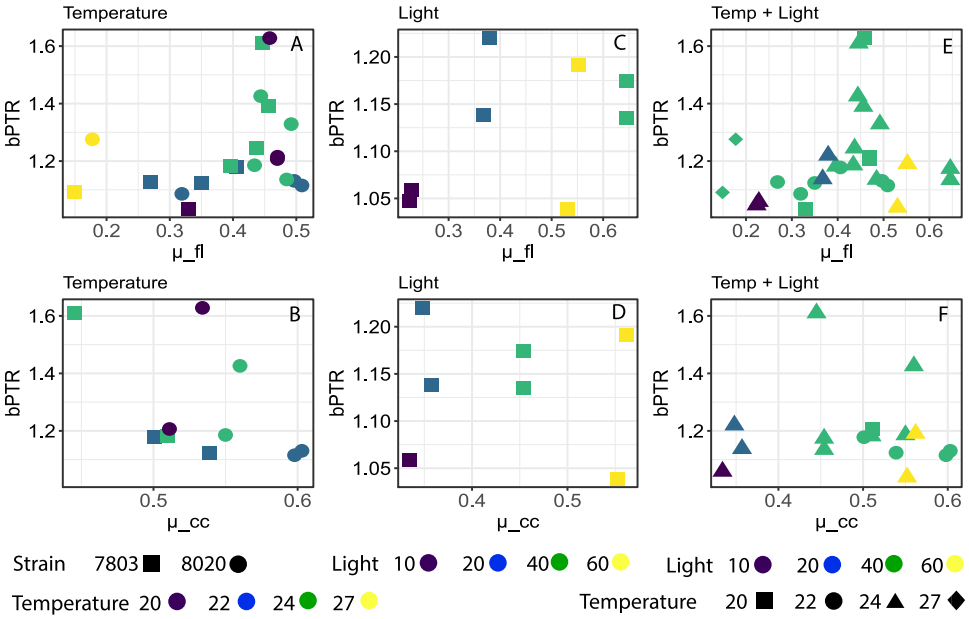

FIG 7 Relationships between bPTR and growth rates estimated by standard measures. A, B: bPTR vs μ_fl or μ_cc for temperature-controlled experiments. C, D: bPTR vs μ_fl or μ_cc for light-controlled experiments. WH8020 only. E, F: bPTR vs μ_fl or μ_cc for temperature- and light-controlled experiments.

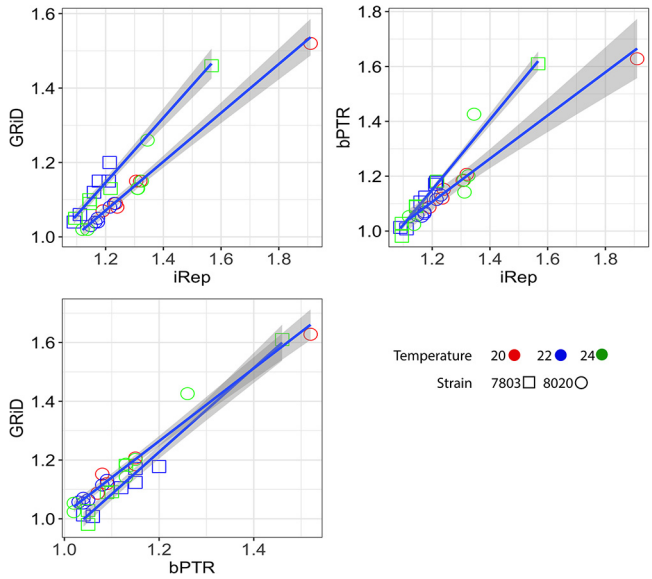

FIG 8 Correlations between the three SBM for each strain from the temperature-controlled experiments. *P* values for all lines are $\ll$ 0.001. (See supplemental Table S4 for statistics on regressions.)

i.e., all three growth temperatures fell on the same line. The tight correlations among the different SBM values suggest that they are all measuring the same thing and provide robustness to their absolute values. Nonetheless, the fact that none of them show significant correlations with standard measures of growth rates μ_fl or μ_cc (Fig. 7) makes the interpretation or ecological significance of an SBM, beyond its correlation with the proportion of cells replicating their DNA, unclear.

DISCUSSION

Could SBM values be a useful proxy for instantaneous growth rate in marine unicellular cyanobacteria such as *Synechococcus*? Here we evaluate each step in the investigation of the potential for SBM as indexes of instantaneous growth rate in synchronous *Synechococcus* cultures: 1) validation of growth rate dependence on the fundamental environmental factors, light and temperature, using the standard growth rate estimates, μ_fl and μ_cc; 2) evaluation of the effect of the diel cycle on growth rate and cell cycle indicators (related to %S); 3) relationship between cell cycle and SBM; and 4) relationships between the SBM and standard growth rate estimates.

(i) Standard growth rate estimates. Both *Synechococcus* strains responded to varying either temperature or light by modulating their growth rate. Both μ_fl and μ_cc were reproducible between replicate cultures, but μ_cc values generally exceeded μ_fl in the temperature-controlled experiments. The variability in μ_fl for the temperature-controlled experiments for both strains at 22°C is unexplained (Fig. 1). The 22°C data represent two different sets of replicated experiments on different days and cannot be explained by small variability in light due to location in the incubator (n = 12, F = 2.69, P value = 0.12). Similarly, the disparity between the two growth rate estimates, μ_fl and μ_cc, at 40 μmol photons m^{-2} s^{-1} in the light-controlled experiment for strain 8020 is unexplained but supported by excellent replication of the individual estimates. A total range of growth rates from $<$0.2 to $>$0.6 per day with an optimum at 22–24°C and 40–60 μmol photons m^{-2} s^{-1} is consistent with the ecological origin and previous characterization of these strains (12). The light optimum for the growth rate of strain WH8020 is reported to be around 40 photons m^{-2} s^{-1} (13). Therefore, we conclude that the standard estimates of growth rate reported here represent an adequate data set for comparison with the recently proposed SBM.

(ii) Diel cycle effects. Because the cultures were synchronized to the diel cycle, it was necessary to sample at the same time every day in order to remove the diel pattern in fluorescence (Fig. S2) to obtain a reproducible measure of μ_fl. Previous characterizations of

the cell cycle in *Synechococcus* had shown that synchronization to the light cycle resulted in predictable patterns in DNA replication for several strains of *Synechococcus* in both field samples and pure culture experiments with light-dark cycles of 12:12 and 14:10 h (14). The timing of division for WH7803 reported by the studies cited by Jacquet et al. (14) varied from midday to just after the beginning of the dark period. In the original report on WH7803 synchronized to a 14:10 h light-dark cycle, Waterbury et al. (15) observed a highly reproducible peak in percent dividing cells (determined by epifluorescence microscopy) after 10 h in the light. DNA replication has been completed prior to division, which is consistent with a maximum %S earlier in the day. Both light- and temperature-controlled experiments reported here detected a strong diel signal in DNA replication, best represented by the maximum %S, which occurred 5–6 h after dawn (Fig. S3). The pattern of %S over the diel cycle was fundamentally the same for all growth rates in the light-controlled experiment, and the dynamic range of %S (i.e., the difference between minimum and maximum %S) was least at the lowest light and greatest at the highest light level treatment (Fig. S5). The positive correlation between μ_cc and maximum %S was significant for WH8020 (Fig. 3) but probably somewhat weakened by uncertainty over the timing and magnitude of maximum %S; our discrete hourly sampling could not precisely detect the exact timing of peak DNA replication.

(iii) Relationship between SBM and cell cycle. The diel pattern in %S represents the pattern in the timing of DNA replication and therefore dictates that the SBM should also vary with time of day. All three SBM were correlated with %S in samples collected over the diel cycle (Fig. 5). Thus, we conclude that the SBM reflect some measure of DNA replication at the population level, which would imply that the SBM should be related to growth rate.

One attraction of the SBM is that they should provide an instantaneous estimate of growth rate. Given the variation in DNA replication (%S) over the diel cycle, the only time that SBM are likely to capture the growth rate accurately is at the time of maximum DNA replication or maximum %S of the synchronized cells. Only iRep was significantly correlated with maximum %S and only when WH8020 was analyzed alone (Fig. 6). The fact that including WH7803 in the analysis weakens the relationship between the SBM and maximum %S is further evidence for the importance of timing. The WH7803 cultures were sampled with less frequency than the WH8020 cultures and were thus less likely to have captured the time of maximum %S accurately.

Alternatively, the dynamic range of population %S might be quantitatively related to growth rate, but determining that would require multiple times points, thus erasing the advantage of instantaneous growth rate estimation.

The three SBM applied here rely on slightly different calculations, but all depend on the bidirectional pattern of replication of the circular bacterial chromosome. iRep is most correlated with %S, perhaps because of the reference genome advantage, but this is both good and bad news. If a perfect reference genome is required for the analysis, then use in field samples is very limited due to the plethora of highly similar but nonidentical sister clades of common microbes such as *Synechococcus*, *Prochlorococcus,* and SAR11.

GRiD is advantageous for complex samples when using an assembled genome from a natural assemblage having lower depth of sequence coverage. After mapping the reads to a given genome sequence, the GRiD algorithm orders the contigs from high to low coverage and then splits the contigs into two temporary files, with one of the files re-sorted from low to high coverage. These sorted, reordered contigs are then concatenated and coverage across each nucleotide is derived. After filtering for outliers, a smoothing curve is then fitted by a redescending M estimator with Tukey's bi-weight function. The original, physical order of the genes may be lost in the process. iRep also reorders the coverage windows using a linear fitting method on $\log_2$ coverage values. Only bPTR uses the full gene order information of the genome (and adds a visual verification of the genome location of *ori* and *ter* based on the transition in cumulative GC bias).

These differences among the SBM algorithms may be responsible for their slightly different relationships with %S, μ_cc and μ_fl. The fact that they all correlate strongly with each other (Fig. 8) but not with growth rate suggests a fundamental gap in translating the genome replication process into a growth rate metric.

(iv) Relationships between the SBM and standard growth rate estimates. Given the breakdown in the relationship between DNA replication and SBM, it is thus not surprising that the SBM also did not correlate with the standard measures of growth rate, μ_cc and μ_fl. According to the model of bidirectional replication of a circular chromosome (16), an SBM value of 1.0 signifies no DNA replication. An SBM value of 1.5 would mean that half of genomes were being actively replicated. The range in absolute values of the SBM was small (factor of ~2) compared to the absolute range in growth rate estimated from μ_fl (factor of ~4.3). The range for μ_cc was small (1.8), but that is explained by the fact that we were not able to obtain cell cycle data for all cultures growing at the lowest rates (<0.3 d^{-1} at 20°C or 27°C or 10 μmol photons m^{-2} s^{-1}), whereas SBM values were obtained for most of the samples covering the whole range of growth rates. None of the SBM values showed a clear relationship with either temperature or light, although all of them exhibited the greatest range at the temperature (24°C) and light (40 μmol photons m^{-2} s^{-1}) expected to be optimal for these two strains (Fig. S9). Note, however, that many more samples were collected under these optimal conditions than on either side of the optima, so we cannot conclude that a similar range might not have occurred with more measurements under all growth conditions.

Based on work with freshwater *Synechococcus* strains *S. elongatus* 7942 and *Synechoscystis* 6803, Watanabe (17) generalized that asynchronous DNA replication, in which multiple copies of the chromosome in the same cell do not initiate replication at the same time, is a common feature of polyploid *Synechococcus*. WH7803 is polyploid (18–20), with an average of ~4 genomes per cell, while as far as we know, WH8020 is monoploid. As was also observed for WH8020, the diel maximum %S in our WH7803 cultures does not support the occurrence of asynchronous DNA replication. Higher growth rates might still be manifest in higher average *ori* to *ter* coverage, and that signal should vary with growth rate but not with time of day. For six experiments with WH7803, in which flow cytometry data were collected at three times during the light period, there was a clear variation in %S with time of day (Fig. S11). Maximum %S occurred at 6–8 h after dawn, which is very similar to the timing of maximum %S observed for WH8020, for which higher resolution sampling was performed. This pattern suggests that even though WH7803 is polyploid (which we did observe in the flow cytometry analysis; only the most abundant S and G2 cell populations with the lowest DNA content were used in μ_cc calculations), it was entrained into synchronous DNA replication. Binder and Chisholm (19) documented multiple peaks in DNA replication associated with different cohorts of cells in one culture, which are at different stages in the cell cycle at different times. Nevertheless, a clear diel pattern occurred in the timing of cell division and the maximum %S for the dominant cohort. It appears that at least for WH7803, polyploidy does not imply asynchronous DNA replication under light-dark cycling.

Watanabe (17) also observed that polyploid *Synechococcus* clades tended to have lower GC skew than monoploid clades, which could make it difficult to identify *ori* and *ter* locations in applications in which those locations are detected on the basis of GC skew. WH7803, however, has a relatively high GC skew, comparable with that of monoploid clades WH8101 and WH8103 (17) and its sister clade, WH7805 (21). Thus, polyploidy does not appear to be an issue in the application of SBM to determination of growth rate in WH7803, nor a source of differential results between WH7803 and WH8020.

Applications of SBM in the environment. Despite its inability to measure growth rate, SBM is informative about DNA replication and might be informative about viability or even potential growth rate, based on the strong relationship between the SBM and percentage of cells in S phase. While it is much easier and faster to evaluate %S from shipboard flow cytometry than to determine SBM values from sequencing, an SBM has the advantage that it may be clade specific. Flow cytometry on field samples cannot reliably detect different populations of closely related clades, but clade level resolution should be possible for SBM.

Prior to next generation sequencing, attempts to exploit patterns in DNA replication to estimate growth rate in both laboratory and environmental settings employed cell cycle analysis based on flow cytometry. Cell division by *Synechococcus* in a variety of marine environments appears to be synchronized to the diel cycle and occurs between 6 and 15 h after

dawn, predominantly within 3 h either side of sundown (14). The timing of the cell cycle in *Synechococcus* is much more variable than in natural populations of *Prochlorococcus* (22), which could make it difficult to capture the most informative sample for analysis of SBM values from natural samples. Flow cytometric analysis could be used to screen samples prior to application of the SBM.

A strong diel pattern in cell division imposes constraints on sampling for the practical determination of the SBM, so it might be supposed that heterotrophic microbes, which are not directly linked to the diel cycle through a direct dependence on photosynthesis, might be easier targets for SBM. Perhaps not surprisingly, however, many surface water microbes exhibit diel periodicity in gene expression (e.g., see references 23 and 24). This gene expression activity probably results from the dependence of heterotrophs on the production of fresh organic matter by autotrophs, perhaps mediated by grazing. This linkage likely means that the cell division of heterotrophs in surface waters is synchronized to some degree with the light cycle. It might be possible to disentangle diel cycle linkages between clades of different trophic levels using multiple temporal sampling points and applying an SBM. Deeper populations might be uncoupled from the light cycle, but their growth rates are probably lower due to a lower supply/production rate and/or quality of organic food source, requiring high sensitivity for detection of growth rate. Digging further into the sequence data, we uncovered three heterotrophic bins in the WH8020 samples and one heterotrophic bin in the WH7803 samples. These numbers of heterotrophic reads present in each sample varied from 0.6% to 23.6% of the total reads. We calculated iRep values for reads mapped to heterotrophic bins to see if any pattern existed between these values and the *Synechococcus* iRep values. Although the majority of the heterotrophic iRep values were higher than those from *Synechococcus*, there was no discernible pattern between the two; furthermore, there was no pattern between heterotrophic iRep values and time of day. This lack of informative patterns in the heterotrophic draft genomes most likely stems from the fact that the DNA samples were all collected in the mid-afternoon when heterotrophic organisms are expected to be active, in response to activity by their photoautotrophic counterparts (24).

Long et al. (25) compared SBM values to an independent measure of growth rate determined from relative abundances of metagenome assembled genomes (MAGs) of mostly heterotrophic microbes in long incubations of natural marine assemblages in the dark. They found that SBM were correlated with growth only for the fastest growing strain in the assemblage and that all other strains (24 total) showed no relationship, or even negative correlations, with the independently measured growth rate. The incubations of Long et al. (25) were several days long in the dark, so the complicating factor of the diel cycle might have been removed, or at least reduced, in their analysis. On the other hand, since surface populations are normally linked to the diel cycle, long incubations in the dark cannot be representative of *in situ* growth rates and activities.

The most prevalent use of SBM in environmental applications appears to be attempts to determine whether microbial populations are alive or dead, and to obtain a relative measure of growth rate, rather than attempting to interpret the results in terms of absolute rates. Schulze-Makuch et al. (26) used iRep to argue that microbes represented by metagenomic bins from transiently wetted Atacama desert sites were replicating (iRep >1), thus arguing for endemic communities in this extremely dry environment, rather than relict cells delivered by atmospheric transport. High quality metagenome assembled genomes (MAGs) from rice paddies, representing novel nitrite oxidizing bacteria that contained sulfur cycling genes, were estimated to be actively dividing based on iRep values $>$ 1 (27). Zhao et al. (28) concluded that 32% of the anammox bacterial population represented by a MAG recovered from 80,000-year-old subsea sediments were actively replicating *in situ*. In these studies and others, (e.g., references 29 to 31), iRep values between 1 and 2 are most commonly reported. The *Synechococcus* cultures investigated in the present study had iRep values between 1 and 2, representing much less range than detected in the independently measured growth rates. Long et al. (25), however, detected iRep values between 2 and 5 for MAGs whose

growth rates were independently determined to be between zero and four per day, and the fastest growing clade had among the lowest measured iRep (and other SBM) values.

Given the lack of correlation between SBM values and independently determined growth rates both for cultures and natural assemblages, it seems unrealistic to assign quantitative significance to absolute SBM values. There is also little basis for robustly concluding that an index of replication (computed by any of the SBM) $>$ 1 actually implies active replication. It is also difficult to evaluate most of the reported SBM values because it is impossible to obtain any estimate of the error or reproducibility associated with one off measurements from a metagenome. The variability in our SBM data from actively growing *Synechococcus* cultures under highly controlled conditions suggests caution in quantitative interpretation of SBM data obtained from single samples.

In contrast, changes in SBM values with time or between samples and stations for the same strains or MAGs within a population context might be informative. It might be possible to detect differences in SBM related to DNA replication based on the relationship with percent S shown in this study, but ground truthing and replication would be needed for robust interpretation.

Conclusions. Growth rates of two *Synechococcus* strains varied in response to light and temperature and were reproducibly estimated by standard growth estimates that rely on *in vivo* fluorescence and flow cytometry measurements of cellular DNA replication. Percentages of cells in S phase varied over the diel cycle and were correlated with the SBM in paired samples, but that relationship did not translate into the ability to estimate growth rate from a single SBM sample collected at the time of maximum %S, where we expected the strongest and most consistent SBM signal. The lack of correlation between the SBM and growth rate may be due to this dependence of the SBM on %S and the failure to sample precisely at the time of maximum %S. Alternatively, even a synchronized cell population may contain enough variability in timing of maximum %S that the relatively small range of SBM values obtained here was not able to differentiate within the distribution of %S among cultures growing under slightly different conditions. We conclude that the SBM is not a robust predictor of growth rate for cells that are synchronized to the diel cycle, despite the 4-fold range of growth rates presented in this study (0.15–0.65 d^{-1}). Since most microbial populations in the surface ocean are synchronized to the diel cycle to some extent, either directly via the light driven circadian rhythm or indirectly based on chemical signals from photosynthetic microbes (24), this finding makes it very unlikely that the SBM value can be interpreted as a growth rate in meaningful ecological ways for most marine microbes. Deeper dwelling populations may be uncoupled from the diel cycle but probably grow much more slowly, thus presenting a challenge in detecting the very small range of SBM values likely to be encountered.

SBM values did nonetheless vary significantly over the diel cycle and did correlate with cell replication, so there might yet be a way to isolate a useful signal from SBM values applied to bacterial populations with a diel cell cycle synchronization. The greatest range in SBM values occurred at growth rates associated with optimal conditions, meaning the highest SBM values occurred near the highest growth rates, so it might be valid to associate high SBM values with high growth rates and non-zero DNA replication rates. The lowest SBM values are also detected at the same growth rates, however, so it is clear that factors other than growth rate must be in play.

MATERIALS AND METHODS

Strains and stock culture maintenance. *Synechococcus* strains WH7803 and WH8020 were obtained from the Woods Hole Culture Collection and maintained at 24° C in sterile 125 ml Erlenmeyer flasks containing 50 ml of SN medium (32). Stock cultures were grown in a Percival incubator under a 12:12 h light-dark cycle at a light level of approximately 45 μmol photons m^{-2} s^{-1}.

Prior to conducting a new experiment, 50 ml stock cultures were acclimated to a particular temperature or light condition for a minimum of 3 days and then inoculated into larger volumes (360–370 ml) of SN medium. Two hundred milliliters of this culture were used to inoculate two replicate 2.8 L Fernbach flasks (total volume 1.5 L) for each *Synechococcus* strain, which were incubated with shaking at 40 μmol photons m^{-2} s^{-1}.

Temperature-controlled growth rate experiments. Four growth temperatures (20°, 22°, 24°, and 26.5° C) were selected based on the isolation temperatures and thermal amplitudes found in Pittera et al. (12): 26.5° C ($\pm$ 5° C to 6.9° C) for WH7803 and 17.5° C ($\pm$ 7° C to 8.9° C) for WH8020. Growth

was monitored by *in vivo* fluorescence over 4 days. Samples were taken at dawn and 6 h post-dawn on the first, third, and fourth days of the experiment targeting the exponential growth phase. More frequent samples were collected on the second day of the experiment at 1–2 h intervals from dawn to 8 h post-dawn for *in vivo* fluorescence and cell cycle analysis (1 ml each) and at dawn, 5—6 h, and 10 h after dawn for DNA extraction (~250 ml).

Light level-controlled experiments. *Synechococcus* WH8020 cultures were used for the light-controlled growth rate experiments. To control the amount of irradiance reaching the cultures, mesh coverings or lab bench paper was used to shade the flasks. Cultures were prepared in 2.8 L Fernbach flasks as for the temperature-controlled experiments and incubated in replicate at each of four different light levels: 10, 20, 40, and 60 μmol photons m^{-2} at 24° C (selected based on results from reference 13 showing optimal growth at 40 μmol photons m^{-2}). Light level in the culturing medium was monitored every other day using a light meter (Walz ULM-500 universal light meter and data logger) at the height of the flasks.

A total of 31 samples were collected for *in vivo* fluorescence and cell cycle analysis over a period of 3–4 days or approximately 2–3 cell generations. Samples were taken at dawn and 6 h post-dawn on the first, third, and fourth day of the experiment and at dawn and every hour between 4 and 11 h post-dawn on the second day of the experiment. Aliquots of the cultures were filtered for subsequent DNA extraction at dawn, 5—6 h, and 10 h post-dawn.

Determination of cell cycle timing in cultures. The light-control experiment was used to characterize the timing of the cell cycle in synchronized cultures during early exponential growth using the samples for *in situ* fluorescence, cell cycle analyses, and DNA collected over the 12-h light period of the second day.

Growth rate determination. Two standard methods were used to quantify growth rates: *in situ* fluorescence (μ_fl) and cell cycle analysis (μ_cc). Chlorophyll fluorescence was measured *in vivo* using a Turner Designs fluorometer using the chlorophyll *a in vivo* module and 1.5 ml cuvettes containing aliquots of *Synechococcus* cultures over the course of 4 days during the temperature- and light-controlled experiments above. The specific growth rate (μ_fl, per day [d^{-1}]) was calculated using the expression: $\frac{\ln(f_2/f_1)}{t_2-t_1}$, where fluorescence values (f_1 and f_2) were measured at time *t*. *R*-squared values for the growth curves of fluorescence vs time averaged > 0.95, implying a 95% confidence interval of $\pm$ 10% or less around the estimated μ_fl. Flow cytometry was used for cell cycle analysis using the BD Accuri C6 Flow Cytometer. Cells were fixed using a final concentration of 2.5% vol/vol glutaraldehyde, flash frozen with liquid nitrogen, and stored at −80°C until analysis. Cell aliquots (500 μL, diluted as necessary to optimize event detection) were stained using 12 μL of 1:100 Sybr green I and 15 μL of 1 M sodium citrate, following an RNase A treatment (33). The flow cytometry settings for cell cycle analysis were set to a core size of 5 μm, at a flow rate of 10 μL/min, and run for the collection of at least 80,000 events. Forward angle light scatter (FALS) and fluorescence of chlorophyll or phycobilin pigments were used to differentiate *Synechococcus* cells from other/heterotrophic bacteria in the cultures and to reproducibly create gates to constrain the two different *Synechococcus* strains. The relative DNA content of the *Synechococcus* cells was detected using the fluorescent signal of the Sybr green-stained cells after exclusion of cell doublets. For cell cycle analysis, the SL CL S0 model in the FCS Express RUO version 7 software was used to determine percentages of cells in each cell cycle phase (G1, S, and G2). Growth rate (μ_cc, d^{-1}) was calculated using the following expression:

$$\mu = \frac{1}{(t_S + t_{G2M})n} \sum_{j=1}^{n} \ln\left[1 + f_s(t_j) + f_{G2M}(t_j)\right]$$

where f_s and f_{G2M} are the fraction of cells in S or G2 phase, respectively; t_j is the time of obtaining the j^{th} sample; *n* is the number of samples; and $(t_s + t_{G2M})$ is the time at which maximum f_s and f_{G2M} were measured (3). The accuracy of the μ_cc calculation depends on the frequency of sampling; the accuracy of estimating $(t_s + t_{G2M})$ cannot exceed 1/2 the frequency of the sampling interval. For cells synchronized to the diel cycle, it was important to sample at the same time in the cell cycle every day and to include the time when the greatest number of *Synechococcus* cells were in S phase (maximum %S). Due to equidistant timing requirements of the growth rate equation from cell cycle (3), sample time points were selected relative to the number of hours before or after dawn (0900).

DNA extraction. Culture aliquots (225—250 ml) were collected by vacuum filtration onto 0.4 μm pore size 47 mm Nuclepore Track-etch Membrane filters on the second day of each experiment. The filter was folded, placed in a cryotube, flash-frozen with liquid nitrogen, and stored at −80°C until extraction. The filter paper was thawed on ice and cut into 4–5 small pieces. Cells were lysed by bead-beating in a 2-ml cryotube using zircon beads twice for 30 s each. DNA extractions were performed using the CTAB method (34), which involves the use of lysozyme, proteinase K, and phenol-chloroform-isoamyl alcohol. The protocol for DNA extraction was optimized to achieve higher genomic yield by proportionally increasing the quantity of each reagent by 10 to 15% over the original recommended amount. DNA extracts were treated with RNase, heated at 40°–45° C to dissolve the DNA, and quantified using the NanoDrop 2000 UV-Vis spectrophotometer.

Calculation of SBM growth rate proxies. DNA extracts were sequenced using next generation Illumina sequencing at Princeton University's Genomics Core Facility. Libraries for genomic sequences were created using the Apollo 324 Robot, and HiSeq sequencing generated short 150 nucleotide reads. The sequences were quality controlled in the Galaxy v.19.09 workflow system using the standard quality reports from FastQC v.0.11.9 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/), the adapter sequences trimmed using Trim Galore! v.0.6.5 (https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) (with the following settings: trim low-quality ends, 0; overlap with adapter sequence, 3; maximum error rate, 0.1; discard reads

with a length shorter than 50 nucleotides), and the resulting trimmed sequences were quality filtered using the "Filter FASTQ reads by quality score and length" (Galaxy version 1.1.1) tool (90% of bases in sequences with a quality cutoff value of at least 30). The trimmed and quality-filtered sequences were screened again using FastQC v.0.11.9, as above. After quality control the sequences were mapped to the published *Synechococcus* genome for each strain (WH7803, https://www.ncbi.nlm.nih.gov/assembly/GCF_000063505.1/; WH8020, https://www.ncbi.nlm.nih.gov/assembly/GCF_001040845.1) using Bowtie2 (35) with default parameters.

The bPTR and iRep scripts (10) identified and generated a visual representation of the origin and terminus sequences based on depth of sequence coverage of the genomic reads on the *Synechococcus* strain's genome sequence. GRiD method v1.3 (source of the script, https://github.com/ohlab/GRiD) was also applied to the complete genomes (11).

Data availability. Sequences used for the SBM calculations are found at https://www.ncbi.nlm.nih.gov, accession number PRJNA795964.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1.6 MB.

ACKNOWLEDGMENTS

We thank undergraduate interns Elisabeth A. Juescher and Srishti Khetan for their help with data collection; Christina DeCoste and Katherine Rittenbach at Princeton University's Flow Cytometry Resource Facility; and the staff at Princeton's Genomics Core Facility for their assistance with cell cycle analysis and DNA sequencing. We greatly appreciate Frederica Valois' generosity in providing the cultures.

This research was supported by NSF EAGER award # 1747511 to B.B.W. and N.V.O.

REFERENCES

1. Flombaum P, Gallegos JL, Gordillo RA, Rincon J, Zabala LL, Jiao NAZ, Karl DM, Li WKW, Lomas MW, Veneziano D, Vera CS, Vrugt JA, Martiny AC. 2013. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. Proc Natl Acad Sci U S A 110:9824–9829. https://doi.org/10.1073/pnas.1307701110.
2. Hunter-Cevera KR, Neubert MG, Solow AR, Olson RJ, Shalapyonok A, Sosik HM. 2014. Diel size distributions reveal seasonal growth dynamics of a coastal phytoplankter. Proc Natl Acad Sci U S A 111:9852–9857. https://doi.org/10.1073/pnas.1321421111.
3. Carpenter EG, Chang J. 1988. Species-specific phytoplankton growth rates via diel DNA synthesis cycles. I. Concept of the method. Mar Ecol Prog Ser 43:105–111. https://doi.org/10.3354/meps043105.
4. Liu H, Campbell L, Landry MR, Nolla HA, Brown SA, Constantinou J. 1998. *Prochlorococcus* and *Synechococcus* growth rates and contributions to production in the Arabian Sea during the 1995 Southwest and Northeast Monsoons. Deep-Sea Res 45:2327–2352. https://doi.org/10.1016/S0967-0645(98)00073-3.
5. Sosik HM, Olson RJ, Neubert MG, Shalapyonok A, Solow AR. 2003. Growth rates of coastal phytoplankton from time-series measurements with a submersible flow cytometer. Limnol Oceanogr 48:1756–1765. https://doi.org/10.4319/lo.2003.48.5.1756.
6. Landry MR, Hassett RP. 1982. Estimating the grazing impact of marine micro-zooplankton. Mar Biol 67:283–288. https://doi.org/10.1007/BF00397668.
7. Redalje DG, Laws EA. 1981. A new method for estimating phytoplankton growth-rates and carbon biomass. Mar Biol 62:73–79. https://doi.org/10.1007/BF00396953.
8. Liu H, Nolla HA, Campbell L. 1997. Prochlorococcus growth rate and contribution to primary production in the equatorial and subtropical North Pacific Ocean. Aquat Microb Ecol 12:39–47. https://doi.org/10.3354/ame012039.
9. Korem T, Zeevi D, Suez J, Weinberger A, Avnit-Sagi T, Pompan-Lotan M, Matot E, Jona G, Harmelin A, Cohen N, Sirota-Madi A, Thaiss CA, Pevsner-Fischer M, Sorek R, Xavier RJ, Elinav E, Segal E. 2015. Growth dynamics of gut microbiota in health and disease inferred from single metagenomic samples. Science 349:1101–1106. https://doi.org/10.1126/science.aac4812.
10. Brown CT, Olm MR, Thomas BC, Banfield JF. 2016. Measurement of bacterial replication rates in microbial communities. Nat Biotechnol 34:1256–1263. https://doi.org/10.1038/nbt.3704.
11. Emiola A, Oh J. 2018. High throughput in situ metagenomic measurement of bacterial replication at ultra-low sequencing coverage. Nat Commun 9:4956–4963. https://doi.org/10.1038/s41467-018-07240-8.
12. Pittera J, Humily F, Thorel M, Grulois D, Garczarek L, Six C. 2014. Connecting thermal physiology and latitudinal niche partitioning in marine *Synechococcus*. ISME J 8:1221–1236. https://doi.org/10.1038/ismej.2013.228.
13. Mackey KRM, Post AF, McIlvin MR, Saito MA. 2017. Physiological and proteomic characterization of light adaptations in marine *Synechococcus*. Environ Microbiol 19:2348–2365. https://doi.org/10.1111/1462-2920.13744.
14. Jacquet S, Partensky F, Lennon JF, Vaulot D. 2001. Diel patterns of growth and division in marine picoplankton in culture. J Phycol 37:357–369. https://doi.org/10.1046/j.1529-8817.2001.037003357.x.
15. Waterbury JB, Watson SW. 1986. Biological and ecological characterization of the marine unicellular cyanobacterium Synechococcus. Can J Fish Aquat Sci 2014:71–120.
16. Helmstetter CE, Cooper S. 1968. DNA synthesis during division cycle of rapidly growing *Escherichia coli* Br. J Mol Biol 31:507–518. https://doi.org/10.1016/0022-2836(68)90424-5.
17. Watanabe S. 2020. Cyanobacterial multi-copy chromosomes and their replication. Biosci Biotechnol Biochem 84:1309–1321. https://doi.org/10.1080/09168451.2020.1736983.
18. Griese M, Lange C, Soppa J. 2011. Ploidy in cyanobacteria. FEMS Microbiol Lett 323:124–131. https://doi.org/10.1111/j.1574-6968.2011.02368.x.
19. Binder BJ, Chisholm SW. 1995. Cell-cycle regulation in marine Synechococcus sp. strains. Appl Environ Microbiol 61:708–717. https://doi.org/10.1128/aem.61.2.708-717.1995.
20. Perez-Sepulveda B, Pitt F, N'Guyen AN, Ratin M, Garczarek L, Millard A, Scanlan DJ. 2018. Relative stability of ploidy in a marine *Synechococcus* across various growth conditions. Environ Microbiol Rep 10:428–432. https://doi.org/10.1111/1758-2229.12614.
21. Dvorak P. 2017. Genome-wide analysis of cyanobacterial evolution: the example of *Synechococcus*, p. 35–54. *In* Los DA (ed), Cyanobacteria: omics and manipulation. Caister Academic Press, Poole, UK.
22. Jacquet S, Partensky F, Marie D, Casotti R, Vaulot D. 2001. Cell cycle regulation by light in *Prochlorococcus* strains. Appl Environ Microbiol 67:782–790. https://doi.org/10.1128/AEM.67.2.782-790.2001.
23. Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, DeLong EF. 2013. Pattern and synchrony of gene expression among sympatric marine microbial populations. Proc Natl Acad Sci U S A 110:E488–E497. https://doi.org/10.1073/pnas.1222099110.
24. Aylward FO, Boeuf D, Mende DR, Wood-Charlson EM, Vislova A, Eppley JM, Romano AE, DeLong EF. 2017. Diel cycling and long-term persistence of viruses in the ocean's euphotic zone. Proc Natl Acad Sci U S A 114:11446–11451. https://doi.org/10.1073/pnas.1714821114.

25. Long ANM, Hou SW, Ignacio-Espinoza JC, Fuhrman JA. 2021. Benchmarking microbial growth rate predictions from metagenomes. ISME J 15:183–195. https://doi.org/10.1038/s41396-020-00773-1.
26. Schulze-Makuch D, Lipus D, Arens FL, Baqué M, Bornemann TLV, de Vera J-P, Flury M, Frösler J, Heinz J, Hwang Y, Kounaves SP, Mangelsdorf K, Meckenstock RU, Pannekens M, Probst AJ, Sáenz JS, Schirmack J, Schloter M, Schmitt-Kopplin P, Schneider B, Uhl J, Vestergaard G, Valenzuela B, Zamorano P, Wagner D. 2021. Microbial hotspots in lithic microhabitats inferred from DNA fractionation and metagenomics in the Atacama Desert. Microorganisms 9:1038. https://doi.org/10.3390/microorganisms9051038.
27. Zecchin S, Mueller RC, Seifert J, Stingl U, Anantharaman K, von Bergen M, Cavalca L, Pester M. 2018. Rice paddy *Nitrospirae* carry and express genes related to sulfate respiration: proposal of the new genus "*Candidatus* Sulfobium." Appl Environ Microbiol 84:e02224-17. https://doi.org/10.1128/AEM.02224-17.
28. Zhao R, Mogollon JM, Abby SS, Schleper C, Biddle JF, Roerdink DL, Thorseth IH, Jorgensen SL. 2020. Geochemical transition zone powering microbial growth in subsurface sediments. Proc Natl Acad Sci U S A 117:32617–32626.
29. Hou JL, Sievert SM, Wang YZ, Seewald JS, Natarajan VP, Wang FP, Xiao X. 2020. Microbial succession during the transition from active to inactive stages of deep-sea hydrothermal vent sulfide chimneys. Microbiome 8:102. https://doi.org/10.1186/s40168-020-00851-8.
30. Hu P, Dubinsky EA, Probst AJ, Wang J, Sieber CMK, Tom LM, Gardinali PR, Banfield JF, Atlas RM, Andersen GL. 2017. Simulation of Deepwater Horizon oil plume reveals substrate specialization within a complex community of hydrocarbon degraders. Proc Natl Acad Sci U S A 114:7432–7437. https://doi.org/10.1073/pnas.1703424114.
31. Mori JF, Chen LX, Jessen GL, Rudderham SB, McBeth JM, Lindsay MBJ, Slater GF, Banfield JF, Warren LA. 2019. Putative mixotrophic nitrifying-denitrifying gammaproteobacteria implicated in nitrogen cycling within the ammonia/oxygen transition zone of an oil sands pit lake. Front Microbiol 10:2435. https://doi.org/10.3389/fmicb.2019.02435.
32. Waterbury JB. 1992. JB. 1992. The Cyanobacteria—isolation, purification, and identification, p 2058–2078. *In* Balows A, Trüper HG, Dworkin M, Harder W, Schleifer K-H (ed), The Prokaryotes, 2nd ed, vol 3. Springer, New York, NY.
33. Marie D, Simon N, Guillou L, Partensky F, Vaulot D. 2000. DNA/RNA analysis of phytoplankton by flow cytometry. Curr Protoc Cytom 11:11.12.1–11.12.14. https://doi.org/10.1002/0471142956.cy1112s11.
34. William S, Feil H, Copeland A. 2012. Bacterial genomic DNA isolation using CTAB. DOE Joint Genome Institute, Washington, DC. https://jgi.doe.gov/wp-content/uploads/2014/02/JGI-Bacterial-DNA-isolation-CTAB-Protocol-2012.pdf.
35. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359. https://doi.org/10.1038/nmeth.1923.