LETTERS

Photosynthesis genes in marine viruses yield proteins during host infection

Debbie Lindell[1], Jacob D. Jaffe[2]†, Zackary I. Johnson[1]†, George M. Church[2] & Sallie W. Chisholm[1,3]

Cyanobacteria, and the viruses (phages) that infect them, are significant contributors to the oceanic 'gene pool'[1,2]. This pool is dynamic, and the transfer of genetic material between hosts and their phages[3–6] probably influences the genetic and functional diversity of both. For example, photosynthesis genes of cyanobacterial origin have been found in phages that infect *Prochlorococcus*[5,7] and *Synechococcus*[8,9], the numerically dominant phototrophs in ocean ecosystems. These genes include *psbA*, which encodes the photosystem II core reaction centre protein D1, and high-light-inducible (*hli*) genes. Here we show that phage *psbA* and *hli* genes are expressed during infection of *Prochlorococcus* and are co-transcribed with essential phage capsid genes, and that the amount of phage D1 protein increases steadily over the infective period. We also show that the expression of host photosynthesis genes declines over the course of infection and that replication of the phage genome is a function of photosynthesis. We thus propose that the phage genes are functional in photosynthesis and that they may be increasing phage fitness by supplementing the host production of these proteins.

Photosynthesis in cyanobacteria, algae and plants requires two photosystems (denoted PSI and PSII). The D1 and D2 proteins (encoded by *psbA* and *psbD*, respectively) form a heterodimer in the reaction centre of PSII and bind the components required for photochemistry. The D1 protein is turned over rapidly owing to light-induced damage[10]; thus, its *de novo* synthesis is required for sustained photosynthesis[10]. High-light-inducible proteins (HLIPs) protect the photosynthetic apparatus from photodamage by dissipating excess light energy[11].

Numerous cyanophages contain photosynthesis genes (*psbA* and at least one other)[5,7–9] with highly conserved amino acid sequences[7,12], suggesting that they encode functional proteins that may be involved in maintaining host photosynthesis during infection. Here we used *Prochlorococcus* MED4 and the podovirus P-SSP7 (a T7-like phage[5]) as a model system to begin exploring this hypothesis. We considered that if these genes are involved in host photosynthesis, then the amount of phage production might be dependent on photosynthetic performance and, conversely, host photosynthesis might be compromised by phage infection. We examined the first part of this hypothesis by inhibiting photosynthesis with darkness or DCMU (3-(3,4-dichlorophenyl)-1,1-dimethylurea), an inhibitor of electron flow from PSII to PSI, which led to a respective four- or twofold reduction in replication of the phage genome (Fig. 1a). Thus, as in other systems[13–15], continued photosynthesis is necessary for maximal phage replication. To determine the converse, that is, whether host photosynthesis is influenced by phage infection, we measured PSII photochemical conversion efficiency (F_v/F_m) and functional cross-sectional area (σ_{PSII}) during the 8-h latent period before lysis. The former decreased only slightly, whereas the latter was constant throughout this period (Fig. 1b, c), indicating that phage infection does not lead to a marked decline in PSII performance, as occurs in some[14–16], but not other[17–19] photosynthetic host–virus systems.

Thus, continued photosynthesis is required for maximum phage production in our system. Moreover, photosynthesis is sustained during infection when a decline in the transcription and translation of host genes might be expected, suggesting that the expression of phage photosynthesis genes might be supplementing host metabolism. To address this hypothesis, we determined whether these phage genes are expressed and, if so, how their expression relates to that of the homologous host genes. Using probes specific for the phage and host *psbA* and *hli* genes, we found that both of the phage genes were transcribed (Fig. 2a, b and Supplementary Fig. 1). Using polymerase chain reaction with reverse transcription (RT–PCR), we determined

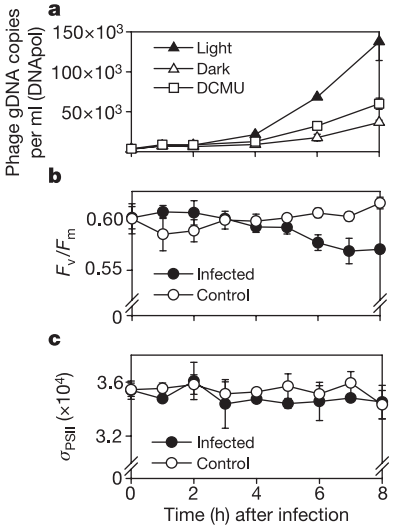

Figure 1 | Photosynthesis and phage infection. a, Replication of the phage genome was assessed by quantifying the phage gene encoding DNA polymerase in host cells kept in the light, transferred to the dark or treated with DCMU while in the light. Significantly lower ($P < 0.05$) phage DNA was detected in dark- and DCMU-treated cells after 4 h. **b**, **c**, PSII photochemical conversion efficiency (F_v/F_m; **b**) and PSII functional absorption cross-sectional area (σ_{PSII}; **c**) in infected and control cells. The absorption cross-section remained constant, although there was a 10% decline in PSII conversion efficiency. Error bars indicate the s.d. from biological replicates.

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. [2]Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA. [3]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. †Present addresses: The Broad Institute of Harvard and MIT, Cambridge, Massachusetts 02141, USA (J.D.J); Department of Oceanography, University of Hawaii, 1000 Pope Road, Honolulu, Hawaii 96822, USA (Z.I.J.).

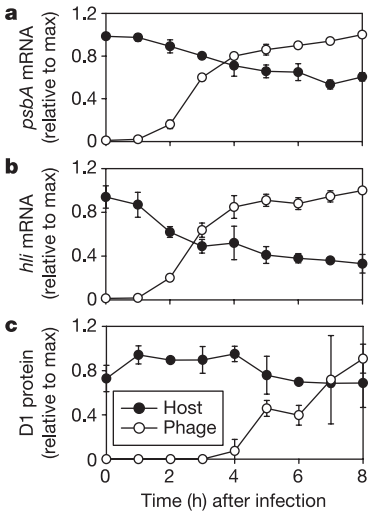

Figure 2 | Expression of phage and host photosynthesis genes. Temporal patterns of *psbA* mRNA (**a**), *hli12* mRNA (**b**) and D1 peptides (**c**). Expression during infection was normalized to expression in uninfected control cells and is presented relative to the maximal level for each individual gene. Data in **a** and **b** were obtained from microarray analyses. See Supplementary Fig. 1 for confirmation of the results shown in **a** by RT–PCR. Host D1 peptides were significantly lower in the second half of the latent period (5–8 h) than in the first half (0–4 h); $P < 0.01$. Error bars indicate the s.d. from biological replicates.

that phage *psbA* mRNA made up 50% of the total (phage plus host) *psbA* transcripts by 7–8 h after infection. By 4–5 h after infection, mRNA from the single host *psbA* gene had dropped to 50–60% of maximal levels (Fig. 2a and Supplementary Fig. 1). Transcription of 16 out of 22 of the host *hli* genes also declined significantly during infection (*hli12* is shown as a representative of this gene family in Fig. 2b). Because the maximal level of first-round infection achieved so far with this host–phage system is 50% (D.L. and S.W.C., unpublished data), the presence of only 50–60% of host *psbA* and *hli* transcripts relative to the control suggests that transcription of these genes had almost ceased by 4 h after infection and that nearly all *psbA* transcripts in infected cells were transcribed from the phage genome. The decline in host gene expression was not specific to photosynthesis genes but was part of a general reduction in host transcription subsequent to infection (D.L. and S.W.C., unpublished data).

The decline in host *psbA* transcription should cause a reduction in translation of the D1 protein and, because this protein is turned over rapidly[10], a decrease in host D1 titre. To verify this, we identified and quantified peptides specific for the host and phage D1 proteins (Fig. 3). The host D1 protein declined during infection to roughly 75% of maximal levels (Fig. 2c), reflecting a decrease of 50% in infected cells. Accompanying this reduction was a steady increase in the homologous protein encoded by the phage (Fig. 2c), which made up about 10% of total D1 in infected cells by the end of the latent period.

Although these results are consistent with our hypothesis that the expression of phage D1 protein helps to bolster host photosynthesis, the amount of phage D1 did not quantitatively compensate for the loss in host D1, even though PSII efficiency in the hosts declined very little during infection (Fig. 1b). This suggests that additional factors may be involved in maintaining host photosynthesis. For example, phage HLIPs, detected from 4 h after infection (data not shown), may reduce photosystem damage as well as being involved in the re-assembly of PSII[20]. Furthermore, phage D1 may be more efficient than host D1 during infection[6]. If photosynthetic antennae are shared among reaction centres, which is consistent with other studies[21], then fewer, more efficient functional reaction centres could lead to an increase in the functional absorption cross-section; however, this was not observed (Fig. 1c). Thus, although we are left with an imperfect balance sheet for the host and phage D1 protein, the inverse temporal expression of host and phage photosynthesis genes is striking and suggests that there is a functional interdependency.

A fitness advantage conferred by phage photosynthesis genes not

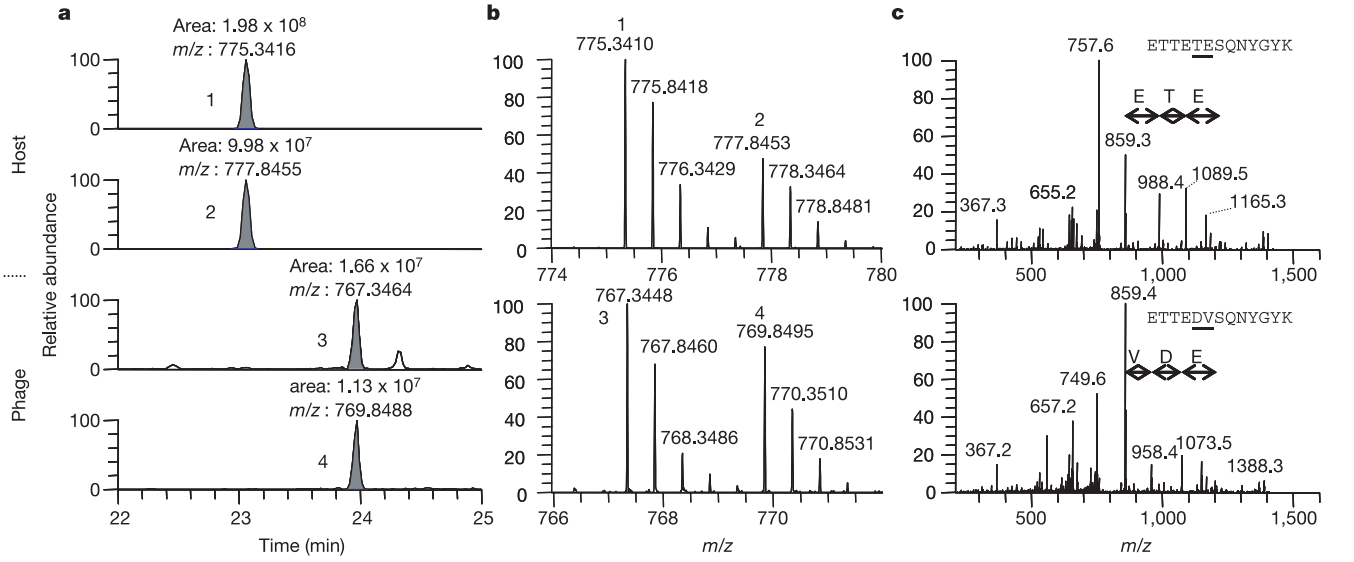

Figure 3 | Analysis of host and phage peptides. **a**, Extracted ion chromatograms of endogenous host (1) and isotopically labelled synthetic host (2) peptides, and endogenous phage (3) and isotopically labelled synthetic phage (4) peptides. Endogenous and corresponding synthetic peptides co-elute, whereas host and phage peptides have different retention times and *m*/*z* values. The area under the peaks of the endogenous peptides was used for quantification. **b**, Mass spectra taken at the apex of the chromatographic peaks show that host and phage peptides have different *m*/*z* values, as do the endogenous and isotopically labelled synthetic peptide pairs. Note that the threonine residue in the third position of the synthetic host and phage peptides was uniformly labelled with ^{15}N and ^{13}C isotopes, adding 5 Da to the mass of the peptides. **c**, Collision-induced dissociation spectra facilitate sequence identification of the host and phage peptides. The amino acid sequence in the partial peptide annotation is reversed because the *y*-ion series is shown. The full sequence of each peptide is shown with each spectrum.

only would explain their presence in cyanophage genomes[7,9], but also supports the modular theory of phage evolution[22], in which phages evolve through the step-wise acquisition of genes from diverse sources. According to this theory, acquired genes are initially expressed autonomously and are integrated into the phage life cycle if they provide a fitness advantage. The photosynthesis genes in cyanophage originate from cyanobacteria[7,9,12], and these phage genomes also contain bacterial, and even archaeal and eukaryotic genes[5,6]. Notably, the *psbA* and *psbD* genes in two *Prochlorococcus* myoviruses[7] have putative promoter and transcriptional terminators flanking the genes, suggesting that they are autonomously expressed. By contrast, photosynthesis genes in the phage used here have overlapping start and stop codons[7] and are co-transcribed with the essential, highly expressed phage capsid genes surrounding the photosynthesis genes (Fig. 4), suggesting that they have become an integral part of the phage genome. Thus, the proposed fitness advantage conferred by these photosynthesis genes, and other genes acquired from their hosts[5,6], may be a forerunner to their becoming *bona fide* members of the cyanophage gene pool.

Although we have not proved that phage photosynthesis gene products are participating in host photosynthesis, we favour this hypothesis, first, because photosynthesis continues during infection despite the decline in expression of host photosynthesis genes; second, because maximal phage DNA replication is dependent on this sustained photosynthesis; and last, because the high conservation of amino acid sequences[7,12] in the phage proteins suggests that these proteins are functioning in the same role as the host proteins. Nonetheless, we cannot rule out alternative functions. One could argue that D1, an efficient manganese-binding protein[10], is obtaining this metal ion for phage enzymatic reactions. Or perhaps phage D1, like the HLIP, is involved in dissipating excess light energy without being actively involved in photosynthesis. Obviously, the next stages in testing the 'same function hypothesis' are to see whether phage D1 localizes to the host PSII complex and to study the behaviour of this phage–host system using phage in which the photosynthesis genes have been inactivated.

The dynamic nature of the oceanic gene pool has led to the genetic diversification of donor and recipient genomes[7] and has apparently guided their functional diversification as well. If phage photosynthesis proteins do indeed function in host photosynthesis, this is a striking example of the interaction of proteins encoded from two distinct genomes in a single metabolic complex. Furthermore, the abundance of cyanobacteria[23] and their phages[16,24] in the oceans suggests that phage photosynthesis proteins have a small but significant role in the conversion of light to chemical energy on a global scale.

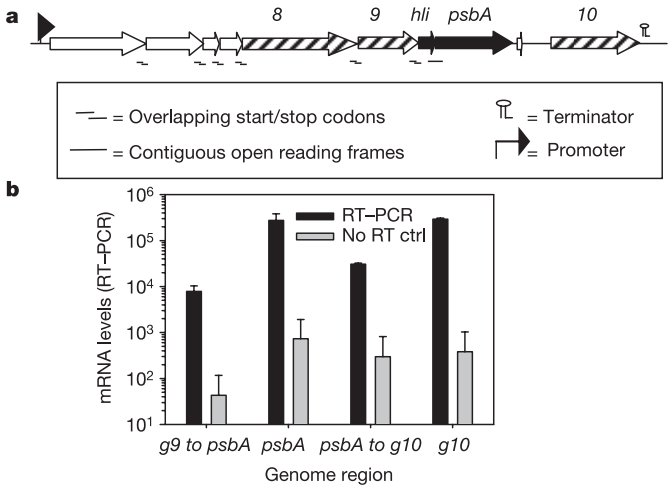

Figure 4 | Co-transcription of phage photosynthesis and capsid genes. **a**, Proposed operon of the photosynthesis gene region determined bioinformatically[7]. **b**, Transcript levels of regions internal to the *psbA* and *g10* genes, and regions spanning from *psbA* to the capsid genes upstream (*g9* to *psbA*) and downstream (*psbA* to *g10*) of *psbA* at 4 h after infection. *g9* encodes capsid assembly protein, *g10* encodes major capsid protein. Amplification from mRNA ('RT–PCR') was more than two orders of magnitude greater than from genomic DNA ('no RT control'). The lower transcript abundance across gene boundaries may be due to posttranscriptional processing or additional autonomous transcription of *psbA* and *g10*. Error bars indicate the s.d. from biological replicates.

METHODS

Experimental conditions. *Prochlorococcus* MED4 was grown at 21 °C under continuous cool white light (25 μmol photon $m^{-2} s^{-1}$) in Sargasso seawater Pro99 medium[25] amended with 10 mM HEPES (pH 7.5) and 12 mM sodium bicarbonate. Cells were concentrated to 10^8 cells per ml by centrifugation, and triplicate cultures were infected with 3×10^8 infective phages per ml for a multiplicity of infection (MOI) of 3 (except for the experiment shown in Fig. 1a; see below). Controls were amended with filter-sterilized spent medium. Cells for RNA and protein analyses were collected by centrifugation (12,400 *g* for 15 min at 20 °C), resuspended in buffer (200 mM sucrose, 10 mM sodium acetate, 5 mM EDTA; pH 5.2), snap frozen in liquid nitrogen and stored at −80 °C. Sample handling took 30 min. For the experiment shown in Fig. 1a, *Prochlorococcus* at 10^8 cells per ml were exposed to 10^7 infective phages per ml for an MOI of 0.1. After 1 h in the light to allow adsorption, cultures were diluted 4,000-fold and incubated in the light or dark, or in the light with 50 μM DCMU. We determined the MOI of phage stocks by the most probable number assay.

Quantitative detection of phage genomic DNA. *Prochlorococcus* cells were collected on 0.2-μm pore-sized polycarbonate filters (Osmonics), washed with sterile seawater followed by 3 ml of preservation solution (10 mM Tris, 100 mM EDTA, 0.5 M NaCl; pH 8) and frozen at −80 °C. We prepared DNA by a heat lysis method[26]. Phage genomic DNA was quantified by real-time PCR (see below), targeting the phage DNA polymerase gene. Primer sequences are given in Supplementary Table 1.

Photosynthesis measurements. A background irradiance gradient single-turnover fluorometer (BIG-STf) was used to measure the photosynthetic conversion efficiency (F_v/F_m) and functional absorption cross-section area (σ_{PSII}) of PSII, which measures the ability of PSII to absorb photons from antennae complexes[27]. Triplicate samples were dark acclimated for 15–30 min before single-turnover fluorescence induction curve measurements. F_v/F_m and σ_{PSII} were estimated by fitting standard models[28] to the data to determine F_o (initial fluorescence), F_m (maximal fluorescence), F_v ($F_m - F_o$) and σ_{PSII}.

RNA extraction and transcript analysis. Total RNA was extracted by a mirVana RNA isolation kit (Ambion). We removed DNA by a Turbo DNA-free kit (Ambion). For microarray analysis, RNA was concentrated by ethanol precipitation, and 2 μg of total RNA was labelled and hybridized to custom-made MD4-9313 arrays (Affymetrix) using the standard Affymetrix protocol for *Escherichia coli* (http://www.affymetrix.com/technology/index.affx). This array contains probe sets for genes and intergenic regions of both host and phage. Standard affymetrix procedures were used for probe design and construction of the array. For RT–PCR, total RNA (0.5–10 ng) was reverse transcribed with gene-specific primers and 100 U of SuperScript II (Invitrogen) in the presence of 200 U of SuperaseIN (Ambion). Triplicate real-time PCR reactions were done with a QuantiTect SYBR Green PCR kit (Qiagen) and primers at 0.3–1.0 μM. After 15 min at 95 °C, 40 cycles of denaturation (95 °C, 15 s), annealing (56 °C, 30 s) and elongation (72 °C, 30 s) were run on an DNA Engine Opticon (MJ Research), which was followed by 5 min at 72 °C and melt curve analysis. Incorporation of SYBR stain into double-stranded DNA was determined subsequent to elongation steps. Standard curves were generated with genomic DNA from *Prochlorococcus* or P-SSP7 phage particles. Primer sequences are given in Supplementary Table 1.

Protein preparation and analysis. Cells were lysed in 3 M urea, 0.05% SDS and 50 mM Tris-HCl (pH 8). Proteins were digested with sequencing grade trypsin (Promega) at a protein to trypsin ratio of 137.5:1, reduced with 10 mM dithiothreitol, alkylated with 50 mM iodoacetamide and acidified to pH < 3. Total protein was purified by solid-phase extraction using Oasis MCX (Waters), concentrated by vacuum centrifugation, resuspended in 0.1% formic acid and purified by solid-phase extraction using Oasis HLB (Waters). Peptides eluted with 70% acetonitrile were concentrated by vacuum centrifugation to dryness and resuspended in 5% acetonitrile and 5% formic acid. Tryptic digestions yielded peptides suitable for differentiation between the host and phage isoforms of D1. Peptides with the sequences NH_2-ETTETESQNYGYK-COOH and NH_2-ETTEDVSQNYGYK-COOH were observed by liquid chromatography mass spectrometry (LC-MS; Fig. 3) and used as surrogates for the host and phage

D1 proteins, respectively. Synthesized stable isotope-containing variants of these peptides were used as mass spectrometric standards to verify the elution time and tandem MS fragmentation patterns of the host and phage peptides (Fig. 3). D1 peptides were identified and quantified by reversed-phase LC-MS as described[29] from duplicate injections of 5.7 μg of total peptides using a hybrid linear ion trap, Fourier transform ion cyclotron resonance mass spectrometer (ThermoElectron) with resolution set to 100,000 and a mass accuracy of ±8 p.p.m. Extracted ion chromatograms were generated by monitoring the mass/charge (m/z) values of 775.3363 ± 0.006 and 767.3388 ± 0.006 for the host and phage peptides, respectively. We used the area under the peaks from XCalibur Software for quantification (see also http://arep.med.harvard.edu/mapquant.html for an alternative method).

Received 27 June; accepted 27 July 2005.
Published online 12 October 2005.

1. Breitbart, M. *et al.* Genomic analysis of uncultured marine viral communities. *Proc. Natl Acad. Sci. USA* **99,** 14250–14255 (2002).
2. Venter, J. C. *et al.* Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304,** 66–74 (2004).
3. Canchaya, C., Fournous, G., Chibani-Chennoufi, S., Dillmann, M. L. & Brussow, H. Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.* **6,** 417–424 (2003).
4. Palenik, B. *et al.* The genome of a motile marine *Synechococcus*. *Nature* **424,** 1037–1042 (2003).
5. Sullivan, M. B., Coleman, M. L., Weigele, P., Rohwer, F. & Chisholm, S. W. Three *Prochlorococcus* cyanophage genomes: signature features and ecological interpretations. *PLoS Biology* **3,** e144 (2005).
6. Mann, N. H. *et al.* The genome of S-PM2, a 'photosynthetic' T4-type bacteriophage that infects marine *Synechococcus*. *J. Bacteriol.* **187,** 3188–3200 (2005).
7. Lindell, D. *et al.* Transfer of photosynthesis genes to and from *Prochlorococcus* viruses. *Proc. Natl Acad. Sci. USA* **101,** 11013–11018 (2004).
8. Mann, N. H., Cook, A., Millard, A., Bailey, S. & Clokie, M. Bacterial photosynthesis genes in a virus. *Nature* **424,** 741 (2003).
9. Millard, A., Clokie, M. R. J., Shub, D. A. & Mann, N. H. Genetic organization of the *psbAD* region in phages infecting marine *Synechococcus* strains. *Proc. Natl Acad. Sci. USA* **101,** 11007–11012 (2004).
10. Adir, N., Zer, H., Shochat, S. & Ohad, I. Photoinhibition—a historical perspective. *Photosynth. Res.* **76,** 343–370 (2003).
11. Havaux, M., Guedeney, G., He, Q. & Grossman, A. R. Elimination of high-light-inducible polypeptides related to eukaryotic chlorophyll *a/b*-binding proteins results in aberrant photoacclimation in *Synechocystis* PCC6803. *Biochim. Biophys. Acta* **1557,** 21–33 (2003).
12. Zeidner, G. *et al.* Potential photosynthesis gene recombination between *Prochlorococcus* and *Synechococcus* via viral intermediates. *Environ. Microbiol.* **7,** 1505–1513 (2005).
13. Adolph, K. W. & Haskelkorn, R. Photosynthesis and the development of blue–green algal virus N-1. *Virology* **47,** 370–374 (1972).
14. MacKenzie, J. J. & Haselkorn, R. Photosynthesis and the development of blue–green algal virus SM-1. *Virology* **49,** 517–521 (1972).
15. Sherman, L. A. Infection of *Synechococcus cedrorum* by the cyanophage AS-1M. *Virology* **71,** 199–206 (1976).
16. Suttle, C. A. & Chan, A. M. Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*—abundance, morphology, cross-infectivity and growth-characteristics. *Mar. Ecol. Prog. Ser.* **92,** 99–109 (1993).
17. Ginzburg, D., Padan, E. & Shilo, M. Effect of cyanophage infection on CO_2 photoassimilation in *Plectonema boryanum*. *J. Virol.* **2,** 695–701 (1968).
18. Rahoutei, J., Garcia-Luque, I. & Baron, M. Inhibitioin of photosynthesis by viral infection: effect on PSII structure and function. *Physiol. Plant.* **110,** 286–292 (2000).
19. Arias, M. C., Lenardon, S. & Taleisnik, E. Carbon metabolism alterations in sunflower plants infected with the sunflower chlorotic mottle virus. *J. Phytopath.* **151,** 267–273 (2003).
20. Xu, H., Vavilin, D., Funk, C. & Vermaas, W. Multiple deletions of small Cab-like proteins in the cyanobacterium *Synechocystis sp.* PCC 6803—consequences for pigment biosynthesis and accumulation. *J. Biol. Chem.* **279,** 27971–27979 (2004).
21. Bruyant, F. *et al.* Diel variations in the photosynthetic parameters of *Prochlorococcus* strain PCC 9511: combined effects of light and cell cycle. *Limnol. Oceanogr.* **50,** 850–863 (2005).
22. Hendrix, R. W., Lawrence, J. G., Hatfull, G. F. & Casjens, S. The origins and ongoing evolution of viruses. *Trends Microbiol.* **8,** 504–508 (2000).
23. Partensky, F., Hess, W. R. & Vaulot, D. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol. Mol. Biol. Rev.* **63,** 106–127 (1999).
24. Sullivan, M. B., Waterbury, J. B. & Chisholm, S. W. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. *Nature* **424,** 1047–1051 (2003).
25. Moore, L. R., Post, A. F., Rocap, G. & Chisholm, S. W. Utilization of different nitrogen sources by the marine cyanobacteria *Prochlorococcus* and *Synechococcus*. *Limnol. Oceanogr.* **47,** 989–996 (2002).
26. Zinser, E. R. *et al. Prochlorococcus* ecotype abundance in the North Atlantic Ocean revealed by an improved quantitative PCR method. *Appl. Environ. Microbiol.* (in the press).
27. Johnson, Z. I. Development and application of the background irradiance gradient–single turnover fluorometer (BIG-STf). *Mar. Ecol. Prog. Ser.* **283,** 73–80 (2004).
28. Kolber, Z. S., Prasil, O. & Falkowski, P. G. Measurements of variable chlorophyll fluorescence using fast repetition rate techniques—defining methodology and experimental protocols. *Biochim. Biophys. Acta* **1367,** 88–106 (1998).
29. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W. & Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl Acad. Sci. USA* **100,** 6940–6945 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank T. Rector and R. Steen for doing the Affymetrix GeneChip experiments; C. Steglich, M. Sullivan, M. Coleman, and E. Zinser for discussions; and M. Sullivan for comments on the manuscript. This research was supported by grants from the National Science Foundation (to S.W.C.), the Gordon and Betty Moore Foundation's Program in Marine Microbiology (to S.W.C), and the Department of Energy Genomes to Life Program (to S.W.C and G.M.C.).

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to S.W.C. (chisholm@mit.edu).