

# Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocospaera*

Samuel T. Wilson<sup>1†</sup>, Frank O. Aylward<sup>1†‡</sup>, Francois Ribalet<sup>2</sup>, Benedetto Barone<sup>1</sup>, John R. Casey<sup>1</sup>, Paige E. Connell<sup>3</sup>, John M. Eppley<sup>1</sup>, Sara Ferrón<sup>1</sup>, Jessica N. Fitzsimmons<sup>4</sup>, Christopher T. Hayes<sup>5</sup>, Anna E. Romano<sup>1</sup>, Kendra A. Turk-Kubo<sup>6</sup>, Alice Vislova<sup>1</sup>, E. Virginia Armbrust<sup>2</sup>, David A. Caron<sup>3</sup>, Matthew J. Church<sup>1‡</sup>, Jonathan P. Zehr<sup>6</sup>, David M. Karl<sup>1</sup> and Edward F. DeLong<sup>1\*</sup>

**The temporal dynamics of phytoplankton growth and activity have large impacts on fluxes of matter and energy, yet obtaining *in situ* metabolic measurements of sufficient resolution for even dominant microorganisms remains a considerable challenge. We performed Lagrangian diel sampling with synoptic measurements of population abundances, dinitrogen (N<sub>2</sub>) fixation, mortality, productivity, export and transcription in a bloom of *Crocospaera* over eight days in the North Pacific Subtropical Gyre (NPSG). Quantitative transcriptomic analyses revealed clear diel oscillations in transcript abundances for 34% of *Crocospaera* genes identified, reflecting a systematic progression of gene expression in diverse metabolic pathways. Significant time-lagged correspondence was evident between *nifH* transcript abundance and maximal N<sub>2</sub> fixation, as well as *sepF* transcript abundance and cell division, demonstrating the utility of transcriptomics to predict the occurrence and timing of physiological and biogeochemical processes in natural populations. Indirect estimates of carbon fixation by *Crocospaera* were equivalent to 11% of net community production, suggesting that under bloom conditions this diazotroph has a considerable impact on the wider carbon cycle. Our cross-scale synthesis of molecular, population and community-wide data underscores the tightly coordinated *in situ* metabolism of the keystone N<sub>2</sub>-fixing cyanobacterium *Crocospaera*, as well as the broader ecosystem-wide implications of its activities.**

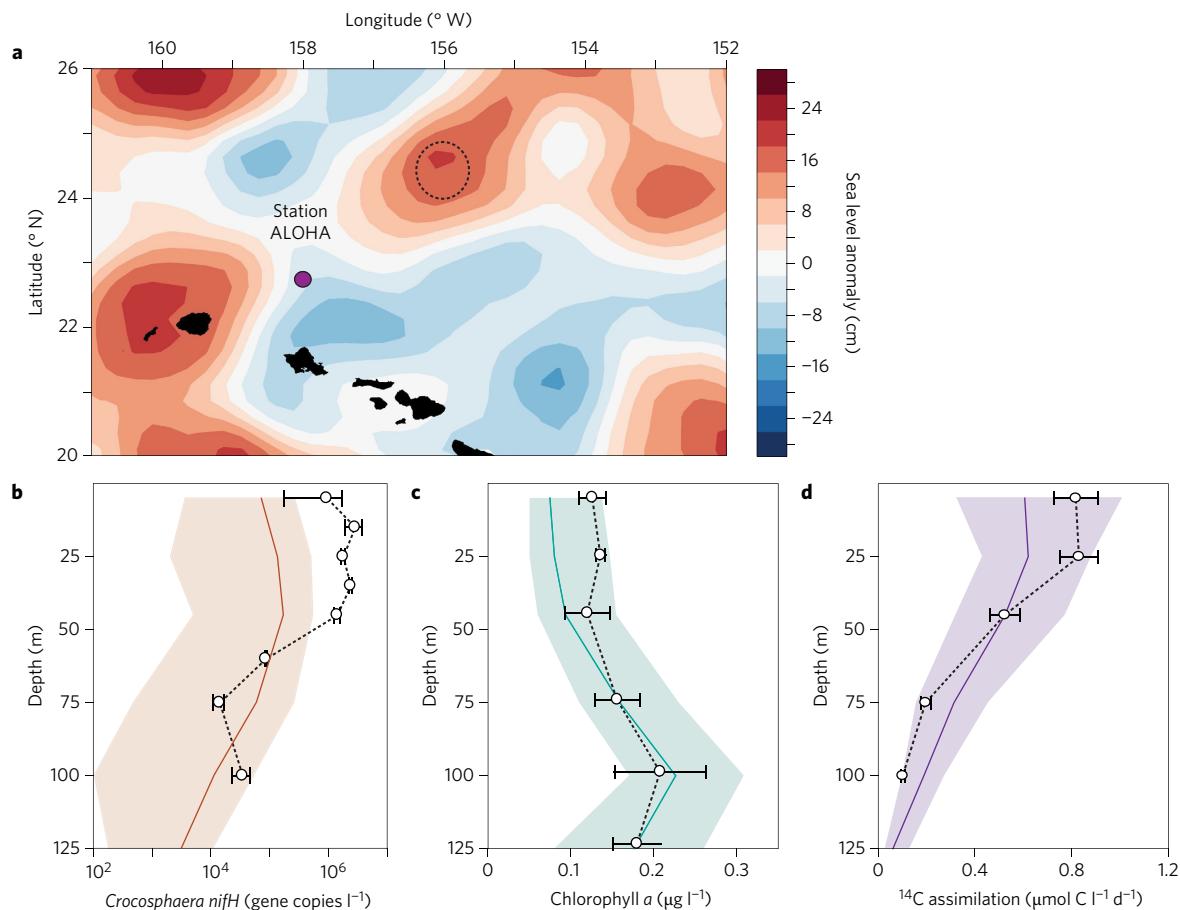
Ecophysiological adaptations that permit cyanobacteria to thrive in the oligotrophic gyres of the open ocean are of paramount interest due to the important role of microbial productivity and their impact on elemental cycling in these vast biomes. Insights into microbial adaptations in this habitat have primarily been obtained from the highly abundant cyanobacterium *Prochlorococcus* and include efficient light capture and dissipation mechanisms, high population microdiversity, small cell sizes and compact genomes<sup>1</sup>. However, alternative cyanobacterial life history strategies are also found in oligotrophic ocean settings. A prominent example is the unicellular cyanobacterium *Crocospaera watsonii*, which is considerably larger than *Prochlorococcus*<sup>2</sup>, with a larger genome<sup>3</sup>, lower genomic microdiversity among shared genes<sup>3,4</sup> and the capability for both oxygenic photosynthesis and N<sub>2</sub> fixation<sup>5</sup>. The biogeographical distribution of *Crocospaera* and its ecological importance in the oligotrophic ocean has become increasingly recognized<sup>6</sup>, although its lower abundance and patchy spatial distributions compared to *Prochlorococcus* make it considerably more difficult to study in the field<sup>7–9</sup>. Given these challenges, most of our knowledge about how this cyanobacterium balances its metabolic demands and maintains its ecological niche within an oligotrophic habitat originates from laboratory-based studies<sup>10–12</sup>.

In this study, we provide an in-depth analysis of natural field populations of *Crocospaera* to assess its metabolic activity and its broader ecological impact within the oligotrophic ecosystem. An intensive multidisciplinary field campaign made use of Lagrangian diel analyses, in which water-column sampling was conducted in close proximity to free-drifting drogues centred at a depth of 15 m. This Lagrangian sampling strategy mitigates the effect of spatial variability and allows microbial populations within the same water parcel to be tracked over time. The combined assessment of our molecular, physiological and biogeochemical measurements implicates *Crocospaera* as a strong driver of ecosystem dynamics, underscoring the importance of low-abundance keystone members to broader community functioning.

## Results

During the period of Lagrangian observations in the North Pacific Subtropical Gyre (NPSG), *Crocospaera* was the third most numerically abundant cyanobacterium and second with respect to biomass (after *Prochlorococcus*) (Supplementary Table 1) within the study area (Fig. 1 and Supplementary Table 1). The abundance of *Crocospaera nifH* genes was greater than other diazotrophs, ranging from  $9.4 \pm 0.7 \times 10^5$  to  $2.8 \pm 0.9 \times 10^6$  gene copies per litre

<sup>1</sup>Daniel K. Inouye Center for Microbial Oceanography: Research and Education, Department of Oceanography, University of Hawaii, Honolulu, Hawaii 96822, USA. <sup>2</sup>School of Oceanography, University of Washington, Seattle, Washington 98195, USA. <sup>3</sup>Department of Biological Sciences, University of Southern California, Los Angeles, California 90089, USA. <sup>4</sup>Department of Oceanography, Texas A&M University, College Station, Texas 77843, USA. <sup>5</sup>School of Ocean Science and Technology, University of Southern Mississippi, Stennis Space Center, Mississippi 39529, USA. <sup>6</sup>Ocean Sciences Department, University of California, Santa Cruz, California 95064, USA. <sup>†</sup>These authors contributed equally to this work. <sup>‡</sup>Present addresses: Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia 24061, USA (F.O.A.); Flathead Lake Biological Station, University of Montana, Polson, Montana 59860, USA (M.J.C.). \*e-mail: edelong@hawaii.edu

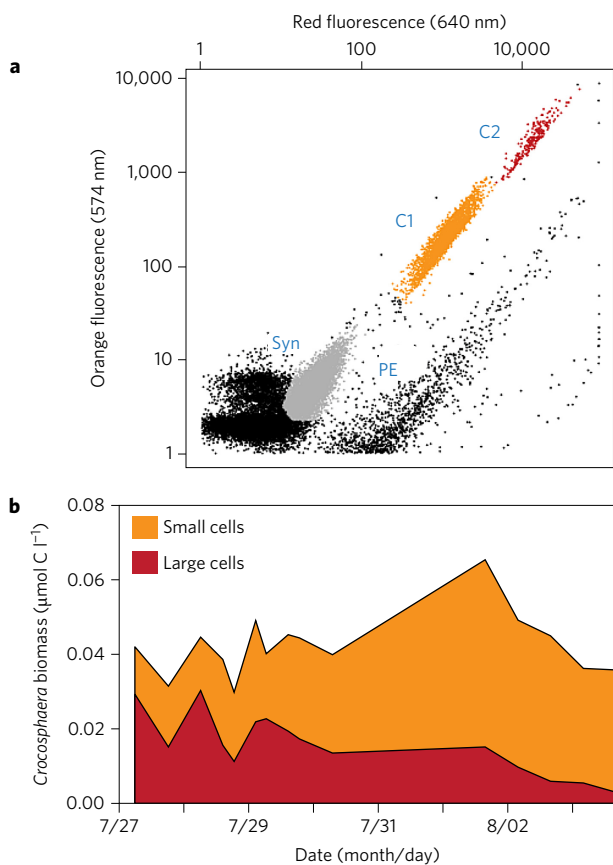


**Figure 1 | High *Crocosphaera* abundances located in an anticyclonic eddy coinciding with elevated water column chlorophyll *a* and productivity.**

**a**, Patterns of the sea level anomaly north of Hawaii, with the cruise location indicated by a dashed circle. **b–d**, Vertical profiles of *Crocosphaera nifH* gene abundance ( $n = 1$ ) (**b**), chlorophyll *a* ( $n = 5$ ) (**c**) and  $^{14}\text{C}$ -bicarbonate assimilation ( $n = 3$ ) (**d**). Values from July to August 2015 are shown as white symbols, and error bars represent standard deviation. Also plotted are the mean (solid line) and 5th and 95th percentile values (shaded area) for the months of July and August during 1988–2014 for chlorophyll *a* and  $^{14}\text{C}$  assimilation and during 2004–2014 for *nifH* gene abundance, measured at Station ALOHA.

(mean  $\pm$  standard deviation (s.d.)) in the upper 50 m of the water column (Fig. 1b and Supplementary Table 1). Similar *nifH* gene abundances for *Crocosphaera* have been observed previously during late summer in the NPSG when *Crocosphaera* reaches its maximum seasonal population abundances<sup>13</sup>, although it is uncommon for *Crocosphaera nifH* gene counts to exceed those of other diazotrophs at this time of year<sup>9,13</sup>. Two distinct populations of *Crocosphaera* were identified, with cell diameters either smaller or larger than 4  $\mu\text{m}$ , and are hereafter referred to as small- and large-sized cells, respectively (Fig. 2a). The mean cell diameter of the small cells was  $2.1 \pm 0.46 \mu\text{m}$  and that of the large cells was  $5.0 \pm 0.76 \mu\text{m}$  (Table 1). The smaller cells were always more abundant ( $100\text{--}600 \text{ cells ml}^{-1}$ ) than larger cells ( $\leq 35 \text{ cells ml}^{-1}$ ) during our observation period. Nonetheless, the large-sized cells contributed significantly ( $\sim 50\%$ ) to the overall *Crocosphaera* biomass during the initial period of measurements (Fig. 2b). There have been few simultaneous measurements of the two *Crocosphaera* populations in the field<sup>14</sup> despite the variability in cell size of the two subpopulations and other phenotypic differences, including the production of exopolysaccharide by the large cells<sup>15,16</sup>. Because the two subpopulations of *Crocosphaera* presented in this study are differentiated according to cell size (Fig. 2a), all the data collected from whole seawater analyses, including transcriptomics,  $\text{N}_2$  fixation, productivity and export, pertain to the total population, and only the microscopic and flow cytometric measurements distinguish the two subpopulations.

Quantitative transcriptomic analyses identified between  $1.5 \times 10^8$  and  $24.8 \times 10^8$  *Crocosphaera* transcripts per litre across all time points sampled during the cruise. Reads from the transcriptomes were mapped to the *Crocosphaera* genes identified in the metagenomes, of which 1,978 were sufficiently abundant for in-depth temporal analyses (Datasets 1–4; see Supplementary Methods for details). These 1,978 transcripts were analysed using a weighted gene correlation network approach<sup>17,18</sup>, which clustered transcripts into six sets, referred to as ‘modules’, based on their temporal abundance profiles (Fig. 3a). Visualization of this network throughout a 24 h period revealed a tight transcriptional progression during the diel cycle, with peak expression of genes involved in ribosome biosynthesis, photopigment synthesis/degradation, photosynthesis and cell division cascading from midnight to noon (Fig. 3b), and a distinct burst of transcription involving the nitrogenase complex at dusk (at approximately 18:00). Although diel periodicity was overwhelmingly the strongest signal in the transcriptional modules (Fig. 3c), day-to-day variability in transcript abundance was also observed, probably due to a combination of population-level dynamics as well as finer-scale changes in gene regulation during the sampling period. Independent of network analyses, the evaluation of peak transcript abundance in key metabolic pathways recapitulated the pattern of diel choreography (Fig. 4). For most pathways, peak transcript abundance occurred between midnight and noon, similar to previous observations of laboratory-maintained cultures<sup>19</sup>. Notably, 23 transposases (out of 105 total identified in the transcriptomes) exhibited



**Figure 2 | Sub-populations of *Crocosphaera*.** **a**, Flow cytogram differentiating the two *Crocosphaera* populations with small cells (C1) and large cells (C2), as well as *Synechococcus* (Syn) and picoeukaryotes (PE). **b**, *Crocosphaera* biomass for the two populations during the period of Lagrangian observations (technical replicate,  $n = 3$  for each time point) using a cellular carbon content of 1.2 and 10.1 pg C per cell for the small and large cells, respectively (Supplementary Section 'Enumeration of *Crocosphaera*').

diel transcriptional periodicity with peak expression between 5:00 and 10:00, slightly later than observed in cultures of *Crocosphaera* strain WH8501<sup>19</sup>. The peak expression of these genes at a time similar to cell division and DNA replication indicates that the activity of these processes may be synchronized.

The diel transcriptional patterns of *Crocosphaera* are similar to those of other photoautotrophs such as *Prochlorococcus*, *Ostreococcus* and *Synechococcus* with respect to the timing of expression of photosynthetic genes<sup>20,21</sup>, which initiate production of the energy harvesting apparatus in the early morning. However, it is the night-time transcriptional activity and subsequent  $N_2$  fixation by *Crocosphaera* that distinguishes this cyanobacterium from other photoautotrophs, as previously noted<sup>19</sup>. Cell division in *Crocosphaera* is delayed until post-dawn, compared to dusk in *Prochlorococcus*, and maximum expression of ribosomal genes occurs near midnight (Fig. 4), compared to about 4:00 in *Prochlorococcus* and *Ostreococcus*<sup>18,20,21</sup>. The overall transcriptional choreography of field populations of *Crocosphaera*, including the timing of cell division near dawn, the shift in ribosomal biosynthesis and the prevalent expression of *nif* genes at dusk, is similar to diel patterns previously observed in *Cyanothece*, another unicellular diazotroph<sup>22,23</sup>. Given that *Crocosphaera* and *Cyanothece* are both unicellular photosynthetic diazotrophs, this further supports the hypothesis that the accommodation of  $N_2$  fixation in unicellular cyanobacteria requires a tightly orchestrated diel cycle distinct from other diazotrophs and non-diazotrophic microorganisms.

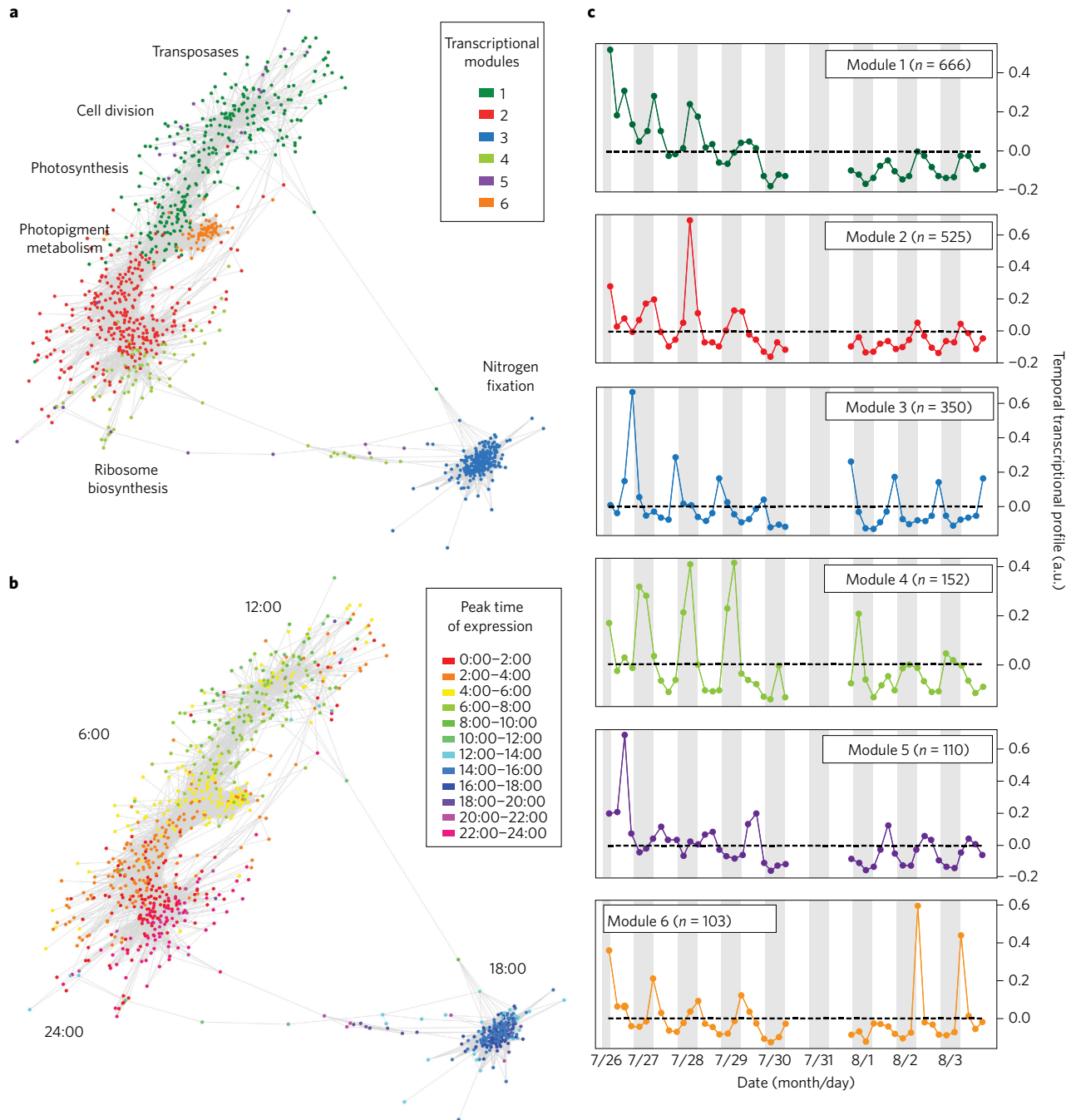
**Table 1 | *Crocosphaera* population dynamics in the context of water column productivity and export.**

Parameter	Values	
	Small cells	Large cells
<b>Stocks</b>		
<i>Crocosphaera</i> abundance ( $10^6$ cells $l^{-1}$ )	0.1–0.7	0.01–0.04
Average cell diameter ( $\mu m$ )	$2.1 \pm 0.46$	$5.0 \pm 0.76$
Carbon content (pg C per cell)	$1.2 \pm 0.6$	$10.1 \pm 3.9$
<b>Rates</b>		
<i>Crocosphaera</i> growth rate ( $d^{-1}$ )	$0.6 \pm 0.2$	
<i>Crocosphaera</i> grazing rate ( $d^{-1}$ )	NS to $-0.7$	
<i>Crocosphaera</i> -specific $N_2$ fixation ( $nmol l^{-1} d^{-1}$ )	$7.3 \pm 1.5$	
NCP ( $\mu mol C l^{-1} d^{-1}$ )	$0.45 \pm 0.03$	
C fixation ( $\mu mol C l^{-1} d^{-1}$ )	$0.82 \pm 0.08$	
$N_2$ fixation ( $nmol N l^{-1} d^{-1}$ )	$10.9 \pm 1.5$	
PC export flux at 150 m ( $\mu mol C m^{-2} d^{-1}$ )	$2353 \pm 217$	
PN export flux at 150 m ( $\mu mol N m^{-2} d^{-1}$ )	$298 \pm 32$	
$\delta^{15}N$ of PN exported at 150 m (‰)	$2.3 \pm 0.7$	

Values are reported either as the range or mean  $\pm$  s.d. Carbon content per cell was computed from biovolume<sup>23</sup> derived from the measured cell diameter using the geometry of a sphere (Supplementary Section 'Biomass and growth rate estimates'). Doubling times were derived from the continual underway measurements and the grazing rates were derived from dilution grazing experiments (Supplementary Section 'Growth and grazing experiments'). NCP was converted to units of C using a photosynthetic quotient of 1.1 (ref. 54) and divided by the depth of the mixed layer to obtain a volumetric rate. Rows 1–6 show stocks and rates applicable to *Crocosphaera*, and rows 7–12 show values applicable to the whole community. NS, not significant.

Transcriptomic measurements coincided with corresponding physiological and population abundance measurements for field populations of *Crocosphaera*. A pronounced diel cycle was observed in the abundance of the small cell *Crocosphaera* population (as measured by SeaFlow<sup>24</sup>) (Fig. 5a). Cell numbers increased sharply one hour after sunrise and increased by 30–40% between 7:00 and 11:00, corresponding to net cell division. Cell division-related transcriptional activity of the *Crocosphaera sepF* gene, which encodes a key cell division protein<sup>25</sup>, exhibited a sharp peak just before 7:00 (Fig. 5a). This rapid increase of *sepF* gene transcripts at dawn was predictive of subsequent increases in cell abundance, and served as a reliable predictor for the timing of cell division. The temporal delay in cell division by  $\sim 1$  h after sunrise could reflect the need to recover energy reserves with light-derived energy after night-time metabolic activity. Alternatively, it could also reflect the balance between cell growth and cell removal due to grazing and/or viral lysis. Cell abundances of small-celled *Crocosphaera* consistently increased from dawn to noon and decreased for the remainder of the day (Fig. 5a). Because the same water mass was being followed during the observational period, it is assumed that the effects of physical processes on cell abundances are negligible. Under those assumptions, the increase and decrease in cell abundances can be attributed to cell division and cell mortality, respectively. The diel changes in cell abundances observed during daylight indicate a tight coupling between cell growth and mortality processes, such as grazing and viral infection. A growth rate of  $0.6 \pm 0.2$  (s.d.) per day for *Crocosphaera* was calculated from the net increase in cell abundances between 6:00 and 11:00 (Supplementary Section 'Biomass and growth rate estimates'), which is similar to the growth rates of  $\sim 0.5$  per day reported for laboratory-maintained strains<sup>5,10,26</sup> and indicates that the field *Crocosphaera* population was dividing at near maximal theoretical rates<sup>27</sup>. We therefore infer that *Crocosphaera* was not limited by macro- or micronutrient availability during the period of Lagrangian observations, and surmise that mortality due to grazing and viral lysis were probably key controls on the abundance of these unicellular cyanobacteria.

$N_2$  fixation was also tightly regulated over the diel cycle, and the peak rates of this metabolic process occurred between 23:00 and 4:00 (Fig. 5b). The diel pattern of nitrogenase activity for this open ocean population of *Crocosphaera* was similar to that observed in

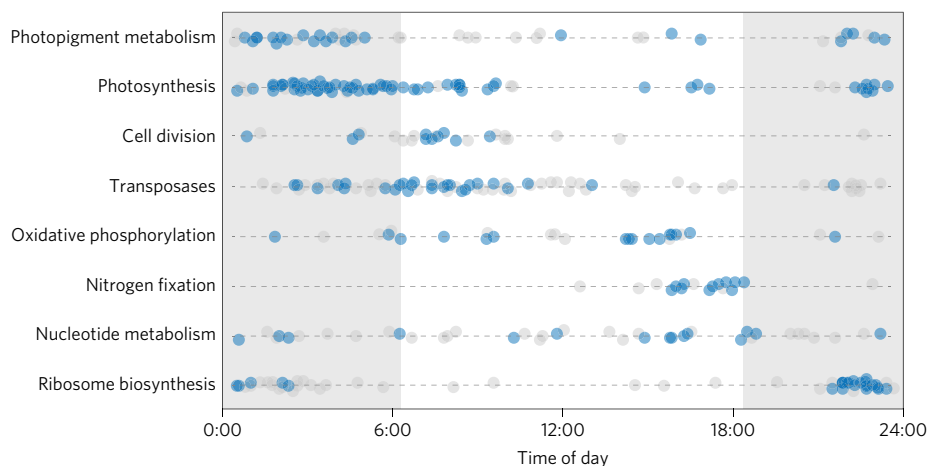


**Figure 3 | Transcriptional network of *Crocosphaera*.** **a**, Representation of a weighted transcriptional network coloured by module. Nodes in the network represent transcripts and grey lines represent edges that link transcripts with high pairwise topological overlap scores. Names of metabolic pathways are listed near the region of the network where genes associated with that process are present. Genes from a particular pathway are in some cases distributed among several modules (see Supplementary Dataset 2 for details). **b**, Depiction of the weighted transcriptional network shown in **a** with nodes coloured according to their peak expression time. **c**, Temporal profiles (module eigengenes, or first-principle components<sup>55</sup>) for the six transcriptional modules depicted in **a**, using the same colour scheme. The y axis denotes the relative abundance trendlines for the eigengenes (see Supplementary Information for details), dashed lines represent module averages and grey shading indicates night time. a.u., arbitrary units.

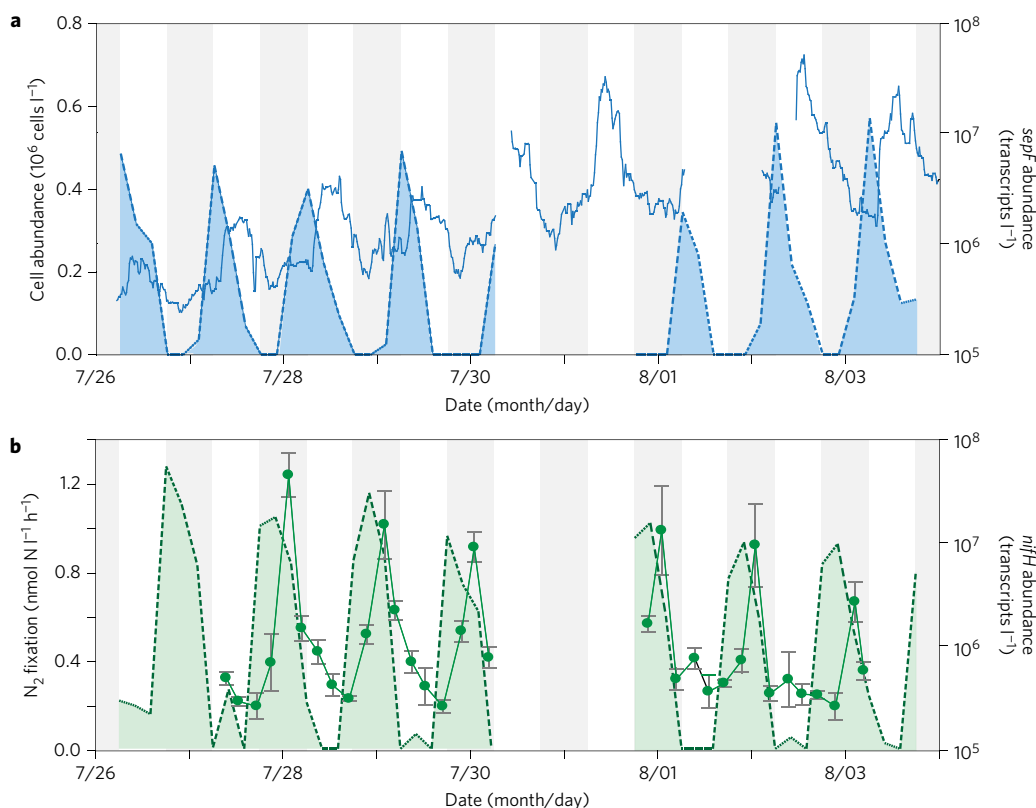
isolated strains in the laboratory<sup>11,12</sup> and at Station ALOHA<sup>28</sup>. The peak in *Crocosphaera*-specific *nifH* transcript abundance occurred just before the peak rate of N<sub>2</sub> fixation, providing further evidence that microorganism-specific transcript abundance yields insight into the physiology of natural populations. The rates of N<sub>2</sub> fixation during the night ( $7.3 \pm 1.5 \text{ nmol l}^{-1} \text{ d}^{-1}$ ) slightly exceed the estimated cellular nitrogen requirements for growth ( $5.1 \pm 3.4 \text{ nmol l}^{-1} \text{ d}^{-1}$ ) calculated from cell abundance, nitrogen quota per cell and doubling times (Supplementary Section ‘Growth requirements’).

The difference could reflect an underestimation of growth rates, which are based on cell abundances and therefore reflect the net balance between growth and mortality. An additional explanation is that *Crocosphaera* fixes N<sub>2</sub> in excess of its growth requirements, as is known to occur with the well-studied marine diazotroph *Trichodesmium*, which maintains an active nitrogenase during the entire daylight period and has been shown to leak fixed nitrogen from the cells<sup>29</sup>. Finally, it is also possible that N<sub>2</sub> was being fixed by other diazotrophs such as *Candidatus* Atelocyanobacterium





**Figure 4 | Diel cascade in *Crocosphaera* transcripts.** The peak expression time of 410 genes from key metabolic processes over the 8 day sampling period are aggregated over a 24 h period with the night period indicated by grey shading (Supplementary Dataset 2). Blue symbols represent the peak times for transcripts that have a statistically significant (RAIN<sup>53</sup>,  $P < 0.05$ ) 24 h periodicity; and grey symbols indicate peak times for transcripts with no significant diel component to their expression.



**Figure 5 | Diel patterns in *Crocosphaera* abundances and  $N_2$  fixation alongside transcript abundances for key genes.** **a**, *Crocosphaera* abundances for small cells (solid line), with single measurements conducted every 3 min. Transcript abundance of *sepF* is shown as a dashed line. **b**, Rates of  $N_2$  fixation (solid line and symbols with error bars representing standard deviation,  $n = 3$ ) and *nifH* transcript abundance (dashed line).

thalassa (UCYN-A) and heterotrophic bacteria, for which the diel periodicity in nitrogenase activity remains unclear. However, given the low *nifH* gene abundance of UCYN-A and heterotrophic bacteria compared to *Crocosphaera* (Supplementary Table 1), their contribution to the pool of fixed nitrogen will be significantly lower than *Crocosphaera*. The relatively compact period of  $N_2$  fixation in *Crocosphaera* facilitates preparation for cell replication and construction of the photosynthetic apparatus, which also occurs before dawn (Fig. 4). Laboratory studies have shown that increased

rates of respiration during the time period when *Crocosphaera* fixes  $N_2$  (ref. 30) not only provide energy to fuel this metabolic process, but also protect nitrogenase from deleterious interactions with  $O_2$  (ref. 31). Consistent with this pattern, in the current study, many transcripts involved in oxidative phosphorylation exhibited a diel peak in the late afternoon (Fig. 4). Overall, the estimated rates of  $N_2$  fixation during the night comprised two-thirds of the total community  $N_2$  fixation ( $10.9 \pm 1.5 \text{ mmol N l}^{-1} \text{ d}^{-1}$ ) (Table 1). These daily rates exceed the average values measured at nearby Station ALOHA during the

months of July and August of  $3.0 \pm 1.6 \text{ nmol N l}^{-1} \text{ d}^{-1}$  at a depth of 25 m (ref. 32), indicating a pronounced role for *Crocospaera* as a prominent diazotroph during the Lagrangian observations.

Our transcriptional, metabolic and physiological measurements of *Crocospaera* conducted at the cellular level can be used to assess the contribution of this unicellular cyanobacterium to water-column productivity and export within the oligotrophic NPSG ecosystem. The contribution of *Crocospaera* to water-column net community production (NCP) can be estimated by conversion of night-time rates of  $\text{N}_2$  fixation ( $7.3 \text{ nmol N l}^{-1} \text{ d}^{-1}$ ) to units of carbon using the Redfield ratio of 6.6, which yields  $0.05 \text{ } \mu\text{mol C l}^{-1} \text{ d}^{-1}$ . The mean rate of mixed-layer NCP measured using the  $\text{O}_2/\text{Ar}$  method was  $0.45 \pm 0.03 \text{ } \mu\text{mol C l}^{-1} \text{ d}^{-1}$  (Table 1) and therefore we calculate that *Crocospaera* carbon fixation was equivalent to 11% of the rate of NCP. This should be considered an upper estimate, because the sole attribution of night-time rates of  $\text{N}_2$  fixation to *Crocospaera* cannot be proven, as described above. It is noteworthy that the rates of NCP measured during this study are in the upper range of previously reported values for the NPSG at this time of year<sup>33,34</sup>. Furthermore, the elevated rates of NCP within the mesoscale eddy are consistent with elevated concentrations of chlorophyll *a* (Fig. 1c) and increased rates of  $^{14}\text{C}$  assimilation (Fig. 1d), which were also observed in the upper 25 m of the water column relative to the corresponding monthly mean values. In the oligotrophic NPSG, over 70% of primary productivity is accounted for by microorganisms  $<3 \text{ } \mu\text{m}$  in size, and by far the greatest contribution ( $\sim 50\%$ ) is by *Prochlorococcus*, with a smaller collective contribution ( $\sim 14\%$ ) by picoeukaryotes and *Synechococcus*<sup>35,36</sup>. Our observations show that during a bloom, *Crocospaera* can account for an important fraction of both phytoplankton biomass and productivity in the oligotrophic gyre.

An assessment of particle export and its  $\delta^{15}\text{N}$  composition also reflected the importance of a diazotroph-dominated community. Free-drifting sediment traps captured a sinking flux of particulate nitrogen (PN) of  $298 \pm 32 \text{ } \mu\text{mol N m}^{-2} \text{ d}^{-1}$  (mean  $\pm$  s.d.) (Table 1), which was lower than the average value of  $348 \pm 19 \text{ } \mu\text{mol N m}^{-2} \text{ d}^{-1}$  measured for the months of July and August at Station ALOHA. The  $^{15}\text{N}$  isotopic composition of sinking PN had a mean  $\delta^{15}\text{N}$  value of  $2.3 \pm 0.7\text{‰}$ , which was lower than the average of  $3.4 \pm 0.7\text{‰}$  measured for the months of July–August at Station ALOHA, indicating that the contribution of  $\text{N}_2$  fixation to total PN flux was higher than normal. Sinking PN, or ‘export production’, should be equivalent to the new sources of nitrogen in the system, including  $\text{N}_2$  fixation, over sufficient timescales. The measured rate of *Crocospaera*  $\text{N}_2$  fixation, integrated over depths between 0 and 50 m based on the vertical profiles of *Crocospaera nifH* abundance (Fig. 1b), was  $400 \text{ } \mu\text{mol N m}^{-2} \text{ d}^{-1}$ , which exceeds the PN flux (Table 1). This indicates that, during the study period, a large fraction of the net organic matter produced in the mixed layer was accumulating and/or being mineralized before reaching the base of the euphotic zone where the sediment trap was situated. In addition, a significant fraction of the export may occur as dissolved organic matter, which is not captured by the sediment traps<sup>37</sup>.

The unusually high abundance of *Crocospaera* in the anticyclonic eddy is intriguing. One potential explanation for the high abundance is the  $1 \text{ } ^\circ\text{C}$  warmer than usual surface seawater temperature (average seawater temperature at 15 m was  $26.8 \pm 0.1 \text{ } ^\circ\text{C}$ ) due to the 2014–2016 El Niño (Supplementary Fig. 2). The elevated temperature probably favoured *Crocospaera*, as their maximal growth temperatures of  $26\text{--}30 \text{ } ^\circ\text{C}$  exceed the optimal growth temperatures for other diazotrophs such as *Trichodesmium*<sup>38,39</sup> or UCYN-A<sup>13</sup>. In support of this, high abundances of *Crocospaera* have been reported in warm areas of the ocean, including the South Pacific Subtropical Gyre and the West Pacific Warm Pool<sup>7,40,41</sup>. Additional conditions conducive to *Crocospaera* were probably

due to the isopycnal downwelling in anticyclonic eddy interiors, which results in nitrate depletion in near-surface waters and thereby enhances  $\text{N}_2$  fixation<sup>13</sup>. Within the eddy highlighted in this study, dissolved nutrient concentrations were characteristic for the oligotrophic gyre, with  $<10 \text{ nmol kg}^{-1}$  of nitrate in the upper 100 m of the water column (Supplementary Table 1). In addition to macronutrient requirements, trace elements such as iron, the metal cofactor for the nitrogenase enzyme, are essential for diazotroph metabolism. Concentrations of dissolved iron ( $<0.4 \text{ } \mu\text{M}$ ) averaged  $0.42 \pm 0.14 \text{ nmol kg}^{-1}$  at a depth of 15 m (Supplementary Table 1). These values are within the range of dissolved iron concentrations ( $0.31 \pm 0.14 \text{ nmol kg}^{-1}$ ) reported from Station ALOHA<sup>42</sup> during a period when *Crocospaera* was less abundant<sup>43</sup>, and indicate that the high abundances of *Crocospaera* are not likely to be explained by changes in the inventory of dissolved iron. Overall, the nutrient regimes and the hydrography, including slightly higher temperatures and the mesoscale eddy, together appear to have provided an environment conducive to growth of the *Crocospaera* population.

## Discussion

Our study provides an insight into the ecophysiology of *Crocospaera* and its ecological function in the open ocean. Field populations of *Crocospaera* are shown to be highly efficient with respect to the provision of nitrogen for cell growth via  $\text{N}_2$  fixation. Rates of  $\text{N}_2$  fixation slightly exceeded the growth requirements, indicating low quantities of nitrogen are released directly from cells into the ambient environment. Nitrogen fixed by *Crocospaera* is most probably made available to the wider microbial community via microzooplankton grazing or viral lysis, which is evidenced by the prevalent decrease in cell abundances each day in the afternoon. It is therefore likely that over the short term, *Crocospaera*-specific nitrogen is recycled within the upper water column, which, given the oligotrophic status of the NPSG, would have probably contributed to the elevated rates of NCP recorded during this study. Furthermore, the estimated 11% direct contribution of *Crocospaera* to NCP strongly implicates *Crocospaera* as important episodic contributors to ecosystem functioning in the NPSG. This assessment is representative of the short-term dynamics during which this study was conducted, but future studies should also consider the longer-term effects of the growth and decline of *Crocospaera* populations on broader ecosystem functioning. Moreover, our study also highlights the importance of top-down control on *Crocospaera* population abundance, which, in comparison to growth parameters such as light, temperature and nutrients, has received little attention as a regulator of field population abundance, and our results suggest that more in-depth analyses are needed.

At the population level, quantitative transcriptomic analyses revealed a tightly coordinated metabolic choreography of photosynthesis, cell division,  $\text{N}_2$  fixation and other key processes, distinct from those of other abundant marine photoautotrophs such as *Prochlorococcus* and *Synechococcus*. The temporal compression of diverse metabolic activities such as  $\text{N}_2$  fixation, DNA replication, ribosome biosynthesis and production of the photosynthetic apparatus into the night period is probably a repercussion of accommodating both photosynthesis and  $\text{N}_2$  fixation within a unicellular microorganism. The expression patterns observed here were similar to those of monocultures growing in controlled laboratory settings<sup>19</sup> and indicate that, despite the variable field conditions, patterns of gene expression appear remarkably conserved. Some aspects of the diel activity observed here remain to be elucidated at a mechanistic level, such as the delay in cell division until shortly after dawn, which could be due to solar energy acquisition before cell division or another light-responsive metabolic process. Overall, the integration of transcriptomic and genomic analyses with physiological

and biogeochemical measurements demonstrates the utility of transcriptomics as a predictive tool for the timing of key physiological processes throughout the diel cycle. This is notable, because several recent studies coupling transcriptomics and proteomics have shown that transcript levels are in general a poor predictor of protein levels<sup>23,44</sup>. The relative instability of mRNA and correspondingly smaller cellular standing stocks relative to proteins may make transcriptomics a particularly sensitive method when determining the timing of physiological processes in a diel cycle, because mRNA levels at a particular time point are likely to have been produced shortly before sampling. It is likely that the success of this approach will vary with the metabolic process of interest, and further work is required to integrate measurements at the community, population and molecular levels to develop a more quantitative understanding of how these processes are intertwined.

Overall, our multiscale synthesis of molecular, population and community dynamics underscores the complexity of the ecological and biogeochemical forces that underpin the ecophysiological role of *Crocospaera* in oceanic gyres as well as the importance of this keystone diazotroph for broader ecosystem functioning.

## Methods

**Fieldwork design and sampling.** The fieldwork was conducted during the summer (25 July to 5 August 2015) in the oligotrophic NPSG. The overall research objective was to holistically examine diel periodicity in microbial metabolism, with a specific focus on microbial physiology and community interactions that contribute to and result from net carbon accumulation during the day and net consumption during the night. A Lagrangian sampling strategy was implemented to sample the same water mass during the observational period, which was facilitated by deployment of World Ocean Circulation Experiment Surface Velocity Profile (WOCE SVP) drifters (Pacific Gyre) with drogues centred at a depth of 15 m. The target deployment site was the centre of an anticyclonic mode water eddy, to mitigate lateral mixing by different water masses and thereby facilitate observations of the biological variability over diel timescales. The mesoscale eddy fields were assessed before the expedition using animations of satellite altimetry covering 15–30° N, 148–170° W and produced from Archiving, Validation and Interpretation of Satellite Oceanographic (AVISO) data. When the field sampling was conducted, the eddy was located north of the Hawaiian Islands at 24.4° N, 156.5° W, and had a diameter of ~100 km (Supplementary Fig. 1a). Over the 12 day sampling period, the shipboard measurements were conducted alongside the drifters as they performed an almost complete circular pattern with a diameter of ~44 km (Supplementary Fig. 1b). Measurements of horizontal velocity using a hull-mounted Acoustic Doppler Current Profiler (ADCP, RDI Ocean Surveyor 300 kHz) between depths of 11 and 21 m confirmed that the drifter trajectory was in agreement with the near-surface current velocity. The cruise track followed the drifter trajectories, and water-column seawater sampling for diel measurements was carried out every 4 h for periods of 4 days (26–30 July) and 3 days (31 July–3 August) at a depth of 15 m corresponding to the depth of the drogue. To ensure the comparability of measurements with those conducted at nearby Station ALOHA, sampling and analytical protocols for vertical profiles of pigments, nutrients, particulates and flow-cytometry-enumerated phytoplankton populations (*Prochlorococcus*, *Synechococcus* and picoeukaryotes) and heterotrophic bacteria were identical to those used by the HOT program (Supplementary Table 1) (<http://hahana.soest.hawaii.edu/index.html>). Full details are provided in Supplementary Section 'Sampling'.

**Enumeration of *Crocospaera* populations.** The unicellular cyanobacteria were counted using continual underway sampling as well as discrete sample analysis via microscopy and flow cytometry. The continual underway sampling was conducted using SeaFlow<sup>24</sup>, which measured the abundances and cell size of the small *Crocospaera* cells. The large cells were enumerated using an Attune Acoustic Focusing Flow Cytometer, and microscopic measurements were used to verify cell sizes. Full details are provided in Supplementary Section 'Enumeration of *Crocospaera*'. To characterize the N<sub>2</sub>-fixing microorganisms, the *nifH* gene, which encodes a subunit of the nitrogenase enzyme, was measured using quantitative PCR (qPCR). The groups of diazotrophs targeted included UCYN-A, *Crocospaera* spp., *Trichodesmium* spp. and two types of heterocystous cyanobacteria that form symbioses with diatoms (Supplementary Table 1). Discrete seawater samples (2 l) were collected using a CTD-rossette, filtered using a peristaltic pump onto 10 µm polyester (GE Osmotics) and 0.2 µm Supor (Cole Parmer) filters in series, frozen in liquid nitrogen and stored at –80 °C until processed. DNA extraction was conducted using published protocols<sup>45</sup> and qPCR analyses were conducted as previously described<sup>46</sup>.

**Rate measurements.** Rates of N<sub>2</sub> fixation were measured during the cruise using the <sup>15</sup>N<sub>2</sub> assimilation technique. <sup>15</sup>N-labelled gas was dissolved in filtered seawater

before its addition using filtered surface seawater collected at Station ALOHA<sup>47</sup>. The quantities of nitrogen isotopes (that is, N masses equivalent to 28, 29 and 30) were measured in each batch of <sup>15</sup>N<sub>2</sub>-enriched seawater using a membrane inlet mass spectrometer (MIMS)<sup>32</sup>. The final atom % enrichment in the seawater incubations averaged 5.72 ± 0.5. To conduct the rate measurements in the field, 200 ml <sup>15</sup>N<sub>2</sub>-enriched seawater was added to a 4 l polycarbonate bottle that had been filled from a depth of 15 m. Rates of N<sub>2</sub> fixation were measured in triplicate every 4 h during 27–30 July and 31 July–3 August 2015. On termination of the incubation, the entire contents of the 4 l bottle were filtered using a peristaltic pump onto a precombusted glass microfibre (Whatman 25 mm GF/F) filter and stored at –20 °C. On land, the filters were analysed for total mass of N, and δ<sup>15</sup>N composition analysis was carried out using an elemental analyser/isotope ratio mass spectrometer (Carlo-Erba EA NC2500 coupled with a ThermoFinnigan Delta S) at the Stable Isotope Facility, University of Hawaii. Productivity measurements included assimilation of <sup>14</sup>C-labelled bicarbonate (NaH<sup>14</sup>CO<sub>3</sub>) into particulate matter and quantification of the *in situ* ratio of oxygen to argon (O<sub>2</sub>/Ar) using MIMS<sup>48</sup>. Growth and mortality (grazing) rates of *Crocospaera* were determined using a modified dilution method<sup>49</sup>.

**Genomics and transcriptomics.** DNA and RNA was collected throughout the diel time course by filtration of 2 l of seawater onto 25 mm, 0.2 µm Supor PES Membrane Disc filters (Pall) using a peristaltic pump. The filtration time ranged from 15 to 20 min and filters were placed in RNALater (Ambion) immediately afterwards, and preserved at –80 °C until processing. Details regarding DNA and RNA extraction are provided in the Supplementary Methods. Molecular standard mixtures for quantitative transcriptomics were prepared as previously described<sup>50</sup>, and, before RNA purification, 50 µl of each standard group was added to the sample lysate. Metatranscriptomic libraries were prepared for sequencing with the addition of 5–50 ng of total RNA to the ScriptSeq cDNA V2 library preparation kit (Epicentre). Metagenomes were prepared for sequencing using Illumina's TruSeq Nano LT library preparation kit. Metagenomic and metatranscriptomic samples were sequenced with an Illumina Nextseq500 system using V2 high output 300 cycle reagent kit with PHIX control added for metagenomic (1%) and for metatranscriptomic (5%) libraries.

The methods for transcriptome processing are similar to those previously described<sup>18</sup>. Reads were mapped to *Crocospaera* genes identified in the metagenomic data using LAST<sup>51</sup> (95% identity cutoff). Transcripts were quantified through normalization of raw hit counts using molecular standards (Supplementary Methods). Diel oscillations of abundant *Crocospaera* genes in the transcriptomes were identified using Rhythmicity Analysis Incorporating Non-parametric Methods (RAIN)<sup>52</sup>. Network analyses were performed using the R package WGCNA using previously described methods<sup>17</sup>.

**Data availability.** Synoptic cruise information and associated data for the HOE Legacy II cruise can be found online at <http://hahana.soest.hawaii.edu/hoelegacy/hoelegacy.html>. Raw sequence data for the metagenomes and metatranscriptomes reported in this study have been deposited in the NCBI Sequence Read Archive under BioProject ID PRJNA358725.

Received 27 January 2017; accepted 23 June 2017;  
published 31 July 2017

## References

1. Biller, S. J., Berube, P. M., Lindell, D. & Chisholm, S. W. *Prochlorococcus*: the structure and function of collective diversity. *Nat. Rev. Microbiol.* **13**, 13–27 (2015).
2. Waterbury, J. B. & Willey, J. M. in *Methods in Enzymology* Vol. 167 (eds Packer, L. & Glazer, A. N) 100–105 (Academic, 1988).
3. Bench, S. R. *et al.* Whole genome comparison of six *Crocospaera watsonii* strains with differing phenotypes. *J. Phycol.* **49**, 786–801 (2013).
4. Zehr, J. P., Bench, S. R., Mondragon, E. A., McCarren, J. & DeLong, E. F. Low genomic diversity in tropical oceanic N<sub>2</sub>-fixing cyanobacteria. *Proc. Natl Acad. Sci. USA* **104**, 17807–17812 (2007).
5. Tuit, C., Waterbury, J. & Ravizza, G. Diel variation of molybdenum and iron in marine diazotrophic cyanobacteria. *Limnol. Oceanogr.* **49**, 978–990 (2004).
6. Zehr, J. P. *et al.* Unicellular cyanobacteria fix N<sub>2</sub> in the subtropical North Pacific Ocean. *Nature* **412**, 635–638 (2001).
7. Moisaner, P. H. *et al.* Unicellular cyanobacterial distributions broaden the oceanic N<sub>2</sub> fixation domain. *Science* **327**, 1512–1514 (2010).
8. Sato, M. *et al.* Distribution of nano-sized *Cyanobacteria* in the western and central Pacific Ocean. *Aquat. Microbial Ecol.* **59**, 273–282 (2010).
9. Robidart, J. C. *et al.* Ecogenomic sensor reveals controls on N<sub>2</sub>-fixing microorganisms in the North Pacific Ocean. *ISME J.* **8**, 1175–1185 (2014).
10. Webb, E. A. *et al.* Phenotypic and genotypic characterization of multiple strains of the diazotrophic cyanobacterium, *Crocospaera watsonii*, isolated from the open ocean. *Environ. Microbiol.* **11**, 338–348 (2009).



11. Mohr, W., Intermaggio, M. P. & LaRoche, J. Diel rhythm of nitrogen and carbon metabolism in the unicellular, diazotrophic cyanobacterium *Crocosphaera watsonii* WH8501. *Environ. Microbiol.* **12**, 412–421 (2010).
12. Saito, M. A. *et al.* Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph *Crocosphaera watsonii*. *Proc. Natl Acad. Sci. USA* **108**, 2184–2189 (2011).
13. Church, M. J. *et al.* Physical forcing of nitrogen fixation and diazotroph community structure in the North Pacific subtropical gyre. *Global Biogeochem. Cycles* **23**, GB2020 (2009).
14. Bench, S. R., Frank, I., Robidart, J. & Zehr, J. P. Two subpopulations of *Crocosphaera watsonii* have distinct distributions in the North and South Pacific. *Environ. Microbiol.* **18**, 514–524 (2016).
15. Bench, S. R., Ilikchyan, I. N., Tripp, H. J. & Zehr, J. P. Two strains of *Crocosphaera watsonii* with highly conserved genomes are distinguished by strain-specific features. *Front. Microbiol.* **2**, 261 (2011).
16. Sohm, J. A., Edwards, B. R., Wilson, B. G. & Webb, E. A. Constitutive extracellular polysaccharide (EPS) production by specific isolates of *Crocosphaera watsonii*. *Front. Microbiol.* **2**, 229 (2011).
17. Langfelder, P. & Horvath, S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
18. Aylward, F. O. *et al.* Microbial community transcriptional networks are conserved in three domains at ocean basin scales. *Proc. Natl Acad. Sci. USA* **112**, 5443–5448 (2015).
19. Shi, T., Ilikchyan, I., Rabouille, S. & Zehr, J. P. Genome-wide analysis of diel gene expression in the unicellular N<sub>2</sub>-fixing cyanobacterium *Crocosphaera watsonii* WH8501. *ISME J.* **4**, 621–632 (2010).
20. Ottesen, E. A. *et al.* Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science* **345**, 207–212 (2014).
21. Ottesen, E. A. *et al.* Pattern and synchrony of gene expression among sympatric marine microbial populations. *Proc. Natl Acad. Sci. USA* **110**, E488–E497 (2013).
22. Stöckel, J. *et al.* Global transcriptomic analysis of *Cyanothece* 51142 reveals robust diurnal oscillation of central metabolic processes. *Proc. Natl Acad. Sci. USA* **105**, 6156–6161 (2008).
23. Welkie, D. *et al.* Transcriptomic and proteomic dynamics in the metabolism of a diazotrophic cyanobacterium, *Cyanothece* sp. PCC 7822 during a diurnal light–dark cycle. *BMC Genomics* **15**, 1185 (2014).
24. Swalwell, J. E., Ribalet, F. & Armbrust, E. V. SeaFlow: a novel underway flow-cytometer for continuous observations of phytoplankton in the ocean. *Limnol. Oceanogr. Methods* **9**, 466–477 (2011).
25. Duman, R. *et al.* Structural and genetic analyses reveal the protein *SepF* as a new membrane anchor for the Z ring. *Proc. Natl Acad. Sci. USA* **110**, E4601–E4610 (2013).
26. Goebel, N. L. *et al.* Growth and carbon content of three different-sized diazotrophic cyanobacteria observed in the subtropical North Pacific. *J. Phycol.* **44**, 1212–1220 (2008).
27. Nielsen, S. L. Size-dependent growth rates in eukaryotic and prokaryotic algae exemplified by green algae and cyanobacteria: comparisons between unicells and colonial growth forms. *J. Plankton Res.* **28**, 489–498 (2006).
28. Church, M. J., Short, C. M., Jenkins, B. D., Karl, D. M. & Zehr, J. P. Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. *Appl. Environ. Microbiol.* **71**, 5362–5370 (2005).
29. Mulholland, M. R., Bronk, D. A. & Capone, D. G. Dinitrogen fixation and release of ammonium and dissolved organic nitrogen by *Trichodesmium* IMS101. *Aquat. Microbial Ecol.* **37**, 85–94 (2004).
30. Wilson, S. T. *et al.* Hydrogen cycling by the unicellular marine diazotroph *Crocosphaera watsonii* strain WH8501. *Appl. Environ. Microbiol.* **20**, 6797–6803 (2010).
31. Inomura, K., Bragg, J. & Follows, M. J. A quantitative analysis of the direct and indirect costs of nitrogen fixation: a model based on *Azotobacter vinelandii*. *ISME J.* **11**, 166–175 (2016).
32. Böttjer, D. *et al.* Temporal variability in dinitrogen fixation and particulate nitrogen export at Station ALOHA. *Limnol. Oceanogr.* **62**, 200–216 (2017).
33. Juranek, L. W. & Quay, P. D. *In vitro* and *in situ* gross primary production and net community production in the North Pacific Subtropical Gyre using labeled and natural abundance isotopes of dissolved O<sub>2</sub>. *Global Biogeochem. Cycles* **19**, GB3009 (2005).
34. Juranek, L. W. *et al.* Biological production in the NE Pacific and its influence on air–sea CO<sub>2</sub> flux: evidence from dissolved oxygen isotopes and O<sub>2</sub>/Ar. *J. Geophys. Res.* **117**, C05022 (2012).
35. Campbell, L. & Vaulot, D. Photosynthetic picoplankton community structure in the subtropical North Pacific Ocean near Hawaii (Station ALOHA). *Deep-Sea Res. I* **40**, 2043–2060 (1993).
36. Rii, Y. M., Karl, D. M. & Church, M. J. Temporal and vertical variability in picophytoplankton primary productivity in the North Pacific Subtropical Gyre. *Mar. Ecol. Prog. Ser.* **562**, 1–18 (2016).
37. Emerson, S. Annual net community production and the biological carbon flux in the ocean. *Global Biogeochem. Cycles* **28**, 14–28 (2014).
38. Boyd, P. W. *et al.* Marine phytoplankton temperature versus growth responses from polar to tropical waters—outcome of a scientific community-wide study. *PLoS ONE* **8**, e63091 (2013).
39. Fu, F. X. *et al.* Differing responses of marine N<sub>2</sub> fixers to warming and consequences for future diazotroph community structure. *Aquat. Microb. Ecol.* **72**, 33–46 (2014).
40. Bonnet, S., Biegala, I. C., Dutrieux, P., Slemmons, L. O. & Capone, D. G. Nitrogen fixation in the western equatorial Pacific: rates, diazotrophic cyanobacterial size class distribution, and biogeochemical significance. *Global Biogeochem. Cycles* **23**, GB3012 (2009).
41. Hewson, I. *et al.* *In situ* transcriptomic analysis of the globally important keystone N<sub>2</sub>-fixing taxon *Crocosphaera watsonii*. *ISME J.* **3**, 618–631 (2009).
42. Fitzsimmons, J. N. *et al.* Daily to decadal variability of size-fractionated iron and iron-binding ligands at the Hawaii Ocean Time-series Station ALOHA. *Geochim. Cosmochim. Acta* **171**, 303–324 (2015).
43. Wilson, S. T. *et al.* Short-term variability in euphotic zone biogeochemistry and primary productivity at Station ALOHA: a case study of summer 2012. *Global Biogeochem. Cycles* **29**, 1145–1164 (2015).
44. Waldbauer, J. R., Rodrigue, S., Coleman, M. L. & Chisholm, S. W. Transcriptome and proteome dynamics of a light–dark synchronized bacterial cell cycle. *PLoS ONE* **7**, e34342 (2012).
45. Moisaner, P. H., Beinart, R. A., Voss, M. & Zehr, J. P. Diversity and abundance of diazotrophic microorganisms in the South China Sea during intermonsoon. *ISME J.* **2**, 954–967 (2008).
46. Goebel, N. L. *et al.* Abundance and distribution of major groups of diazotrophic cyanobacteria and their potential contribution to N<sub>2</sub> fixation in the tropical Atlantic Ocean. *Environ. Microbiol.* **12**, 3272–3289 (2010).
47. Wilson, S. T., Böttjer, D., Church, M. J. & Karl, D. M. Comparative assessment of nitrogen fixation methodologies conducted in the oligotrophic North Pacific Ocean. *Appl. Environ. Microbiol.* **78**, 6491–6498 (2012).
48. Ferrón, S., Wilson, S. T., Martínez-García, S., Quay, P. D. & Karl, D. M. Metabolic balance in the mixed layer of the oligotrophic North Pacific Ocean from diel changes in O<sub>2</sub>/Ar saturation ratios. *Geophys. Res. Lett.* **42**, 3421–3430 (2015).
49. Landry, M. R., Kirshtein, J. & Constantinou, J. A refined dilution technique for measuring the community grazing impact of microzooplankton, with experimental tests in the central equatorial Pacific. *Mar. Ecol. Prog. Ser.* **120**, 53–63 (1995).
50. Gifford, S. M., Becker, J. W., Sosa, O. A., Repeta, D. J. & DeLong, E. F. Quantitative transcriptomics reveals the growth- and nutrient-dependent response of a streamlined marine methylotroph to methanol and naturally occurring dissolved organic matter. *mBio* **7**, e01279–16 (2016).
51. Kielbasa, S. M., Wan, R., Sato, K., Horton, P. & Frith, M. C. Adaptive seeds tame genomic sequence comparison. *Genome Res.* **21**, 487–493 (2011).
52. Thaben, P. F. & Westermark, P. O. Detecting rhythms in time series with RAIN. *J. Biol. Rhythms* **29**, 391–400 (2014).
53. Menden-Deuer, S. & Lessard, E. J. Carbon to volume relationships for dinoflagellates, diatoms, and other protist plankton. *Limnol. Oceanogr.* **45**, 569–579 (2000).
54. Laws, E. A. Photosynthetic quotients, new production and net community production in the open ocean. *Deep-Sea Res.* **38**, 143–167 (1991).
55. Langfelder, P. & Horvath, S. Eigengene networks for studying the relationships between co-expression modules. *BMC Syst. Biol.* **1**, 54 (2007).

## Acknowledgements

The dataset presented here resulted from the efforts of many scientists who contributed to the success of the 2015 expedition. The authors thank T. Clemente for cruise leadership of KOK1507, J. Collins, J. Ossolinski and B. Van Mooy for net trap samples used for  $\delta^{15}\text{N}$  isotope analysis and E. Boyle for support of the dissolved iron measurements. For assistance with field and laboratory work, the authors thank the operational staff of the Simons Collaboration on Ocean Processes and Ecology (SCOPE), D. Böttjer, P. Den Uyl, L. Jensen, N. Lanning, M. Linney and A. Nelson. This work was supported by grants from the Simons Foundation (no. 329108 to D.M.K. and E.F.D.), the Gordon and Betty Moore Foundation (no. 3777 to E.F.D., no. 3794 to D.M.K., no. 3776 to E.V.A.) and the Balzan Prize for Oceanography (to D.M.K.). In addition, the authors acknowledge the National Science Foundation for support of the HOT programme (OCE1260164 to M.J.C. and D.M.K.) and the Center for Microbial Oceanography: Research and Education (C-MORE; EFO424599 to D.M.K. and E.F.D.). This work is a contribution of SCOPE and C-MORE.

## Author contributions

All authors contributed to the design of the study. S.T.W. and D.M.K. measured nitrogen fixation and provided the water-column hydrography and biogeochemical data. F.O.A., A.E.R., A.V., J.M.E. and E.F.D. sampled, prepared and analysed the metatranscriptomic and metagenomic data. F.R. and E.V.A. conducted the underway enumeration of *Crocosphaera* abundances. B.B. and F.R. quantified the abundances of larger size *Crocosphaera*. D.A.C. and P.E.C. performed the microscopy analyses and dilution grazing experiments. S.F. and J.R.C. conducted the productivity measurements. M.J.C. provided the time-series *nifH* abundances and measured particle export. J.R.C. and B.B. collected and analysed the isotopic composition of sinking particles. K.A.T.-K. and J.P.Z. analysed the *nifH*



abundances. C.T.H. and J.N.F. measured dissolved iron concentrations. S.T.W., F.O.A., D.M.K. and E.F.D. wrote the manuscript with contributions from all coauthors.

### Additional information

Supplementary information is [available for this paper](#).

Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints).

Correspondence and requests for materials should be addressed to E.F.D.

**How to cite this article:** Wilson, S. T. *et al.* Coordinated regulation of growth, activity and transcription in natural populations of the unicellular nitrogen-fixing cyanobacterium *Crocospaera*. *Nat. Microbiol.* **2**, 17118 (2017).

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

### Competing interests

The authors declare no competing financial interests.