

Specialized proteomic responses and an ancient photoprotection mechanism sustain marine green algal growth during phosphate limitation

Jian Guo¹, Susanne Wilken^{1,7}, Valeria Jimenez^{1,2}, Chang Jae Choi¹, Charles Ansong³, Richard Dannebaum^{1,4}, Lisa Sudek¹, David S. Milner⁵, Charles Bachy¹, Emily Nahas Reistetter¹, Virginia A. Elrod¹, Denis Klimov¹, Samuel O. Purvine³, Chia-Lin Wei^{4,8}, Govindarajan Kunde-Ramamoorthy^{4,8}, Thomas A. Richards⁵, Ursula Goodenough⁶, Richard D. Smith³, Stephen J. Callister³ and Alexandra Z. Worden^{1,2*}

Marine algae perform approximately half of global carbon fixation, but their growth is often limited by the availability of phosphate or other nutrients^{1,2}. As oceans warm, the area of phosphate-limited surface waters is predicted to increase, resulting in ocean desertification^{3,4}. Understanding the responses of key eukaryotic phytoplankton to nutrient limitation is therefore critical^{5,6}. We used advanced photo-bioreactors to investigate how the widespread marine green alga *Micromonas commoda* grows under transitions from replete nutrients to chronic phosphate limitation and subsequent relief, analysing photosystem changes and broad cellular responses using proteomics, transcriptomics and biophysical measurements. We find that physiological and protein expression responses previously attributed to stress are critical to supporting stable exponential growth when phosphate is limiting. Unexpectedly, the abundance of most proteins involved in light harvesting does not change, but an ancient light-harvesting-related protein, LHCSR, is induced and dissipates damaging excess absorbed light as heat throughout phosphate limitation. Concurrently, a suite of uncharacterized proteins with narrow phylogenetic distributions increase multifold. Notably, of the proteins that exhibit significant changes, 70% are not differentially expressed at the mRNA transcript level, highlighting the importance of post-transcriptional processes in microbial eukaryotes. Nevertheless, transcript–protein pairs with concordant changes were identified that will enable more robust interpretation of eukaryotic phytoplankton responses in the field from metatranscriptomic studies. Our results show that P-limited *Micromonas* responds quickly to a fresh pulse of phosphate by rapidly increasing replication, and that the protein network associated with this ability is composed of both conserved and phylogenetically recent proteome systems that promote dynamic phosphate homeostasis. That an ancient mechanism for mitigating light stress is central to sustaining growth during extended phosphate limitation highlights the possibility of interactive effects arising from combined stressors under

ocean change, which could reduce the efficacy of algal strategies for optimizing marine photosynthesis.

Understanding phytoplankton acclimation to variations in surface ocean conditions is important for predicting future trajectories of phytoplankton lineages, photosynthetic uptake of CO₂ and marine food webs^{5,6}. However, these processes are poorly understood, especially for the dominant eukaryotic algae in the open-ocean regions where desertification is expected, which belong to the picoplankton size class ($\leq 2\mu\text{m}$ cell diameter)^{3,4}. The picoeukaryote *Micromonas commoda* provides an attractive model for investigating acclimation because it carries a small genome (21 Mb) containing just 10,306 protein-encoding genes, and is related to well-characterized land plants and chlorophyte algae (Fig. 1a)⁷. It belongs to the widespread prasinophyte class Mamiellophyceae⁸, which are abundant in coastal, frontal and open-ocean environments, such as Station ALOHA in the North Pacific, where they comprise $34 \pm 13\%$ (annual average) of photosynthetic eukaryotes at the deep chlorophyll maximum^{9–11}. These algae can also perform a significant fraction of picoplanktonic primary production¹². *Micromonas* itself is a tiny, flagellated, fast-swimming genus found in varied ecosystems from tropical to polar^{13,14}, including the phosphate-limited Sargasso Sea in the Atlantic^{15–17}. Despite their ecological importance, molecular remodelling has not been elucidated in picoeukaryotic phytoplankton growing under nutrient limitation, in part due to the challenges entailed in provisioning algae with continuous-but-limiting nutrient supplies.

To capture how *Micromonas* acclimates to periods of sustained but growth-limiting phosphate supplies that typify ocean seasonality³, we performed experiments in custom-built photo-bioreactors, with controlled, continuously flowing medium, that also provided in situ monitoring of photosynthetic efficiency and other parameters (Fig. 1b and Supplementary Fig. 1). *M. commoda* was shifted from nutrient-replete conditions to a limiting phosphate supply, ultimately resulting in a stable, acclimated P_{limited} phase with slower exponential growth ($0.28 \pm 0.05 \text{ d}^{-1}$; $P < 0.001$) than the initial P_{replete} phase ($0.82 \pm 0.02 \text{ d}^{-1}$), and differing photosynthetic efficiency of photosystem (PS) II (Fig. 1c,d and Supplementary Data 1). During

¹Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA. ²Ocean Sciences Department, University of California Santa Cruz, Santa Cruz, CA, USA. ³Pacific Northwest National Laboratory, Richland, WA, USA. ⁴Joint Genome Institute, Lawrence Berkeley National Laboratory, Walnut Creek, CA, USA. ⁵University of Exeter, Exeter, UK. ⁶Department of Biology, Washington University, St Louis, MO, USA. ⁷Present address: Department of Freshwater and Marine Ecology, University of Amsterdam, Amsterdam, the Netherlands. ⁸Present address: The Jackson Laboratory, Farmington, CT, USA. *e-mail: azworden@mbari.org

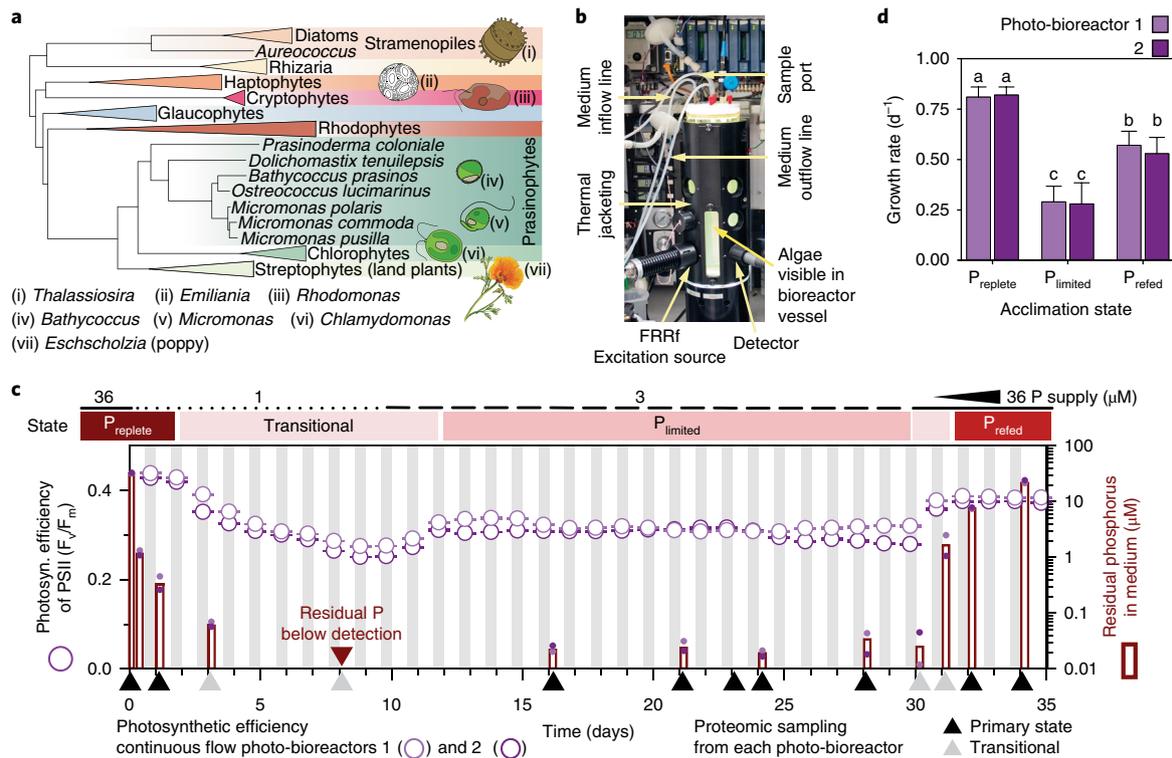


Fig. 1 | Algal relationships and *Micromonas* growth under changes in phosphate availability. **a**, Relationships between photosynthetic eukaryotes, including *Micromonas*, other Mamiellophyceae (class II prasinophytes, specifically *Ostreococcus*, *Bathycoccus* and *Dolichomastix*) and more distant phytoplankton groups. Together, chlorophytes and prasinophytes are termed green algae. This schematic tree is based on a multigene phylogenetic analysis²⁶ and transcriptome sequencing project²⁷. **b**, The photo-bioreactors, of which one growth vessel is shown, are designed to examine the impact of environmental factors on algal growth and physiology, incorporating aspects of both traditional turbidostat and chemostat systems. **c**, Residual phosphate in the medium after *M. commoda* growth (right y axis, magenta bars) and photosynthetic efficiency of PSII (F_v/F_m , left y axis, circles). Simulated day (14 h, white) and night (10 h, grey) periods are indicated by vertical bars. Horizontal bars at top indicate physiological states (coloured) resulting from P supply (black) changes. Cells in two independent photo-bioreactors were first shifted from nutrient-replete conditions to low phosphate (36-fold reduction); over the next 8 days, biological drawdown reduced residual phosphate in the medium to below detection limits ($0.01 \mu\text{M}$; transitional period). The P supply was then increased 3-fold, resulting in acclimated P-limited growth (17-day P_{limited} phase). Triangles below the x axis mark primary (black) and transitional (grey) phase proteomic sampling time points. Bars (magenta, note log scale) represent mean soluble reactive phosphate from biological duplicates and dots show individual measurements in photo-bioreactors 1 (light purple) and 2 (purple). F_v/F_m is shown as the mean and standard deviation of 27 FRRf measurements performed 30 to 60 min after lights-off each night. **d**, Flow cytometry counts showed an average growth rate of $0.82 \pm 0.02 \text{ d}^{-1}$ over 10 generations leading up to and through the initial days in photo-bioreactors (P_{replete}), which started to decline 3 days after switching to low P (day 3, $0.74 \pm 0.02 \text{ d}^{-1}$). Growth rate during P_{limited} was $0.28 \pm 0.05 \text{ d}^{-1}$, increased upon initiation of refeeding (day 30), and stabilized at a lower rate in P_{refed} ($0.55 \pm 0.04 \text{ d}^{-1}$) than P_{replete} due to increased dilution (pump rate increased to 0.5 d^{-1}) starting two days after the supply vessel change for refeeding. Bars and error bars represent geometric means and standard deviations of growth rates from within phase biological replicates in each photo-bioreactor for P_{replete} ($n=7$), P_{limited} ($n=13$) and P_{refed} ($n=3$) and results of a two-way ANOVA comparing the effect of phosphate supply (on growth) per photo-bioreactor vessel are indicated (letters, $P < 0.001$). Additionally, the P supply effect had a significant impact on growth rates between phases ($P < 0.005$, two-tailed one sample t -test, $n=4$).

the P_{limited} phase, residual phosphate was similar to surface concentrations in the Sargasso Sea^{2,17} and the *M. commoda* growth rate was similar to in situ Sargasso rates for a *Micromonas*-containing population (0.22 d^{-1}) during early summer¹⁵. This allowed us to investigate physiological and proteomic responses in a marine alga growing under chronic nutrient limitation. Previous studies, by contrast, evaluate abrupt nutrient deprivation^{18–21}, and hence monitor the transition into stationary phase and/or death²².

Expression of abundant proteins that changed significantly (≥ 2 -fold, $Q < 0.05$) under P_{limited} growth relative to replete cells followed two distinct patterns. Of those that decreased in relative abundance under P limitation ('DOWN' cluster, Fig. 2a), 94% had known functions or recognizable domains, such as ribosomal proteins (48%), and other proteins anticipated to change under reduced growth (Supplementary Fig. 2 and Supplementary Data 2). Proteins involved in nitrogen acquisition and assimilation (4%) also

decreased (Supplementary Fig. 3). Modulation of nitrogen-metabolism-related proteins has been shown in other marine algae when phosphate or micronutrients become unavailable, including important stramenopiles (Fig. 1a; specifically diatoms and the pelagophyte *Aureococcus*)^{18,23} and the cyanobacterium *Prochlorococcus*²⁴.

Among proteins with higher relative abundance in the P_{limited} phase (Fig. 2a, UP cluster), the largest increase (77-fold change) entailed a homologue of a phosphate-repressible alkaline phosphatase secreted by the chlorophyte *Chlamydomonas reinhardtii* to cleave phosphate from organic compounds²⁵, followed by an uncharacterized protein with a domain present in some phosphatase classes (Supplementary Fig. 4a). Other phosphate metabolism and transport proteins also increased, alongside unknown-function transporters, a putative chromate transporter and a nitrogen regulatory protein (Supplementary Fig. 4b–e). Protein expression profiles indicated that during P_{limited} , *Micromonas* bypassed the

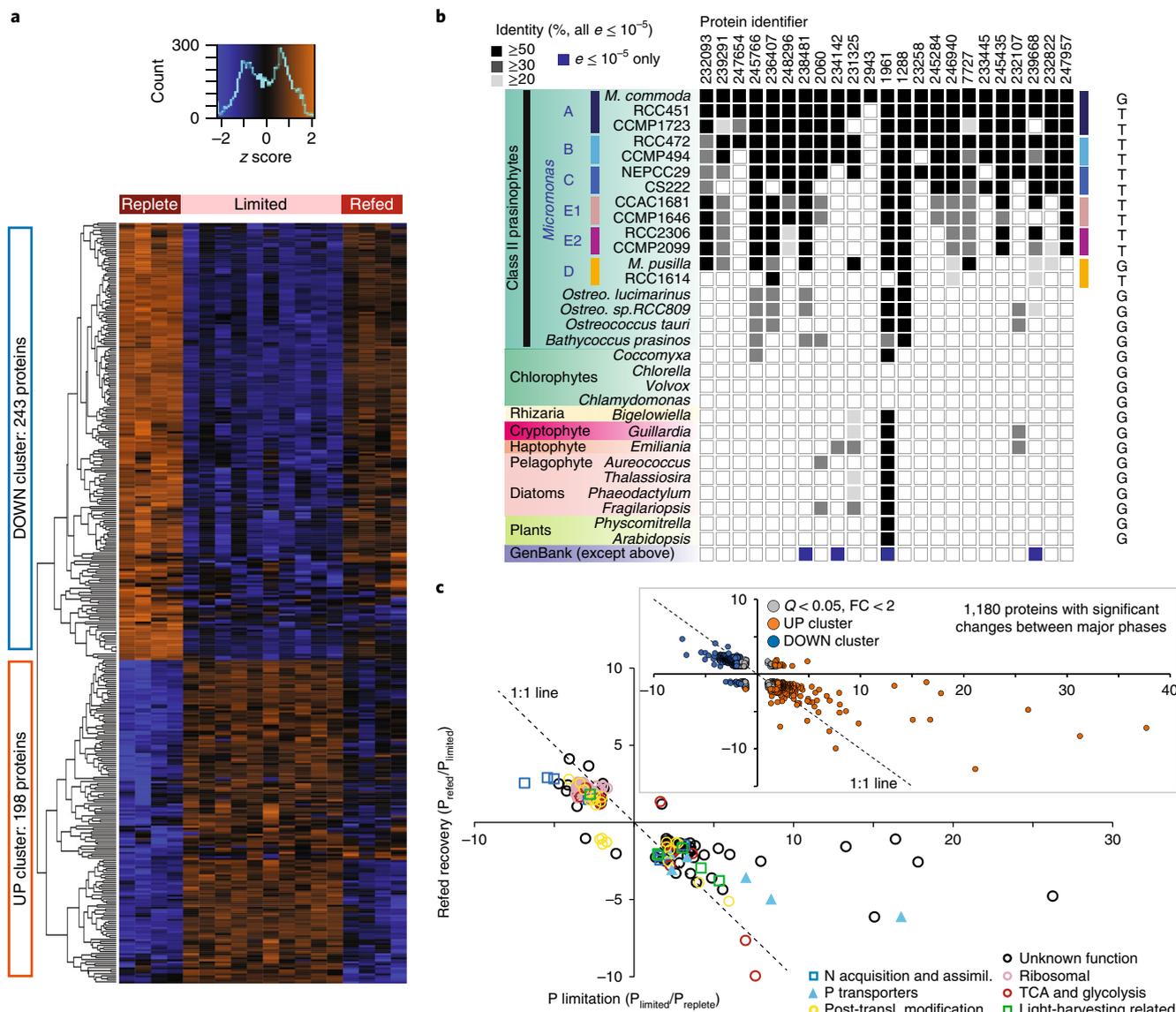


Fig. 2 | Proteomic changes under acclimated P-limited growth. **a**, Hierarchical clustering of z-scored log₂-transformed protein abundance data (LC-MS) from 18 samples (9 per photo-bioreactor) identified two response groups (UP versus DOWN with limitation) among the 441 proteins that were differentially expressed (≥2-fold change, Q < 0.05; Q value being the false discovery rate adjusted P value) between the P_{replete}, P_{limited} or P_{refed} phases out of the 2,118 detected reliably. Observed proteins: 4,954; proteins with different relative abundances between major phases (Q < 0.05): 1,180. P_{replete} (T_{0r}, T₁; n = 4) and P_{refed} (T₃₂, T₃₄; n = 4) consist of within phase biological proteomic replicates from each photo-bioreactor sampled on two days per phase, while P_{limited} (n = 10) is represented by sampling on 5 days (T₁₆, T₂₁, T₂₃, T₂₄ and T₂₈) per photo-bioreactor. A two-tailed t-test was used to determine significant differences between phases (Supplementary Data 2a,b). In addition to many ribosomal proteins (48%), the DOWN cluster included, for example, proteins involved in translation initiation/elongation (8%) and the light-independent reactions of photosynthesis (4%). In the DOWN cluster the greatest reductions were exhibited (>−60 fold change) by a DNA repair ATPase and a unique class II prasinophyte protein lacking recognizable domains; the two proteins with highest fold changes in the UP cluster were a phosphatase and an uncharacterized protein with a domain present in some phosphatase classes and hence considered here a potential phosphatase (Supplementary Fig. 4a). Z-scored data and hierarchical clustering of all samples (26 total), including transitional periods, showed greater variability during transitions, as expected (Supplementary Fig. 2). **b**, Taxonomic distributions of proteins that increased strongly (≥3 fold change, Q < 0.005) under P limitation and lack recognized domains, shows some are restricted to *Micromonas* (Supplementary Data 2c). Distributions are based on amino acid identities of ≥50% (black), ≥30% (dark grey) and ≥20% (light grey) with e-values ≤ 10^{−5}, e-value ≤ 10^{−5} only (blue) and no hit (white) in predicted proteomes from genome-sequenced organisms (G) and transcriptomes²⁷ (T) (Supplementary Data 2d) as well as GenBank’s non-redundant database. **c**, Recovery observed for proteins in select pathways and for unknown function proteins (259 proteins, r² = 0.8769) after increasing the P supply (P_{refed} relative to P_{limited}, y axis) as compared to P_{limited} relative to P_{replete} (x axis). The inset shows recovery patterns for all proteins (r² = 0.5068) that changed significantly (1,180 in total); six are excluded from the figure, two with >40 and four with <−10 fold changes in P_{limited} relative to P_{replete}. Fold change (FC) on the 1:1 line indicates full recovery and values in this figure are not log transformed.

final pyruvate-producing step of glycolysis and instead used three non-phosphorus-dependent reactions to produce pyruvate (Supplementary Fig. 5), as seen in plants²⁶ and proposed for

*Aureococcus*¹⁸. Increased UDP-sulfoquinovose synthase indicated membrane remodelling from phospholipids to sulfolipids (Supplementary Fig. 6), an adaptation also documented in

P-starved stramenopile algae^{18,20}, marine cyanobacteria and heterotrophic bacteria⁶, plants and *Chlamydomonas*^{25,26}. Several of these observed adjustments thus appear to reflect evolutionarily similar strategies, spanning two domains of life, to acclimate to phosphate scarcity. However, in contrast to the algal-wide commonalities seen for proteins in the DOWN cluster^{18,20}, half the UP cluster proteins had unknown functions (Supplementary Data 2b). Among those that increased strongly (≥ 3 -fold change, $Q < 0.005$) under P limitation, 20% had 'domains of unknown function' and 35% had no recognizable domains. Homologues of the latter were generally lacking in non-prasinophyte algae and other organisms, and were largely restricted to the *Micromonas* genus¹³ (Fig. 2b and Supplementary Data 2c,d), despite the availability of considerable genomic information from other prasinophytes²⁷. *Micromonas* therefore encodes specialized responses that are either co-regulated with P limitation, or reflect specific adaptations for growth under P limitation.

Following the P_{limited} phase, the P supply was slowly increased (re-feeding), and after 2 days the photo-bioreactor dilution rate was adjusted to hold growth at 0.5 d^{-1} (Fig. 1c,d). Protein levels rapidly readjusted such that P_{refed} relative abundances were closer to P_{replete} than P_{limited} (Fig. 2c) even though residual phosphate had yet to reach P_{replete} levels. Within 6 hours after the change to replete medium in the photo-bioreactor supply vessel, the proteomic profile adjusted in the direction of relative abundances observed in P_{replete} (Supplementary Fig. 2b), even though the photo-bioreactors had at that point received at most two-fold higher P than during P_{limited} . Moreover, the growth rate increased less than 24 hours after the supply switch, and doubled within 48 hours. These responses are markedly faster than in studies of P-deprived marine algae^{18,20}, even though we used a light-dark cycle while others typically use continuous light, which enables faster replication²⁸. Thus, P-limited *Micromonas* cells were poised to shift rapidly into faster exponential growth following a pulse of the limiting nutrient, as might occur in nature.

The relationship between transcript and protein expression is poorly characterized in marine eukaryotes. In plants^{26,29} and *Chlamydomonas*^{19,25}, many gene transcripts and their protein products have ill-matched expression profiles, as do clade D species *Micromonas pusilla* over the diurnal cycle³⁰, and the diatom *Thalassiosira pseudonana* under nutrient deprivation²⁰. Herein, most transcripts (70%) encoding *M. commoda* proteins that changed between P_{limited} and P_{replete} (Fig. 2a) showed < 2 -fold change, not meeting our differential expression criterion (Supplementary Figs. 7 and 8), including all those involved in amino acid transport and metabolism and most ribosomal proteins. The set of significantly changed proteins with differentially expressed transcripts (30%, Fig. 3a,b) was enriched for proteins involved in nitrogen acquisition and metabolism, post-translational modifications and turnover, and the picoprasinophyte proteins with limited phylogenetic distributions noted in the P_{limited} UP cluster. Additionally, for this set, transcript fold change between P_{limited} and P_{replete} was correlated with fold change in the encoded protein ($r^2 = 0.69$), and 95% changed concordantly (Fig. 3a and Supplementary Data 3). That most transcripts did not exhibit differential expression, although their protein products did, highlights the importance of post-transcriptional regulation in marine algae, possibly including RNA methylation, as seen in other eukaryotes^{30,31}. These processes impose limitations on interpretation of field metatranscriptomes, making it essential to identify concordant transcript-protein pairs.

We also studied *M. commoda* under nitrogen-depletion (Supplementary Fig. 9) to distinguish the effects of phosphate-limitation from generalized responses to growth reduction for the concordant transcript-protein pairs. RNA-seq showed that transcripts for most (72%) of the P-experiment concordant set changed ($P < 0.05$) under N depletion, and 55% showed ≥ 2 -fold change relative to the N-replete control (Fig. 3c, Supplementary Datafile 3b).

Those changing in the same direction (51%) presumably reflect adjustments for reducing growth rate, while those in the opposite direction (49%) reflect cellular balancing related to macronutrient availability. The latter includes the chromate transporter family protein that had shown > 10 -fold protein increase under P limitation relative to P_{replete} (Supplementary Fig. 4b). Given its opposing responses under N depletion and P limitation, we tested whether this protein transports phosphate. The *Micromonas* gene indeed rescued growth of a *Saccharomyces cerevisiae* mutant devoid of the Pho84 phosphate transporter and maintained in low phosphate medium (Supplementary Fig. 10). We therefore named it Mipt1 (*Micromonas* phosphate transporter 1). Notably, Mipt1 expression decreased with N depletion, unlike the classical sodium-dependent phosphate transporter in *M. commoda* which was expressed constitutively in the N-depletion experiment (Supplementary Datafile 3b). Hence, Mipt1 may provide an alternative to Pho4, the high affinity phosphate transporter present in other prasinophytes^{32,33} but lacking or highly diverged in *Micromonas*. Overall, this and other concordant transcript-protein pairs provide baseline information for deriving conclusions from metatranscriptomes on algal cellular modifications and growth changes in nature, where the controlling nutrients themselves are often below assay detection limits².

We also investigated photosynthetic processes because of their importance for understanding primary production and because of the unusual patterns³⁴ observed for photosynthetic efficiency of PSII during the photo-bioreactor time course (F_v/F_m , Fig. 1c). Reduced rates of carbon fixation, as might occur under nutrient limitation, pose the risk of photo-damage^{25,26}. One way to avoid such damage is to adjust the major Light Harvesting Complex (LHC) proteins, which bind chlorophyll and form the antenna complex that collects light energy for photosynthesis; by reducing the overall antenna size, photo-damage is diminished. Prior results on expression of major LHC proteins, or genes encoding them, are highly inconsistent in nutrient-stress studies on other marine algae, including the picoprasinophyte *Ostreococcus*, haptophytes and stramenopiles (Fig. 1a), perhaps reflecting differences in the starvation protocols or analysis methods used^{18,20,21,23,33,35,36}. Photosynthetic efficiency of PSII (F_v/F_m) itself in algae often shows a temporary reduction, then recovery upon acclimation to the limiting nutrient^{34,37}. Under transient high-light stress in plants and *Chlamydomonas*, non-photochemical quenching (NPQ) drives similar temporary F_v/F_m shifts^{38,39}. NPQ dissipates harmful excess absorbed light energy as heat, reducing high-light stress during shade-sun transitions or until major adjustments are made to the photosynthetic apparatus during long-term shifts to high-light. In green algae and diatoms NPQ is mediated by an ancient, distant LHC protein family member, LHCSR (LHC-stress related/LHCX; Supplementary Fig. 11)^{28,39,40}. Xanthophyll cycle-based NPQ (involving the pigments violaxanthin, antheraxanthin and zeaxanthin; VAZ) also occurs in these organisms and multiple prasinophytes^{41,42}.

In our study of *M. commoda*, major LHC proteins did not change significantly during acclimated P_{limited} growth, nor did detected photosystem I and II proteins. However, photosynthetic efficiency of PS II changed significantly as a function of P availability ($P < 0.001$) and was stable within each phase, being lower during P_{limited} (0.310 ± 0.002) than P_{replete} (0.429 ± 0.008) or P_{refed} (0.378 ± 0.006). Biophysical, pigment and protein measurements exposed a large, coordinated NPQ response that tracked exhaustion of internal P stores and stabilized at a higher level (3.4 ± 0.4 fold, $P < 0.001$) during P_{limited} than P_{replete} and P_{refed} (Fig. 4a and Supplementary Fig. 12). VAZ pigments increased 2-fold relative to chlorophyll *a* during P_{limited} (Fig. 4b, $P < 0.01$), but protein abundances of VAZ biosynthetic enzymes did not change, probably reflecting post-translational control⁴¹. Remarkably, just 4 of the 38 LHC-related proteins in *Micromonas* (all 38 were detected), responded to long term P limitation, three exhibiting higher (FAS-ELIP, LIL6-like, LHCSR)

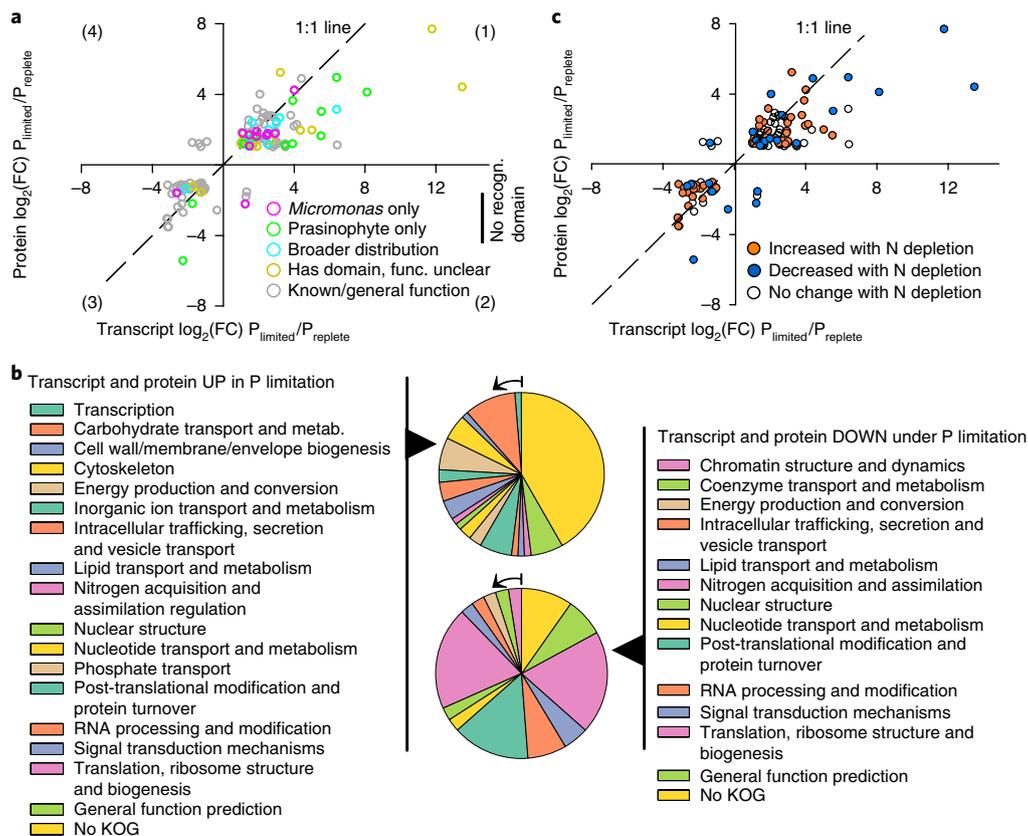


Fig. 3 | Relationships between transcript fold change and protein fold change for differentially expressed transcripts encoding proteins that changed significantly between P_{limited} and P_{replete} . **a**, The relationship was strong ($r^2=0.69$) for those that changed between P_{limited} and P_{replete} ($\log_2(\text{fold change}) > 1$, $Q < 0.05$). P_{limited} transcriptomes in biological triplicate and proteomes (as in Fig. 2) were analysed using statistics implemented in Cuffdiff. The strength of the correlation between proteins and transcripts that increased (top right, quadrant 1) versus decreased (quadrant 3) under P limitation suggests the latter are more influenced by half-life differences of the mRNA versus protein than direct modulation. Circle colours indicate taxonomic affiliations of proteins. **b**, Functional annotations of proteins for which the transcript and protein increased (left) or decreased (right) under P limitation (for the protein subset with $\log_2(\text{fold change}) > 1$, $Q < 0.05$). Arrows on pie charts indicate the starting colour category corresponding to the coloured legend lists. **c**, A separate, simplified N-supply experiment rendered the same *M. commoda* growth rates in N-replete controls ($0.8 \pm 0.1 \text{ d}^{-1}$; Supplementary Fig. 9) as P_{replete} (Fig. 1d), and reduced growth 24 hours after transfer into N-lacking medium ($0.5 \pm 0.1 \text{ d}^{-1}$, $P < 0.01$). Here, the same data as in **a** is shown with coloured fills indicating directionality of differentially-expressed transcripts ($\log_2(\text{fold change}) > 1$, $P < 0.05$) under N depletion relative to N replete. Collectively, these analyses led to identification of the Mipt1 phosphate transporter characterized herein (Supplementary Fig. 10) and other concordant transcript-protein pairs. In addition to those that changed in the same or opposite direction under the N- and P-nutrient changes, 15 unknown-function proteins (12 lacking recognizable domains) and 18 other proteins in the P-limitation concordant transcript-protein set were unchanged between N-replete and N-deplete conditions and serve as candidate P-limitation indicators (Supplementary Data 3b).

and one lower (LIL3-like) relative protein abundances (Fig. 4c and Supplementary Fig. 13). None of these four are major LHC proteins of the antenna complex and only LHCSR^{39,40} has been functionally characterized (Supplementary Fig. 13 and Fig. 4c). Importantly, modifications in LHCSR protein abundances tracked phosphate availability closely, demonstrating a sustained LHCSR response throughout phosphate limitation.

Our results highlight LHCSR-driven NPQ as a long-term strategy in *Micromonas* for exponential growth during P limitation, enabling a plastic response capacity. *Micromonas* circumvents the ensemble of varied changes in major LHC and photosystem proteins (or their transcripts) observed in global expression studies of nutrient-deprived marine algae^{18,21,33,35,36} and *Chlamydomonas*^{19,43}, some of which exhibit concurrent LHCSR activation. Unlike the non-motile, generally larger, marine algae studied to date, *Micromonas* is a fast swimmer and very small⁷; smaller phytoplankton are better optimized for uptake of low-level nutrients⁵, and under patchy resource fields, motile bacteria are known to outcompete non-motile taxa⁴⁴. These characteristics and associated adaptations may

therefore underpin its differentiated response to varying nutrients. An alternative hypothesis, testable using approaches enabled by continuous-flow photo-bioreactors, is that our results reveal a universal strategy for acclimated phosphate-limited growth, spanning the divergent marine algal groups with LHCSR homologues, that remains to be documented.

Adaptive differences among diverse unicellular algae underlie seasonal dynamics and will shape the emergence of new successional patterns, or dominant species, in future oceans. The strategies discovered herein, including activation of unknown function proteins with restricted phylogenetic distributions and an ancient photoprotective protein formerly implicated in high-light stress responses, leave *Micromonas* poised for rapid pivots when encountering nutrient changes. This underscores how understanding cell biology and behaviour can help interpret climate-induced community changes, such as increasing *Micromonas* abundance alongside reductions in larger algae in the Canadian Arctic⁴⁵. Importantly, stronger stratification in warmer waters is expected to reduce nutrient inputs to the photic-zone^{3,4}, and shallowing of the surface mixed

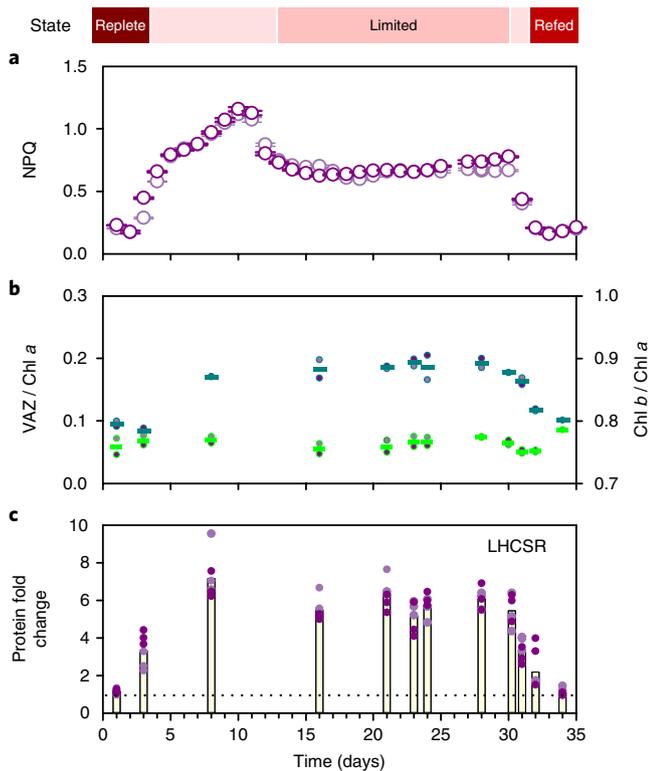


Fig. 4 | Non-photochemical quenching as a strategy for sustained growth potential during P limitation. **a**, Non-photochemical quenching (NPQ) based on FRRf measurements was significantly higher during the P_{limited}

phase than the P_{replete} and P_{refed} phases ($P < 0.001$, $n = 4$, two-sided one sample t -tests on \log_2 NPQ fold changes between major P phases). The geometric mean (symbols) and propagated error based on the standard deviation of F_m and F'_m measurements ($n = 8$) are shown for photo-bioreactors 1 (light purple) and 2 (purple). Fast dark relaxation of this P-limitation-induced NPQ characterizes it as quickly reversible, energy-dependent quenching qE (Supplementary Fig. 12). **b**, Mean violaxanthin, antheraxanthin and zeaxanthin (VAZ)-cycle pigments (dark cyan bars) and chlorophyll *b* (green bars) expressed relative to chlorophyll *a* from HPLC measurements of cells during the experiment, with 2 (P_{replete}), 10 (P_{limited}) and 2 (P_{refed}) samples analysed per primary phase (biological duplicates at each time point). Individual values from photo-bioreactors 1 (light purple) and 2 (purple) are indicated by dots (dark cyan rim, VAZ; green rim, chlorophyll *b*). P availability had a significant effect on VAZ-cycle pigments ($P < 0.01$, $n = 4$; t -tests as in **a**, see Supplementary Data 1). **c**, NPQ responses to transient light stress are mediated by the protein PsbS in moss (*Physcomitrella patens*) and vascular plants, and by a functionally similar but more ancient member of the light-harvesting complex protein family, LHCSR (LHC-stress related/LHCX) in *Chlamydomonas* and diatoms^{28,38–40} (moss has both LHCSR and PsbS; Supplementary Fig. 11). Here, LHCSR protein fold changes (relative to T_0 P_{replete}) were significantly higher from day 8 to day 31 than during P_{replete} and P_{refed} phases ($P < 0.001$, $n = 4$, two-sided one-sample t -test of \log_2 fold changes between major P phases). The mean of LHCSR fold-changes from the two photo-bioreactors (yellow bars) and distribution of technical triplicates (dots) from photo-bioreactors 1 (light purple) and 2 (purple) are shown. No change (dotted line) occurred if fold change values are one. Note that LHCSR and LIL6-like, which tracked NPQ and phosphate changes (Supplementary Fig. 13), as well as FAS-ELIP and LIL3, are not major LHCS of the antennae complex; rather, LIL6-like and LIL3 are similar to *Chlamydomonas* and plant LHC-Like (LIL) proteins activated by light-stress but of unclear function, akin to the picoprasinophyte-specific FAS-ELIP²⁸. Genes encoding these proteins had < 2 transcript fold change, suggesting post-transcriptional regulation.

layer would trap algae at higher light intensities. Thus, similarities in the biological mechanism for mitigating light stress and maximizing growth potential under nutrient limitation highlight the probability of unknown interactive effects caused by multifactorial environmental changes expected in future oceans.

Methods

Continuous flow photo-bioreactors. Photo-bioreactors were custom built at Monterey Bay Aquarium Research Institute (MBARI); the initial build has been partially described previously⁴⁶. Each photo-bioreactor has two 21 cylindrical glass culture vessels with specialized flat glass-plate bottoms that ensure uniform light delivery. Each vessel is encased in a black anodized two-part aluminium clamp covered inside with thermally conductive material (Supplementary Fig. 1), equipped with six optical sensor ports, two vertical viewports for visual monitoring, and bolted to an in-house designed Thermo-Electric temperature control unit with a heat sink and fan for heat dissipation, and two temperature sensors (one attached to the clamp and the other coupled to the wall of the vessel) for accurate control of medium temperature. The white PTFE autoclavable vessel lid contains conical ports for oxygen and pH electrodes as well as five threaded inlet and outlet ports for air, medium and discrete sampling. The vessel and clamp are positioned over a sealed 'lower unit' containing a shaded pole motor with cooling propeller and round two-pole magnet to run a magnetically coupled stirrer bar positioned inside of the vessel, and a cluster of eight warm white LEDs (Lumileds LXHL-NWG8) projecting light into each vessel with full internal reflection for evenness of light distribution. The thermo-electric unit with the two-part aluminium clamp and the lower unit permit accurate manipulation of light level as needed ($\pm 2\%$ column-to-column consistency, $\pm 1\%$ repeatability), temperature ($\pm 0.05^\circ\text{C}$) and culture mixing. Two vessels and associated assemblies are positioned within the custom designed photo-bioreactor housing and share an industrial controller, interfaced with a computer running a Labview program, for advance programming of light, temperature and mixing regimes down to 1 second temporal resolution, as well as gathering, displaying and saving data from sensors. In addition to sensors inside the vessels, each has a fast repetition rate fluorometer (FRRf) that uses remote sensors mounted in optical ports (Fig. 1b and Supplementary Fig. 1).

Micromonas culturing and sample collection. All pre-experiment culturing and experiments used sterile polycarbonate flasks or photo-bioreactor vessels. Axenic *M. commoda* RCC299 was grown under a 14:10 h light:dark cycle at 21°C . This species is a member of the *Micromonas* A/B/C lineage, which harbours clades A (represented by *M. commoda*), B and C, three clades that are more closely related to each other than to other *Micromonas* clades/species¹³, such as *M. pusilla* (Fig. 1a, a clade D member). Cultures were monitored daily by flow cytometry using an Accuri C6 (BD biosciences, San Jose, CA) before and during experiments. Axenicity was verified by DAPI staining⁴⁷ followed by visual inspection using epifluorescence microscopy and by inoculation into an organic rich test medium with incubation for up to one week in the dark. Pre-experiment cultures were acclimated and maintained in mid-exponential growth for 15 (phosphate-limitation experiment) and 10 (nitrogen-deprivation experiment) generations.

Cultures were harvested discretely 1 h after lights on (8:30 a.m.) at multiple points across the photo-bioreactor time course and at fewer time points for the nitrogen-deprivation experiment (only flow cytometry and RNA-seq performed on the latter). Live and fixed flow cytometry samples were harvested daily. The latter were fixed with glutaraldehyde (0.23% final concentration) for 20 min in the dark and frozen in liquid nitrogen. For protein samples 50 ml of culture were spun at 10,000g for 12 min, most supernatant was removed, and cells were spun again in a 1.5 ml tube at 8,000 g for 10 min, and the pellet then frozen on dry ice. For RNA, 90–100 ml of culture was filtered onto 47 mm Supor filters and flash frozen in liquid nitrogen (photo-bioreactor experiment) or frozen at -80°C (nitrogen-deprivation experiment). For pigment analysis, 45 ml of sample were filtered onto 0.65 μm pore-size DVPP filters. All of these samples were then stored at -80°C until further processing. The filtrate remaining from pigment sampling was stored at -20°C and then analysed for soluble reactive phosphate.

Photosynthetic properties. PS II fluorescence kinetics were measured by FRRf. The excitation protocol consisted of a series of 150 flashlets of 1 μs duration each followed by 4 μs dark intervals. The relaxation protocol consisted of a series of 90 flashlets with exponentially increasing time intervals between flashlets. Over the course of the excitation protocol fluorescence rose from a minimum value (F_0 or F_s in the absence or presence of actinic light, respectively) to the maximum value (F_m or F'_m in the absence or presence of actinic light, respectively) as the reaction centres closed during light saturation. Maximum photochemical yield of PSII was then calculated as $F_v/F_m = (F_m - F_0)/F_m$ from night-time measurements. Shown (Fig. 1c) are mean F_v/F_m values from 27 individual readings per night, performed between 30 and 60 min after lights off (note the value at day 26 is approximated, as daylight disrupted measurements). NPQ was calculated as $(F_m - F'_m)/F'_m$, using a mean of nine dark acclimated readings (F_m) performed before the lights went on (end of dark period) and a mean of nine readings in the presence of actinic

light (F'_m) performed 20–30 min after lights on. To study dark-relaxation of NPQ it was also calculated around the end of the light period using a mean of the five measurements immediately before lights went off for light-acclimated NPQ or the reading 5 min after lights went off for the slowly relaxing component of NPQ. Both of these were calculated relative to the dark acclimated maximum F_m reached about 45 min after lights went off.

Nutrient and pigment measurements. Soluble reactive phosphorus (SRP) was measured at discrete time points during the photo-bioreactor experiment and in the medium and seawater used before inoculation according to the methods described previously⁴⁸. Briefly, after performing a turbidity blank measurement, and four to six measurements with the vitamin addition (reagent blank), samples were measured in technical duplicate in the presence of the ammonium molybdate solution⁴⁸. SRP was measured in this manner on 13 different days in each bioreactor. Nitrogen concentrations were quantified in the seawater used for preparation of media using the methods described previously⁴⁹.

Concentrations of chlorophylls *a* and *b*, violaxanthin, antheraxanthin and zeaxanthin were measured on 12 independent days during the experimental period by HPLC analysis performed at the Horn Point Analytical Services Laboratory at the University of Maryland Center for Environmental Science (<http://www.umces.edu/hpl/analytical-services>). Acetone extraction and HPLC methodology was as described previously⁵⁰.

Details of limitation experiments. The phosphate limitation experiment was performed in duplicate columns (vessels) in the custom-built photo-bioreactors. Pre-culturing and the experiment were performed at $150 \mu\text{E m}^{-2} \text{sec}^{-1}$ photosynthetically active radiation (PAR). Autoclaved L1 medium (-Si)⁵¹ was made with $0.2 \mu\text{m}$ pore-size filtered natural seawater ($1.48 \pm 0.33 \mu\text{M PO}_4^{3-}$, $8.12 \pm 0.18 \mu\text{M NO}_3^{-}$, $n=3$) as a base. Replete phosphate was supplied at the standard concentration in L1 medium ($36 \mu\text{M}$) with a measured final concentration of $37.66 \mu\text{M}$ (before inoculation). As detailed below, the bioreactors were operated in a way that incorporates aspects of turbidostats and chemostats^{52,53}, allowing us to capture algal responses to changing environmental factors (phosphate in this case). Thus, the concentration of amended phosphate in the medium (in the sterile supply vessel) was changed three times and the pump rate was adjusted to prevent washout of cells (or overgrowth) associated with slowed (or faster) growth related to these changes.

The photo-bioreactors were inoculated with stable mid-exponential growth cells that had been moved from replete phosphate into low phosphate medium by centrifuging the replete cultures, largely removing the supernatant, resuspending the cells in P-limited medium ($1 \mu\text{M PO}_4^{3-}$) and then adding them to the photo-bioreactors (final volume 1.8 l per column). Real-time temperature, photosynthetic parameters (measured every 65 s), pH and dissolved oxygen were monitored in situ and actual growth rates were measured by drawing a discrete sample for flow cytometric cell enumeration each morning ~1 h after the onset of light as described below. While the photo-bioreactor dilution rate sets the nominal growth rate during steady state conditions, this approach is not valid for transitional periods. Hence, the reported growth rates (μ) were computed using flow cytometric cell counts and the formula:

$$\mu = D_1 + \ln(N_2/N_1)/(t_2 - t_1)$$

where D_1 is dilution rate at time t_1 , and N_1 and N_2 are cell number at time 1 (t_1) and time 2 (t_2), respectively. Changes in growth rate and F_v/F_m were not observed between T_0 and T_2 (48 h). On day 3, F_v/F_m and growth rates began to decline in both photo-bioreactor columns and on day 7 the pump rate was adjusted from 0.5 d^{-1} to 0.18 d^{-1} to avoid cell washout. On day 8, concurrent with SRP dropping below the detection limit (10 nM), growth became unstable with an average of $0.11 \pm 0.13 \text{ d}^{-1}$ from day 8 to 10. On day 10 the medium was amended with an additional $2 \mu\text{M PO}_4$ for a $3 \mu\text{M}$ supply concentration. Days 3 to 12 were considered transitional. Exponential growth was evident by day 12, and the pump rate was adjusted to 0.26 d^{-1} on day 13 to avoid overgrowth resulting from the additional phosphate. Growth rate was stable for the 17-day period from T_{13} to T_{30} , which was defined as the period of acclimated P-limited growth (P_{limited}). Refeeding was accomplished by switching the supply carboys from medium concentrations of $3 \mu\text{M}$ to $37 \mu\text{M}$ phosphate on day 30. Peak F_v/F_m and growth increased by nightfall and after 24 hours 400 ml of 1.8 l photo-bioreactor volume had been replaced so that the total P supply pool should have increased by $2.1 \mu\text{M}$ (6 h), $7.6 \mu\text{M}$ (24 h) and $13.4 \mu\text{M}$ (48 h) after the supply level change on day 30, in addition to the $3 \mu\text{M}$ that was in the system (largely in biological material) during P_{limited} . Measured phosphate concentrations after resupplying P-replete medium corresponded with estimates computed by accounting for the photo-bioreactor dilution rate, and theoretical displacement of the P-limited media in the vessels, which should have been complete on T_{33} (72 h after the supply switch). Days 30–31 were considered transitional and the pump dilution rate was turned up 2 days after refeeding due to increasing cell abundances. The P_{replete} phase (defined as starting at day 32) was set to a lower growth rate than under the initial replete conditions (Fig. 1b) by the photo-bioreactor pump dilution rate (0.5 d^{-1}). Cell abundances were $1.31 (\pm 0.04) \times 10^7$, $0.93 (\pm 0.07) \times 10^7$ and $1.38 (\pm 0.33) \times 10^7$ cells ml^{-1} for P_{replete} , P_{limited} and P_{replete} , respectively.

Two-way ANOVA with the all pairwise multiple comparison procedures (Holm-Sidak method) was used to compare the changes in growth rate between columns and/or changes between the stable phases (Fig. 1d). Using measurements from 7, 14 and 3 days, from within each phase (P_{replete} , P_{limited} and P_{replete} , respectively) and from each column, resulting in a total of 14, 28 and 6 measurements, respectively) the difference in means among different levels of phosphate was statistically significant ($P < 0.001$, $n=2$) after allowing for effects of differences in column, whereas differences between columns were not significant ($P=0.726$). For additional statistical comparisons, the initial removal of phosphate in the medium and hence transition to P limitation and the resupply of P that initiated recovery were treated as two independent alterations of the P supply. The effect of P supply on growth rates between the stable phases was compared using a two-tailed *t*-test ($n=4$). To test for responses in F_v/F_m , NPQ and VAZ-pigments to P availability, response ratios between the stable phases (as defined above) before and after the transitions in P supply were calculated (again $n=4$). A one-sample *t*-test was used on log₂-transformed response ratios to test for a significant effect of changes in P supply ($n=4$).

The nitrogen depletion experiment was performed under $180 \mu\text{E m}^{-2} \text{sec}^{-1}$ PAR. After centrifugation of pre-experiment cultures from replete conditions, cells were re-suspended in replete K^{54} medium (control) or K medium in artificial seawater with the nitrogen amendment omitted (N-deplete treatment). Controls and treatments were performed in biological triplicates in a volume of 800 ml per replicate and growth rates were compared using a *t*-test ($n=3$; Supplementary Fig. 9).

Protein extraction and digestion. Protein was extracted in a randomized order from samples collected at T_0 (3 hours before inoculation into vessels, 1 hour after lights on) and day 1 (T_1), representing the P_{replete} phase, days 3 (T_3) and 8 (T_8) representing the transition to P depletion, days 16 (T_{16}), 21 (T_{21}), 23 (T_{23}), 24 (T_{24}) and 28 (T_{28}) representing acclimated P-limited growth, 6 h after resupply (T_{30} , 6 h), and 1 day (T_{31}), 2 days (T_{32}) and 4 days (T_{34}) after resupply, with the latter two termed P_{replete} . The 26 samples ($n=2$ per time point, that is, biological duplicates, two to five time points per primary phase) were washed and suspended in 100 mM NH_4HCO_3 (pH 8.4) containing 2% SDS and 1% DTT, and lysed via Pressure Cycling Technology as previously described⁴⁵. Lysates were centrifuged at $14,000g$, 4°C , for 30 min, and proteins within the supernatant were then precipitated and washed using a chloroform/methanol procedure⁵⁶. The pellet was suspended in a solution of 8 M urea and 50 mM NH_4HCO_3 (pH 8.4) and a bicinchoninic acid assay (BCA; Pierce Chemical Co.) was performed ($10\times$ and $20\times$ dilution) to determine protein concentrations. Proteins were then reduced using fresh dithiothreitol to reach a final concentration of 5 mM; the solution was incubated at 60°C for 30 min. Following incubation, a volume of 0.4 M Iodoacetamide was added to each fraction (alkylation step) to obtain a final concentration of 40 mM then incubated in the dark for 1 h at 37°C . Proteins were digested using sequencing-grade trypsin (Roche) at a unit-to-protein ratio of 1:50. The resulting peptides were desalted using a strong cation-exchange, C-18 SPE column (Supelco) following established protocols⁵⁷. Peptide concentrations in each of the extracts were measured using BCA assays.

LC-MS data generation. The measurement of peptides by mass spectrometry (MS) used protocols that have been applied to *M. pusilla*⁶⁰. Briefly, each peptide sample was analysed in technical (instrument) triplicate, under a randomized run order, using an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) coupled to a reversed-phase HPLC dual column system using in-house manufactured columns ($60 \text{ cm} \times 360 \mu\text{m}$ outer diameter (o.d.) $75 \mu\text{m}$ inner diameter (i.d.) fused silica capillary tubing) packed with 3 μm Jupiter C18 stationary phase (Phenomenex). A detailed description of HPLC operating conditions including mobile phase composition is described elsewhere⁵⁸. In brief, $5 \mu\text{l}$ of peptides ($0.1 \mu\text{g} \mu\text{l}^{-1}$) were injected into the HPLC and subjected to an exponential gradient. A 40 cm length of $360 \mu\text{m}$ o.d. \times $15 \mu\text{m}$ i.d. fused silica tubing was used to split $\sim 20 \mu\text{l} \text{min}^{-1}$ of flow before it reached the injection valve (sample loop), which controlled the gradient speed under conditions of constant flow ($\sim 300 \text{ nl} \text{min}^{-1}$).

Separated peptides were ionized (positive) using an electrospray ionization interface (manufactured in-house; no sheath gas or make-up liquid was used) that consisted of chemically etched electrospray emitters⁵⁹ ($150 \mu\text{m}$ o.d., $20 \mu\text{m}$ i.d.). The mass spectrometer was operated using a heated capillary temperature and spray voltage of 325°C and 2.2 kV, respectively. Data were acquired for 100 min beginning ~ 60 min after sample injection. High-resolution parent ion mass spectra (Automatic Gain Control (AGC) 1×10^6) were collected from 400–2,000 m/z at a resolution of 60,000 followed by generation of data dependent CID MS/MS high-resolution spectra, centroid mode, at a resolution of 7500 (collision energy 35%, activation time 30 ms (AGC 5×10^4) of the 10 most abundant ions using an isolation width of 3 Da). Charge state screening was not enabled. A dynamic exclusion time of 60 sec was used to discriminate against previously analysed ions.

Peptide/protein identification and analysis. LC-MS spectra were de-isotoped using DeconTools⁶⁰, after which mass and normalized elution time features (LC-MS feature) were identified using VIPER⁶¹. A total of 2,689,319 LC-MS features were identified (an average of 17,129 per dataset) and matched to an

empirically derived accurate mass and time (AMT) tag database according to published methods^{30,57,62}. This database was used as a peptide sequence library for matching LC-MS features to peptide sequences. Matched LC-MS features were then filtered on a false discovery rate (FDR) of $\leq 5\%$ resulting in a total of 1,081,094 (an average of 12,012 per dataset) non-redundant matches. The FDR associated with the AMT tag proteomics approach was calculated using a statistical algorithm for assigning confidence to matched mass and elution time features⁶³. The total number of proteins identified was 4,954 (that is, at least one peptide match) when data from all time points were considered. Only the 2,118 proteins with at least two peptide matches were analysed further. Instrument data files used to build the AMT tag database, plus the data files used by VIPER, are available at ProteomeXchange (under PXD006320).

Relative peptide abundances were calculated from the integration of ion intensities (ion current) measured across instrument scans, then \log_2 -transformed and technical replicates scaled and normalized using the linear regression algorithm in DAnTE⁶⁴. The linear regression algorithm corrects for systematic bias that is linearly dependent on the magnitude of peptide abundance as can occur with, for example, sample carry-over on the LC column⁶⁵. The mean Pearson correlation coefficient between peptide abundances measured for technical replicates was calculated as 0.93 ± 0.03 . Protein abundance from the observed peptides was estimated using the DAnTE Rollup function. Default parameters were utilized for this function including the requirement that a peptide be measured in at least 50% of all generated datasets, and a minimum of five peptides per protein was required for the Grubb's test (to determine peptide outliers), performed with a *P* value cutoff of 0.05.

Protein abundance values for each time point were compared using a two-tailed *t*-test. Additionally, samples from each of the three physiological states (P_{replete} , P_{limited} or P_{refed}) were compared to the other primary states using two-tailed *t*-tests. Differentially expressed proteins were determined by the *t*-test and retained for biological interpretation if exhibiting a fold change ≥ 2 and *Q* value cutoff < 0.05 in at least one treatment relative to the other two. Fold changes of proteins between different phases were computed from average protein abundances from biological duplicates (technical triplicate) over sampled time points. Fold changes at each time point were also calculated relative to T_0 (P_{replete}) and compared using a two-tailed *t*-test. MultiExperiment Viewer was used (MeV4.1; <http://www.tm4.org/mev.html>) to perform a *z*-score transformation across samples. The \log_2 protein abundances used to calculate *z*-scores for each sample were the mean of instrument triplicates ($n = 3$). Hierarchical clustering was performed on data from the three primary phases using the Pearson coefficient to construct a correlation matrix and average linkage as the linkage method.

RNA-sequencing and analysis. RNA was extracted using the ToTALLY RNA kit (Ambion) with the addition of a mechanical step involving glass beads⁶⁶. PolyA-selected RNA was isolated from 5 μg total RNA by two rounds of isolation, with purification and reverse transcription of fragmented RNA (200 to 300 nt) and further purification and amplification of cDNA fragments as described previously⁶⁶. Protocols for strand-specific, paired-end sequencing (150 bp each read) were performed on the Illumina HiSeq platform. Reads were aligned to the genome of *M. commoda* (RCC299 genome build v3)⁷ using TopHat version 1.4.0⁶⁷. Relative transcript abundances for each protein-encoding gene were estimated using Cufflinks version 2.2.1⁶⁸ and calculated as FPKM (fragments per kilobase of exon per million reads mapped). Genes with FPKM ≥ 2 in one or more sample were further analysed. When all data was taken together, 9,727 of 10,306 predicted protein encoding genes met this cutoff criterion. Cuffdiff and statistical methods therein were used to perform differential gene expression analysis⁶⁸; 14% of protein encoding genes increased and 11% decreased in relative transcript abundance with *P* limitation (\log_2 fold change > 1 , $Q < 0.05$; Supplementary Fig. 8). For the nitrogen deprivation experiment D₁ nutrient-replete mid-exponential growth cultures (as a control) and D₁ N-deplete cultures were sequenced in biological triplicate (Supplementary Fig. 9). Reads sequenced per sample ranged from 5.8×10^7 to 6.9×10^7 , with a mean and standard deviation of $6.2 \pm 0.4 \times 10^7$ reads per sample.

Protein taxonomic distributions. Proteins that changed ≥ 2 -fold under P_{limited} growth were analysed based on the functional annotation of the *M. commoda* genome to classify functions into families or detect any recognizable domains and important sites. Additional functional annotation of proteins was performed using NCBI's Conserved Domain Database (CDD) with a default setting to detect any conserved domain footprints that were not predicted by the JGI Annotation Pipeline. Proteins with no characterized domains were further analysed for organismal distributions using BLAST (Basic Local Alignment Search Tool) standalone version 2.6.0+. BLASTP was used to search against available phytoplankton and plant genomes and GenBank's non-redundant database. Two custom BLAST databases were built for the protein distribution analysis. The first included predicted proteomes from 56 prasinophyte transcriptomes available from the Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP; Supplementary Data 2)²⁷. The other database used publically available predicted proteomes from 16 eukaryotic phytoplankton genome projects (*M. pusilla* CCMP1545, *Ostreococcus lucimarinus*, *Ostreococcus* sp. RCC809, *Ostreococcus tauri*, *Bathycoccus prasinos* RCC1105, *Coccomyxa* sp. C-169,

C. reinhardtii, *Chlorella* sp. NC64A, *Volvox carterii nagariensis*, *Bigelowiella natans* CCMP2755, *Aureococcus anophagefferens* CCMP1984, *Thalassiosira pseudonana* CCMP1335, *Phaeodactylum tricornutum* CCAP1055/1, *Fragilariopsis cylindrus* CCMP1102, *Guillardia theta* CCMP2712, *Emiliania huxleyi*) and two plant genomes (*Physcomitrella patens* ssp. *patens* and *Arabidopsis thaliana*) all from JGI except *A. thaliana* from TAIR (The *Arabidopsis* Information Resource). If a protein was not found in MMETSP peptide files, then the corresponding transcriptome assembly was analysed using the TBLASTN algorithm; this approach detected proteins that had been missed during the MMETSP protein prediction/filtering steps. A low stringency criterion was used to conservatively assess the 'unique' set: *e*-value $\leq 10^{-5}$ and % identity over 50%, 30% and 20% (Fig. 2b and Supplementary Data 2c) of probable orthologs. For searching the nr database in GenBank, only the criteria of *e*-value $\leq 10^{-5}$ was used. These proteins were also analysed using Phyre2⁶⁹. However, Phyre2 predictions did not meet significance levels deemed robust and therefore the results were inconclusive for the proteins with phylogenetic distribution restricted to prasinophytes.

***S. cerevisiae* Δ pho84 complementation and metal toxicity assays.** The open reading frame of wlab.237202 was codon-optimized for expression in *Saccharomyces cerevisiae*, synthesized with a C-terminal His₆ tag (Synbio Technologies) and cloned into the BamHI/HindIII sites of p416 GPD (ATCC) to generate pDM008. *S. cerevisiae* strain BY2959⁷⁰ (*pho84::HIS3*) was then transformed with either p416 GPD or pDM008, while BY5676 was transformed with p416 GPD and used as a positive control. Cells were grown at 30 °C, 200 rpm shaking for 16 h in SC-Ura medium (0.79% yeast nitrogen base without amino acids (Formedium), 700 mg l⁻¹ complete supplement mix-ura (Formedium), 2% (w/v) glucose), centrifuged at 1,810g for 3 min, and washed twice with low-phosphate (LP) medium⁷¹, pH 4.5. Cells were re-suspended in LPi medium and diluted to an OD₆₀₀ of 0.1 before 100 μl was inoculated into a 96-well plate. Cells were grown at 30 °C in a Fluostar Omega Lite microplate reader (BMG Labtech Ltd.), with OD₅₉₅ assessed at 10-minute intervals and continuous double-orbital shaking (200 rpm) between reads. BY2959 strains were assessed in biological triplicate, alongside a single positive control strain (BY5676 p416 GPD); each strain consisted of eight technical replicates. For metal toxicity assays on *S. cerevisiae* strains BY5676 p416 GPD, BY2959 p416 GPD and BY2959 pDM008 were grown in Sc-ura medium and prepared as above, before being diluted to an OD₆₀₀ of 1.0 and a dilution series prepared. Toxicity resulting from transport of copper, cobalt, zinc, manganese, arsenate and chromate were tested in duplicate at each of the indicated concentrations (Supplementary Fig. 10B). Ten μl of each dilution was spotted onto Sc-ura plates containing each metal assayed. Plates were incubated for 72 h at 30 °C before being imaged on a G:BOX imaging system (Syngene) to identify compounds on which toxicity was restored by the addition of pDM008.

Phylogenetic reconstruction of LHCSR proteins. Homologues to the LHCSR protein from *M. commoda* were retrieved using a BLASTp search against the nr-DB from NCBI using the LHCSR2²⁸ sequence as a seed, and all protein sequences with an *e*-value $\leq 10^{-15}$ were kept. Function in NPQ has been established for the LHCSR of green algae³⁹ (in *Ostreococcus* and *Chlamydomonas*) and for the homologous diatom protein LHCX⁷². For identical sequences (100% amino acid identity) coming from the same species, only one representative sequence was retained. The protein sequence from *Chromera velia* (accession AIN76774) was too short (117 aa) and therefore discarded. Sequences were aligned using MAFFT⁷³ with default parameters and the resulting multiple alignment were masked by discarding all the positions allowing 5% of gapped positions. LHCZ protein sequences were used as outgroup as they were shown to be the most closely related LHC protein sequences⁷⁴. Phylogenetic reconstructions were conducted using RAXML⁷⁵ with WAG matrix, and 1,000 bootstrap replicates.

Reporting Summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Data supporting the findings of this study that are not included in the manuscript are included in the Supplementary Information. Protein data are available at ProteomeXchange under accession PXD006320 and RNA-seq data is available from the NCBI short read archive (SRA) under Bioproject no. PRJNA445307. Gene models used are available in a downloadable format at <https://genome.jgi.doe.gov/MicpuN3v2/MicpuN3v2.home.html>. The *M. commoda* strain used is available on request from the corresponding author and has been deposited under CCMP2709 at the National Center for Marine Algae and Microbiota.

Received: 9 December 2017; Accepted: 16 May 2018;
Published online: 25 June 2018

References

- Field, C. B., Behrenfeld, M. J., Randerson, J. T. & Falkowski, P. Primary production of the biosphere: integrating terrestrial and oceanic components. *Science* **281**, 237–240 (1998).

2. Wu, J., Sunda, W., Boyle, E. A. & Karl, D. M. Phosphate depletion in the western North Atlantic. *Ocean. Sci.* **289**, 759–762 (2000).
3. Behrenfeld, M. J. et al. Climate-driven trends in contemporary ocean productivity. *Nature* **444**, 752–755 (2006).
4. Flombaum, P. et al. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proc. Natl Acad. Sci. USA* **110**, 9824–9829 (2013).
5. Worden, A. Z. et al. Environmental science. Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. *Science* **347**, 1257594 (2015).
6. Sebastian, M. et al. Lipid remodelling is a widespread strategy in marine heterotrophic bacteria upon phosphorus deficiency. *ISME J.* **10**, 968–978 (2016).
7. Worden, A. Z. et al. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science* **324**, 268–272 (2009).
8. Monier, A., Worden, A. Z. & Richards, T. A. Phylogenetic diversity and biogeography of the Mamiellophyceae lineage of eukaryotic phytoplankton across the oceans. *Environ. Microbiol. Rep.* **8**, 461–469 (2016).
9. Clayton, S., Lin, Y.-C., Follows, M. J. & Worden, A. Z. Co-existence of distinct *Ostreococcus* ecotypes at an oceanic front. *Limnol. Oceanogr.* **62**, 75–88 (2017).
10. Limardo, A. J. et al. Quantitative biogeography of picoprasinophytes establishes ecotype distributions and significant contributions to marine phytoplankton. *Environ. Microbiol.* **19**, 3219–3234 (2017).
11. Countway, P. D. & Caron, D. A. Abundance and distribution of *Ostreococcus* sp. in the San Pedro Channel, California, as revealed by quantitative PCR. *Appl. Environ. Microbiol.* **72**, 2496–2506 (2006).
12. Worden, A. Z., Nolan, J. K. & Palenik, B. Assessing the dynamics and ecology of marine picophytoplankton: the importance of the eukaryotic component. *Limnol. Oceanogr.* **49**, 168–179 (2004).
13. Simmons, M. P. et al. Intron invasions trace algal speciation and reveal nearly identical Arctic and Antarctic *Micromonas* populations. *Mol. Biol. Evol.* **32**, 2219–2235 (2015).
14. Not, F. et al. Late summer community composition and abundance of photosynthetic picoeukaryotes in Norwegian and Barents Seas. *Limnol. Oceanogr.* **50**, 1677–1686 (2005).
15. Cuvelier, M. L. et al. Targeted metagenomics and ecology of globally important uncultured eukaryotic phytoplankton. *Proc. Natl Acad. Sci. USA* **107**, 14679–14684 (2010).
16. Treusch, A. H. et al. Phytoplankton distribution patterns in the northwestern Sargasso Sea revealed by small subunit rRNA genes from plastids. *ISME J.* **6**, 481–492 (2012).
17. Lomas, M. W. et al. Two decades and counting: 24-years of sustained open ocean biogeochemical measurements in the Sargasso Sea. *Deep Sea Res. Part II Top. Stud. Oceanogr.* **93**, 16–32 (2013).
18. Wurch, L. L., Bertrand, E. M., Saito, M. A., Van Mooy, B. A. & Dyhrman, S. T. Proteome changes driven by phosphorus deficiency and recovery in the brown tide-forming alga *Aureococcus anophagefferens*. *PLoS ONE* **6**, e28949 (2011).
19. Hsieh, S. I. et al. The proteome of copper, iron, zinc, and manganese micronutrient deficiency in *Chlamydomonas reinhardtii*. *Mol. Cell Proteom.* **12**, 65–86 (2013).
20. Dyhrman, S. T. et al. The transcriptome and proteome of the diatom *Thalassiosira pseudonana* reveal a diverse phosphorus stress response. *PLoS ONE* **7**, e33768 (2012).
21. Dyhrman, S. T. et al. Long serial analysis of gene expression for gene discovery and transcriptome profiling in the widespread marine coccolithophore *Emiliania huxleyi*. *Appl. Environ. Microbiol.* **72**, 252–260 (2006).
22. Himeoka, Y. & Kaneko, K. Theory for transitions between exponential and stationary phases: universal laws for lag time. *Phys. Rev. X* **7**, 021049 (2017).
23. Yang, Z. K. et al. Systems-level analysis of the metabolic responses of the diatom *Phaeodactylum tricornerutum* to phosphorus stress. *Environ. Microbiol.* **16**, 1793–1807 (2014).
24. Thompson, A. W., Huang, K., Saito, M. A. & Chisholm, S. W. Transcriptome response of high- and low-light-adapted *Prochlorococcus* strains to changing iron availability. *ISME J.* **5**, 1580–1594 (2011).
25. Grossman, A. R. & Aksoy, M. in *Annual Plant Reviews Vol. 48: Phosphorus Metabolism in Plants* (eds. Plaxton, W. C. & Lambers, H.) Ch. 12 (Wiley, Hoboken, 2015).
26. Plaxton, W. C. & Tran, H. T. Metabolic adaptations of phosphate-starved plants. *Plant Physiol.* **156**, 1006–1015 (2011).
27. Keeling, P. J. et al. The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol.* **12**, e1001889 (2014).
28. Cuvelier, M. L. et al. Responses of the picoprasinophyte *Micromonas commoda* to light and ultraviolet stress. *PLoS ONE* **12**, e0172135 (2017).
29. Lan, P., Li, W. & Schmidt, W. Complementary proteome and transcriptome profiling in phosphate-deficient *Arabidopsis* roots reveals multiple levels of gene regulation. *Mol. Cell Proteom.* **11**, 1156–1166 (2012).
30. Waltman, P. H. et al. Identifying aspects of the post-transcriptional program governing the proteome of the green alga *Micromonas pusilla*. *PLoS ONE* **11**, e0155839 (2016).
31. Harvey, R., Dezi, V., Pizzinga, M. & Willis, A. E. Post-transcriptional control of gene expression following stress: the role of RNA-binding proteins. *Biochem. Soc. Trans.* **45**, 1007–1014 (2017).
32. Monier, A. et al. Phosphate transporters in marine phytoplankton and their viruses: cross-domain commonalities in viral-host gene exchanges. *Environ. Microbiol.* **14**, 162–176 (2012).
33. Le Bihan, T. et al. Shotgun proteomic analysis of the unicellular alga *Ostreococcus tauri*. *J. Proteom.* **74**, 2060–2070 (2011).
34. Halsey, K. H., Milligan, A. J. & Behrenfeld, M. J. Contrasting strategies of photosynthetic energy utilization drive lifestyle strategies in ecologically important picoeukaryotes. *Metabolites* **4**, 260–280 (2014).
35. Bender, S. J., Durkin, C. A., Berthiaume, C. T., Morales, R. L. & Armbrust, E. V. Transcriptional responses of three model diatoms to nitrate limitation of growth. *Front. Marine Sci.* <https://doi.org/10.3389/fmars.2014.00003> (2014).
36. Lommer, M. et al. Genome and low-iron response of an oceanic diatom adapted to chronic iron limitation. *Genome Biol.* **13**, R66 (2012).
37. Parkhill, J. P., Maillet, G. & Cullen, J. J. Fluorescence-based maximum quantum yield for PSII as a diagnostic of nutrient stress. *J. Phycol.* **37**, 517–529 (2001).
38. Kromdijk, J. et al. Improving photosynthesis and crop productivity by accelerating recovery from photoprotection. *Science* **354**, 857–861 (2016).
39. Peers, G. et al. An ancient light-harvesting protein is critical for the regulation of algal photosynthesis. *Nature* **462**, 518–521 (2009).
40. Lepetit, B. et al. The diatom *Phaeodactylum tricornerutum* adjusts NPQ capacity in response to dynamic light via fine-tuned Lxhc and xanthophyll cycle pigment synthesis. *New Phytol.* **214**, 205–218 (2016).
41. Goss, R. & Lepetit, B. Biodiversity of NPQ. *J. Plant Physiol.* **172**, 13–32 (2015).
42. Maat, D. S., Crawford, K. J., Timmermans, K. R. & Brussaard, C. P. Elevated CO₂ and phosphate limitation favor *Micromonas pusilla* through stimulated growth and reduced viral impact. *Appl. Environ. Microbiol.* **80**, 3119–3127 (2014).
43. Moseley, J. L., Chang, C. W. & Grossman, A. R. Genome-based approaches to understanding phosphorus deprivation responses and PSRI control in *Chlamydomonas reinhardtii*. *Eukaryot. Cell* **5**, 26–44 (2006).
44. Smriga, S., Fernandez, V. I., Mitchell, J. G. & Stocker, R. Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria. *Proc. Natl Acad. Sci. USA* **113**, 1576–1581 (2016).
45. Li, W. K. W., McLaughlin, F. A., Lovejoy, C. & Carmack, E. C. Smallest algae thrive as the Arctic Ocean freshens. *Science* **326**, 539 (2009).
46. Wilson, S. T. et al. Hydrogen cycling by the unicellular marine diazotroph *Crocospaera watsonii* strain WH8501. *Appl. Environ. Microbiol.* **76**, 6797–6803 (2010).
47. Porter, K. & Feig, Y. S. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* **25**, 943–948 (1980).
48. Murphy, J. & Riley, J. P. A modified single solution method for the determination of phosphate in natural waters. *Anal. Chim. Acta* **27**, 31–36 (1962).
49. Pennington, J. T. & Chavez, F. P. Seasonal fluctuations of temperature, salinity, nitrate, chlorophyll and primary production at station H3/M1 over 1989–1996 in Monterey Bay, California. *Deep Sea Res. Part II* **47**, 947–974 (2000).
50. Van Heukelem, L. & Thomas, C. S. Computer-assisted high-performance liquid chromatography method development with applications to the isolation and analysis of phytoplankton pigments. *J. Chromatogr. A* **910**, 31–49 (2001).
51. Anderson, R. *Algal Culturing Techniques* (Elsevier, New York, 2005).
52. McCarthy, J. J. in *Physiological Bases of Phytoplankton Ecology* (ed. Platt, T.) 211–233 (Canadian Bulletin of Fisheries and Aquatic Sciences, 1981).
53. McCarthy, J. J. & Goldman, J. C. Nitrogenous nutrition of marine phytoplankton in nutrient-depleted waters. *Science* **203**, 670–672 (1979).
54. Keller, M. D., Selvin, R. C., Claus, W. & Guillard, R. R. L. Media for the culturing of oceanic ultraplankton. *J. Phycol.* **23**, 633 (1987).
55. Callister, S. J. et al. Analysis of biostimulated microbial communities from two field experiments reveals temporal and spatial differences in proteome profiles. *Environ. Sci. Technol.* **44**, 8897–8903 (2010).
56. Fic, E., Kedracka-Krok, S., Jankowska, U., Pirog, A. & Dziedzicka-Wasylewska, M. Comparison of protein precipitation methods for various rat brain structures prior to proteomic analysis. *Electrophoresis* **31**, 3573–3579 (2010).
57. Lipton, M. S. et al. Global analysis of the *Deinococcus radiodurans* proteome by using accurate mass tags. *Proc. Natl Acad. Sci. USA* **99**, 11049–11054 (2002).

58. Robidart, J. et al. Characterizing microbial community and geochemical dynamics at hydrothermal vents using osmotically driven continuous fluid samplers. *Environ. Sci. Technol.* **47**, 4399–4407 (2013).
59. Kelly, R. T. et al. Chemically etched open tubular and monolithic emitters for nano-electrospray ionization mass spectrometry. *Anal. Chem.* **78**, 7796–7801 (2006).
60. Crowell, K. L. et al. LC-IMS-MS feature finder: detecting multidimensional liquid chromatography, ion mobility and mass spectrometry features in complex datasets. *Bioinformatics* **29**, 2804–2805 (2013).
61. Monroe, M. E. et al. VIPER: an advanced software package to support high-throughput LC-MS peptide identification. *Bioinformatics* **23**, 2021–2023 (2007).
62. van Baren, M. J. et al. Evidence-based green algal genomics reveals marine diversity and ancestral characteristics of land plants. *BMC Genomics* **17**, 267 (2016).
63. Stanley, J. R. et al. A statistical method for assessing peptide identification confidence in accurate mass and time tag proteomics. *Anal. Chem.* **83**, 6135–6140 (2011).
64. Polpitiya, A. D. et al. DANTE: a statistical tool for quantitative analysis of -omics data. *Bioinformatics* **24**, 1556–1558 (2008).
65. Callister, S. J. et al. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* **5**, 277–286 (2006).
66. Duanmu, D. et al. Marine algae and land plants share conserved phytochrome signaling systems. *Proc. Natl Acad. Sci. USA* **111**, 15827–15832 (2014).
67. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
68. Trapnell, C. et al. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat. Protoc.* **7**, 562–578 (2012).
69. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
70. Bun-Ya, M., Nishimura, M., Harashima, S. & Oshima, Y. The PH084 gene of *Saccharomyces cerevisiae* encodes an inorganic phosphate transporter. *Mol. Cell. Biol.* **11**, 3229–3238 (1991).
71. Rubin, G. M. Three forms of the 5.8-S ribosomal RNA species in *Saccharomyces cerevisiae*. *Eur. J. Biochem.* **41**, 197–202 (1974).
72. Bailleul, B. et al. An atypical member of the light-harvesting complex stress-related protein family modulates diatom responses to light. *Proc. Natl Acad. Sci. USA* **107**, 18214–18219 (2010).
73. Katoh, K., Kuma, K., Toh, H. & Miyata, T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* **33**, 511–518 (2005).
74. Neilson, J. A. D., Rangrikhphoti, P. & Durnford, D. G. Evolution and regulation of *Bigelowiella natans* light-harvesting antenna system. *J. Plant Physiol.* **217**, 68–76 (2017).
75. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
76. Burki, F., Shalchian-Tabrizi, K. & Pawlowski, J. Phylogenomics reveals a new ‘megagroup’ including most photosynthetic eukaryotes. *Biol. Lett.* **4**, 366–369 (2008).

Acknowledgements

We thank D. Au, L. Bird, Z. Kolber, R. Thompson and S. Tozzi for contributions to photo-bioreactor design, construction and refinement. We thank M. Ares, K. Halsey, K. Hoadley and J. Barry for helpful discussions and S. Sudek for data deposition. Proteomics was performed at EMSL, a PNNL facility sponsored by DOE's Office of Biological and Environmental Research. This research was supported by the Packard Foundation, Gordon and Betty Moore Foundation Award GBMF3788 (A.Z.W.), National Science Foundation NSF-IOS0843119 (A.Z.W. and U.G.) and US Department of Energy DOE-DE-SC0004765 (A.Z.W., S.J.C. and R.D.S.).

Author contributions

A.Z.W., E.N.R., J.G. and V.J. designed the experiments, E.N.R., J.G. and V.J. performed the experiments, flow cytometry and nutrient measurements with contributions from L.S. and V.A.E. S.W. performed NPQ and pigment analyses. C.K.A., S.O.P., R.D.S., S.J.C. performed proteomic analyses and J.G. and S.J.C. further analysed proteomics data with input from A.Z.W. R.D., G.K.R. and C-L.W. performed RNA-seq. R.D. and V.J. analysed RNA-seq with input from A.Z.W. and U.G., C.J.C. performed protein taxonomic distribution studies, C.B. performed phylogenetic analysis and T.A.R. and D.S.M. performed yeast complementation. D.K. was responsible for all technical aspects of the photo-bioreactors. J.G., S.W., U.G. and A.Z.W. wrote the manuscript and all authors read or edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41564-018-0178-7>.

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence and requests for materials should be addressed to A.Z.W.

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

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

▶ Experimental design

1. Sample size

Describe how sample size was determined.

Sample sizes could be interpreted in two ways for our study: number of cells sampled within a time point, determined based on the amounts needed for RNAseq and Proteomics, generally 300 million cells per sample); additionally sample size could refer to samples representing different phases, and these differ depending on the experimental procedure, e.g. for proteomics at least two different time points (each in biological duplicate) were sampled for the stable acclimated phases, for photobiology measurements were continuous (~ every minute measurements for 35 days). All the details for sample sizes are provided in the methods.

2. Data exclusions

Describe any data exclusions.

One out of the 35 days sampled in the bioreactor was excluded where there was a time change (e.g., daylight savings, unrecognized by the individual sampling) so the data was collected one hour different than for all other time points - which meant the cells were at a different cell cycle phase (as they are light synchronized cultures on a day: night cycle). Additionally this point was at the end of the experiment, not in the two major phases compared throughout the MS.

3. Replication

Describe the measures taken to verify the reproducibility of the experimental findings.

The photo-bioreactors were run in duplicate, non-bioreactor experiments were in triplicate, proteomics were in technical triplicate for biological duplicates at each time point, and multiple time points were sampled within each physiological state.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Proteomics samples were processed in a randomized fashion

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Yes the investigators were blinded to group allocation through proteomics data collection and initial analysis.

Note: all in vivo studies must report how sample size was determined and whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- Test values indicating whether an effect is present
*Provide confidence intervals or give results of significance tests (e.g. *P* values) as exact values whenever appropriate and with effect sizes noted.*
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars in all relevant figure captions (with explicit mention of central tendency and variation)

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

All are publicly available (published) or for purchase from other groups: MultiExperiment Viewer, Winlist, R, SigmaStat, SigmaPlot, TopHat suite

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a third party.

There are no unique materials used, apart from the photo-bioreactors (non-biological) and all part numbers are provided in Figure S1 for constructing these.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

not applicable

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

RCC299 available from the National Center for Marine Algae and Microbiota under CCMP2709

b. Describe the method of cell line authentication used.

Mapping of RNAseq to the RCC299 genome confirms the culture used was RCC299

c. Report whether the cell lines were tested for mycoplasma contamination.

not applicable

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

Provide a rationale for the use of commonly misidentified cell lines OR state that no commonly misidentified cell lines were used.

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide all relevant details on animals and/or animal-derived materials used in the study.

For laboratory animals, report species, strain, sex and age OR for animals observed in or captured from the field, report species, sex and age where possible OR state that no animals were used.

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

Provide all relevant information on human research participants, such as age, gender, genotypic information, past and current diagnosis and treatment categories, etc. OR state that the study did not involve human research participants.