

Proton-pumping rhodopsins in marine diatoms

Susumu Yoshizawa^{a,b,c,†}, Tomonori Azuma^d, Keiichi Kojima^{e,f}, Keisuke Inomura^g,
Masumi Hasegawa^{a,b}, Yosuke Nishimura^a, Masuzu Kikuchi^e, Gabrielle Armin^g, Hideaki
Miyashita^d, Kentaro Ifuku^h, Takashi Yamanoⁱ, Adrian Marchetti^j, Hideya Fukuzawa^l,
Yuki Sudo^{e,f} and Ryoma Kamikawa^h

^aAtmosphere and Ocean Research Institute, The University of Tokyo, Chiba 277–8564,
Japan

^bGraduate School of Frontier Sciences, The University of Tokyo, Chiba 277–8563, Japan

^cCollaborative Research Institute for Innovative Microbiology, The University of Tokyo,
Tokyo 113–8657, Japan

^dGraduate School of Human and Environmental Studies, Kyoto University, Kyoto 606-
8501, Japan

^eDivision of Pharmaceutical Sciences, Okayama University, Okayama 700-8530, Japan.

^fGraduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama
University, Okayama 700-8530, Japan.

^gGraduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA

^hGraduate School of Agriculture, Kyoto University, Kyoto 606-8502, Japan

ⁱGraduate School of Biostudies, Kyoto University, Kyoto, 606-8502, Japan

^jDepartment of Earth, Marine and Environmental Sciences, University of North
Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

[†] Corresponding author Email: yoshizawa@aori.u-tokyo.ac.jp

Abstract

Diatoms are a major phytoplankton group responsible for about 20% of Earth's primary production. They carry out photosynthesis inside the plastid, an organelle obtained through eukaryote-eukaryote endosymbiosis. Recently, microbial rhodopsin, a photoreceptor distinct from chlorophyll-based photosystems, has been identified in certain diatoms. However, the physiological function of diatom rhodopsin is not well understood. Here we show that the diatom rhodopsin acts as a light-driven proton pump and localizes to the outermost membrane of the four membrane-bound complex plastids. Heterologous expression techniques were used to investigate the protein function and subcellular localization of diatom rhodopsin. Using model simulations, we further evaluated the physiological role of the acidic pool in the plastid produced by proton-transporting rhodopsin. Our results propose that the rhodopsin-derived acidic pool may be involved in a photosynthetic CO₂-concentrating mechanism and assist CO₂ fixation in diatom cells.

Introduction

Diatoms are unicellular, photosynthetic algae found throughout aquatic environments, and are responsible for up to 20% of annual net global carbon fixation¹. As their contribution to primary production in the ocean is significant, their light utilization mechanisms are essential to correctly understand marine ecosystems. Diatoms contain chlorophyll a and c and carotenoids such as fucoxanthin as photosynthetic pigments in the secondary plastids acquired by eukaryote-eukaryote endosymbiosis². In addition, certain diatoms have recently been shown to contain microbial rhodopsin (henceforth rhodopsins), a light-harvesting antenna distinct from the chlorophyll-containing antenna for photosynthesis³. Although rhodopsin-mediated light-harvesting may somehow support the survival of these diatoms in marine environments, the physiological role of rhodopsin in diatom cells remains unclear.

Microbial rhodopsins are a large family of seven transmembrane photoreceptor proteins⁴. Rhodopsin has an all-*trans* retinal as the light-absorbing chromophore, and its protein function is triggered by light-induced isomerization of the retinal. The first microbial rhodopsin, the light-driven proton pump bacteriorhodopsin (BR), was discovered in halophilic archaea⁵. Although rhodopsin was initially thought to occur only in halophilic archaea inhabiting hypersaline environments, subsequent studies have shown that the rhodopsin gene is widely distributed in all three domains of life^{6,7}. Rhodopsins can be classified based on their functions into light-driven ion pumps, light-activated signal transducers, and light-gated ion channels. The two former functional types of rhodopsin have so far been found in prokaryotes⁸. The rhodopsins in prokaryotes, regardless of function, are localized to the plasma membrane in which they operate. For example, proton-pumping rhodopsins export protons from the cytosol across the plasma membrane to convert light energy into a proton motive force (PMF)⁹. The PMF induced by rhodopsin ion transport is utilized by various physiological functions such as ATP synthesis, substrate uptake, and flagellar movement.

In eukaryotic microorganisms, rhodopsins functioning as light-driven ion pumps and light-gated ion channels have been reported^{8,10}. Light-gated ion channels are well studied; they exclusively localize in the eyespot of phytoplankton and are involved in phototaxis¹¹. The other type of rhodopsin in eukaryotes, light-driven ion-pumping rhodopsins have been found in a number of organisms belonging to both photoautotrophic and heterotrophic protists^{3,12}. Since the intracellular membrane structure of eukaryotic cells is more complex than that of prokaryotes, containing a variety of organelles, even light-driven ion-pumping rhodopsins may have distinct physiological roles depending on their subcellular localization¹². However, due to the difficulty in determining the exact localization of rhodopsins in eukaryotic cells, most rhodopsins are not even known for their subcellular localization.

In this study, to clarify the physiological function of rhodopsin in a marine pennate diatom, we investigated the phylogeny, protein function, spectroscopic characteristics, and subcellular localization of rhodopsin from a member of the genus *Pseudo-nitzschia*. Heterologous expression techniques were used to analyze the protein function and spectroscopic features. The expression eGFP-fusion rhodopsin also revealed the subcellular localization in the model diatom (*Phaeodactylum tricornutum*). Furthermore, a model-based analysis was performed to evaluate the impacts of the potential roles of the rhodopsin for cellular biology.

Results and Discussion

Rhodopsin sequences and phylogenetic analysis

We performed phylogenetic analysis using the rhodopsin (named PngR, accession no. AJA37445.1) of the diatom *Pseudo-nitzschia granii* and the microbial rhodopsin sequences reported to date³. This phylogenetic tree revealed that PngR is not included in the proteorhodopsin (PR) clade commonly found in the ocean but belongs to the Xanthorhodopsin (XR) -like rhodopsin (XLR) clade, which is presumed to have an outward proton transporting function (Fig. 1 and Extended Data Fig. 1). The

comparison of motif sequences necessary for ion transport showed that the amino acids in the putative proton donor and acceptor sites of XR and PR were conserved in PngR, suggesting that PngR functions as an outward proton pump (Extended Data Fig. 2). Furthermore, the homology search for rhodopsin sequence in the XLR clade from Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP) revealed that not only diatoms (Ochrophyta, Stramenopiles), but also dinoflagellates (Dinophyceae, Alveolata) and haptophytes have rhodopsin genes in the same XLR clade (Supporting Information Table S1). These results indicate that rhodopsin of the XLR clade is widely distributed among the major phytoplankton groups, which are important primary producers in the ocean.

Function and spectroscopic features of Diatom rhodopsin.

To characterize the function of the PngR, we heterologously expressed the synthesized rhodopsin gene in *Escherichia coli* cells. A light-induced decrease in pH was observed in the suspension of the *E. coli* cells expressing the PngR, and this decrease was almost completely abolished in the presence of the protonophore carbonyl cyanide M-chlorophenylhydrazone (CCCP) (Fig. 2A). The observed pH changes clearly showed that PngR exports protons from the cytoplasmic side across the plasma membrane.

Next, we examined the spectroscopic characteristics of PngR using recombinant protein purified from *E. coli*. The absorption maximum of PngR was located at 511 nm (Fig. 2B), which was explicitly shorter than that of XR (565 nm) and GR (*Gloeobacter* rhodopsin 541nm) in the XLR clade¹³. It is noteworthy that while *P. granii* is a marine species, while XR and GR are both distributed in terrestrial freshwater organisms. Thus, our observations are consistent with the shorter wavelength of the absorption maximum of rhodopsin in marine environments than in the terrestrial environment¹⁴, indicating that PngR is well adapted to the light conditions in the ocean.

We then examined the retinal configuration in PngR by high-performance liquid chromatography. In both in light- and dark-adapted samples, the isomeric state of

retinal was predominantly all-*trans* (Extended Data Fig. 3), which was similar to the isomeric state of retinal in prokaryotic GR in XLR clade but different from that in BR¹⁵. Since the pK_a value of the proton acceptor residue (Asp85 in BR) is an indicator of the efficiency of proton transport by rhodopsin, we estimated the pK_a values of the putative proton acceptor in PngR (Asp91) by pH titration experiment (Extended Data Fig. 4). This experiment estimated that the pK_a of this residue acceptor is about 5.0, indicating that the proton acceptor of PngR works well in marine and intracellular environments. Furthermore, the photochemical reactions that proceed behind the ion-transportation mechanism of PngR were examined by flash-photolysis analysis (Extended Data Fig. 5). All photocycle of PngR were completed in about 300 ms, suggesting that the cycle is fast enough to transport protons in a physiologically significant time scale. Detailed results and discussion of flash-photolysis analysis are described in the supplementary information.

Subcellular localization of PngR in a model diatom.

The PngR sequence bears neither apparent N-terminal extension nor detectable N-terminal signal peptide, and thus *in silico* analyses are not able to predict a possible subcellular localization of PngR. To identify the subcellular localization of PngR, a C-terminal eGFP-fusion PngR was expressed in the model diatom *Phaeodactylum tricornutum*, which can be transformed by electroporation and is often used for heterologous expression analysis^{16,17}. The transformed *P. tricornutum* cell was observed with differential interface contrast and epifluorescent microscopes (Fig. 3A). From these microscopic observations, we examined the morphology of *P. tricornutum* and the localization of recombinant PngR, nucleus, chlorophyll and other cellular organelles in multiple cells (Extended Data Fig. 6 and 7). The fluorescence signal of PngR:eGFP transformant was most likely localized in the periphery of chlorophyll fluorescence and DAPI signals, corresponding to the outermost plastid membrane, called the chloroplast endoplasmic reticulum membrane (CERM), which is physically connected to the

nuclear membrane (Fig. 3A). A few cells also exhibited GFP signals within vacuolar membranes in addition to CERM (Extended Data Fig. 7). Insertion of the complete sequence of the PngR:eGFP gene in the transformant DNA was confirmed by PCR followed by the Sanger sequencing.

Given the above heterologous expression experiment and microscopic observation, we expect that PngRs primarily localize to the outermost membrane of the plastid, although to a lesser extent, may also be found in the vacuolar membrane, which may be a function of other factors such as cell growth conditions. In addition, PngR should be oriented in the direction that the inside of the CERM is regarded as the outside of an analogous prokaryotic cell. These results imply that the light-driven proton transport by PngR could acidify the inner region of the CERM (Fig. 3B). If correct, what physiological role does the acidification of this region have in diatoms? The electrochemical gradient formed by rhodopsin could be a driving force for various secondary transport processes. Alternatively, since the primary purpose of plastids is photosynthesis, the formation of acidic pools is likely to be somehow related to photosynthesis. One possibility is that the appearance of this acidic pool influences carbon fixation, since in the carbonate system, acidification favors a higher concentration of CO₂ relative to the other dominant carbonate species in seawater (HCO₃⁻ and CO₃²⁻).

In general, under weakly alkaline conditions in the ocean, most of the dissolved inorganic carbon (DIC) is present in the form of HCO₃⁻, and only approximately 1% is present in the form of CO₂. However, ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO) localized in the stroma can only react with inorganic carbon (Ci) in the form of CO₂, but not of HCO₃⁻. The RuBisCO enzyme in diatoms has a low affinity even for CO₂ (K_m of 25~68 μM while CO₂aq in the ocean is about 10 μM at 25°C) and thus requires concentrated CO₂ for efficient fixation at the site of RuBisCO. In other words, the ocean has always been a CO₂-limited environment for most phytoplankton¹⁸. Consequently, due to membrane impermeability of HCO₃⁻, phytoplankton have

developed a variety of CO₂-concentrating mechanism (CCM) to efficiently transport Ci to the site of RuBisCO by placing HCO₃⁻ transporters in appropriate membranes and carbonic anhydrase (CA) in these compartments, the latter of which catalyzes the rapid interconversion between HCO₃⁻ and CO₂. This interconversion can be facilitated if the balance of the carbonate system is biased toward CO₂ with acidification.

In marine phytoplankton, especially those with secondary plastids, since RuBisCO is located in the stroma surrounded by three or four membranes, it is necessary to localize HCO₃⁻ transporters in each membrane¹⁹. Diatoms that possess four membrane-bound plastids are known to have a HCO₃⁻ transporter (e.g., SLC4) present in the plasma membrane¹⁶, but it is not well understood how CO₂ is concentrated in close proximity to RuBisCO in the stroma²⁰. In sharp contrast to the transporter-based mechanism, if the equilibrium of Ci can be shifted toward CO₂ by acidifying the inner region of the outermost membrane, it may be possible to transport CO₂ to RuBisCO in a membrane-permeable form through gas diffusion. However, since it is difficult to directly investigate the role of membrane protein in eukaryotic microbial organelles, the combination of observation and model simulation is a powerful approach²¹.

A quantitative model of Carbon concentration in diatoms: CFM-CC

Our subcellular localization analysis suggests that proton transport by rhodopsin acidifies the inner side of the outermost membrane of the plastid (hereafter "the middle space"). To quantitatively test the effect of the decreased pH in the middle space on C fixation, we developed a simple quantitative model of carbonate chemistry combined with membrane transport and C fixation (CFM-CC: Cell Flux Model of C Concentration) (Fig. 4 upper panel). Previously, a comprehensive model of CO₂ concentration within diatoms was developed²². CFM-CC uses a conceptually similar structure to this model, focusing on more specific membrane layers, designed to test the effect of altered pH in the middle space.

Our model result shows that the concentrations of CO₂ in the middle space ([CO₂]_m) strongly depend on the pH (pH_m), suggesting that proton pumping by rhodopsin could enhance C fixation (Fig. 4 bottom panel). The calculation of C chemistry in the middle space shows that decreasing pH_m favors higher [CO₂]_m at a given DIC concentration (Fig. 4A) (here we used 993 μmol L⁻¹ ²³). We note that the potential leaking of CO₂ to the cytosol may change the DIC in the middle space, but here we use a constant DIC value, since such effect has not been experimentally shown and is difficult to quantify due to unknown factors (e.g., the balance of DIC uptake and CO₂ leaking). At the reference point (here pH_m = 7.59 ²³), [CO₂]_m is 17 μmol L⁻¹ [y007 03], but it increases to 64, 410, and 870 μmol L⁻¹ for pH_m values of 7, 6, and 5, respectively [y007 04] (Fig. 4A).

Due to the increased [CO₂]_m, CO₂ concentration in the plastid [CO₂]_p also increases with lowered pH_m, but the level of increase depends on V_{max}/D (the ratio of maximum C fixation rate to the diffusion constant) (Fig. 4B). When V_{max}/D is small, diffusion of CO₂ from the middle space to the plastid dominates the influence, resulting in [CO₂]_p similar to [CO₂]_m. In contrast, when V_{max}/D is large, CO₂ uptake dominates, and there is small influence of [CO₂]_m on [CO₂]_p.

The rate of C fixation increases with lowered pH_m, since the increased [CO₂]_m leads to faster transport of CO₂ into the plastid (Fig. 4C). However, the magnitude of such increase depends on V_{max}/D . The model shows that the effect of pH_m on C fixation is larger when V_{max}/D is large, because the rate of C fixation is more strongly affected by [CO₂]_m [eq. 6], which is directly influenced by pH_m. However, when V_{max}/D is small, the C fixation rate is more strongly influenced by the C uptake kinetics [eq. 4], which saturates at relatively low [CO₂]_p (K value of 44 μmol L⁻¹ ²⁴). The increase in C fixation ranges 2.1-3.8, 3.2-24.2 and 3.4-51.2 times when pH_m decreases from 7.59 to 7, 6, 5 respectively [y008 13]. These results suggest two things. First, any V_{max}/D , pH_m has a considerable effect on C uptake. Second, it would be highly favorable for the cells to have high V_{max} , relative to the diffusivity of CO₂ across the membrane. This is most

likely the case since when there are multiple membranes between the middle space and plastid, which would lower the D value^{25,26}. Thus, this simple yet elegant system with rhodopsin to manipulate the pH_m can provide a powerful mechanism in C concentration and thus support a high C fixation in some rhodopsin-containing diatoms, enabling them to be more productive phytoplankton in the ocean.

Such a CCM based on CO_2 diffusion (termed pump-leak type) is known to be a possible mechanism by placing CAs in appropriate locations. For example, *Nannochloropsis oceanica* (*Ochrophyta*), possessing the same four membrane-bound complex plastids as found in diatoms, is thought to generate CO_2 by placing CA in the middle space²⁷. Besides, centric diatom *Chaetoceros gracilis* is thought to use externally placed CAs to allow CO_2 to flow into the cell²⁸. In contrast to the CA-based model, acidification-based model was formerly proposed to facilitate CO_2 fixation of RuBisCO in the thylakoid lumen of plastids; HCO_3^- is converted into CO_2 by acidification by photosynthetic proton pumping into the thylakoid lumen²⁹. Our model further expands the model and proposes that the middle space acidified by proton-pumping rhodopsin also plays the role of a CO_2 generator. Our proposed mechanism would be useful in most parts of the ocean where CO_2 chronically limits photosynthesis, but would be even more valuable in specific environments. For example, since CA, which plays a central role in CCM, requires cobalt or zinc ions as the reaction center, and photosynthetic proton pumping systems require iron, rhodopsin-derived acidic pools may be useful for the transport of Ci in oceans where these metal ions are depleted (such as the HNLC region of the North Pacific Ocean)³⁰. Indeed, in the HNLC region of the North Pacific, where the *P. granii* rhodopsin-containing cells were initially isolated and sequenced³¹, it has been pointed out that primary production may be limited by iron and influenced by other trace metals³². In other words, this mechanism seems to be particularly effective in the ocean, where trace metals involved in CCM are depleted.

Conclusion

In this study, we have clarified the function and subcellular localization of PngR in a photosynthetic diatom. Our results suggested that proton transport by rhodopsin creates a local acidic pool inside the outermost membrane of the plastid. Based on the quantitative simulation, we propose a conceptual model of rhodopsin-contributing CCM (Extended Data Fig. 8). The acidic pool created by light is likely involved in the CCM and provides positive feedback on carbon fixation efficiency. Future analyses of cultured rhodopsin-bearing microbial eukaryotes would corroborate our results and allow for the evaluation of how rhodopsin-mediated proton transport promote their growth in the ocean.

References

- 1 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 Keeling, P. J. Diversity and evolutionary history of plastids and their hosts. *Am. J. Bot.* **91**, 1481-1493 (2004).
- 3 Marchetti, A., Catlett, D., Hopkinson, B. M., Ellis, K. & Cassar, N. Marine diatom proteorhodopsins and their potential role in coping with low iron availability. *The ISME Journal* **9**, 2745-2748 (2015).
- 4 Spudich, J. L., Yang, C. S., Jung, K. H. & Spudich, E. N. Retinylidene proteins: Structures and functions from archaea to humans. *Annu. Rev. Cell Dev. Biol.* **16**, 365-392 (2000).
- 5 Oesterhelt, D. & Stoeckenius, W. Rhodopsin-like protein from the purple membrane of *Halobacterium halobium*. *Nature New Biol.* **233**, 149-152 (1971).
- 6 Beja, O. *et al.* Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science* **289**, 1902-1906 (2000).
- 7 Sineshchekov, O. A., Jung, K. H. & Spudich, J. L. Two rhodopsins mediate phototaxis to low- and high-intensity light in *Chlamydomonas reinhardtii*. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 8689-8694 (2002).
- 8 Ernst, O. P. *et al.* Microbial and animal rhodopsins: structures, functions, and molecular mechanisms. *Chem. Rev.* **114**, 126-163 (2014).
- 9 Yoshizawa, S., Kawanabe, A., Ito, H., Kandori, H. & Kogure, K. Diversity and functional analysis of proteorhodopsin in marine *Flavobacteria*. *Environ. Microbiol.* **14**, 1240-1248 (2012).
- 10 Kikuchi, M. *et al.* Functional expression of the eukaryotic proton pump rhodopsin OmR2 in *Escherichia coli* and its photochemical characterization. *Sci. Rep.* **11**, 1-13 (2021).
- 11 Nagel, G. *et al.* Channelrhodopsin-1: A light-gated proton channel in green algae. *Science* **296**, 2395-2398 (2002).
- 12 Slamovits, C. H., Okamoto, N., Burri, L., James, E. R. & Keeling, P. J. A bacterial proteorhodopsin proton pump in marine eukaryotes. *Nat. Commun.* **2** (2011).
- 13 Balashov, S. P. *et al.* Xanthorhodopsin: a proton pump with a light-harvesting carotenoid antenna. *Science* **309**, 2061-2064 (2005).
- 14 Man, D. L. *et al.* Diversification and spectral tuning in marine proteorhodopsins. *EMBO J.* **22**, 1725-1731 (2003).
- 15 Miranda, M. R. M. *et al.* The photocycle and proton translocation pathway in a cyanobacterial ion-pumping rhodopsin. *Biophys. J.* **96**, 1471-1481 (2009).
- 16 Nakajima, K., Tanaka, A. & Matsuda, Y. SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1767-1772

313 (2013).

314 17 Dorrell, R. G. *et al.* Principles of plastid reductive evolution illuminated by
315 nonphotosynthetic chrysophytes. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 6914-6923 (2019).

316 18 Riebesell, U., Wolgastrow, D. A. & Smetacek, V. Carbon-dioxide limitation of marine-
317 phytoplankton growth-rates. *Nature* **361**, 249-251 (1993).

318 19 Tsuji, Y., Nakajima, K. & Matsuda, Y. Molecular aspects of the biophysical CO₂-
319 concentrating mechanism and its regulation in marine diatoms. *J. Exp. Bot.* **68**, 3763-3772
320 (2017).

321 20 Matsuda, Y., Hopkinson, B. M., Nakajima, K., Dupont, C. L. & Tsuji, Y. Mechanisms of
322 carbon dioxide acquisition and CO₂ sensing in marine diatoms: a gateway to carbon
323 metabolism. *Philos. Trans. R. Soc. B* **372**, 20160403 (2017).

324 21 Hopkinson, B. M., Dupont, C. L., Allen, A. E. & Morel, F. M. M. Efficiency of the CO₂-
325 concentrating mechanism of diatoms. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 3830-3837
326 (2011).

327 22 Hopkinson, B. M. A chloroplast pump model for the CO₂ concentrating mechanism in
328 the diatom *Phaeodactylum tricornutum*. *Photosynth. Res.* **121**, 223-233 (2014).

329 23 Burns, B. D. & Beardall, J. Utilization of Inorganic Carbon by Marine Microalgae. *J. Exp.*
330 *Mar. Bio. Ecol.* **107**, 75-86 (1987).

331 24 Jensen, E. L., Maberly, S. C. & Gontero, B. Insights on the functions and ecophysiological
332 relevance of the diverse carbonic anhydrases in microalgae. *Int. J. Mol. Sci.* **21**, 2922
333 (2020).

334 25 Eichner, M. *et al.* N₂ fixation in free - floating filaments of *Trichodesmium* is higher than
335 in transiently suboxic colony microenvironments. *New Phytol.* **222**, 852-863 (2019).

336 26 Inomura, K., Wilson, S. T. & Deutsch, C. Mechanistic model for the coexistence of
337 nitrogen fixation and photosynthesis in marine *Trichodesmium*. *mSystems* **4**, e00210-
338 00219 (2019).

339 27 Gee, C. W. & Niyogi, K. K. The carbonic anhydrase CAH1 is an essential component of
340 the carbon-concentrating mechanism in *Nannochloropsis oceanica*. *Proc. Natl. Acad. Sci.*
341 *U. S. A.* **114**, 4537-4542 (2017).

342 28 Tsuji, Y. *et al.* Characterization of a CO₂-Concentrating Mechanism with Low Sodium
343 Dependency in the Centric Diatom *Chaetoceros gracilis*. *Mar. Biotechnol.* **23**, 456-462
344 (2021).

345 29 Raven, J. A. CO₂-concentrating mechanisms: A direct role for thylakoid lumen
346 acidification? *Plant Cell Environ.* **20**, 147-154 (1997).

347 30 Moore, C. M. *et al.* Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**,
348 701-710 (2013).

349 31 Marchetti, A. *et al.* Comparative metatranscriptomics identifies molecular bases for the
350 physiological responses of phytoplankton to varying iron availability. *Proc. Natl. Acad. Sci.*
351 *U. S. A.* **109**, E317-E325 (2012).

352 32 Saito, M. A., Goepfert, T. J. & Ritt, J. T. Some thoughts on the concept of colimitation:
353 Three definitions and the importance of bioavailability. *Limnol. Oceanogr.* **53**, 276-290
354 (2008).

Figures and Figure Legends

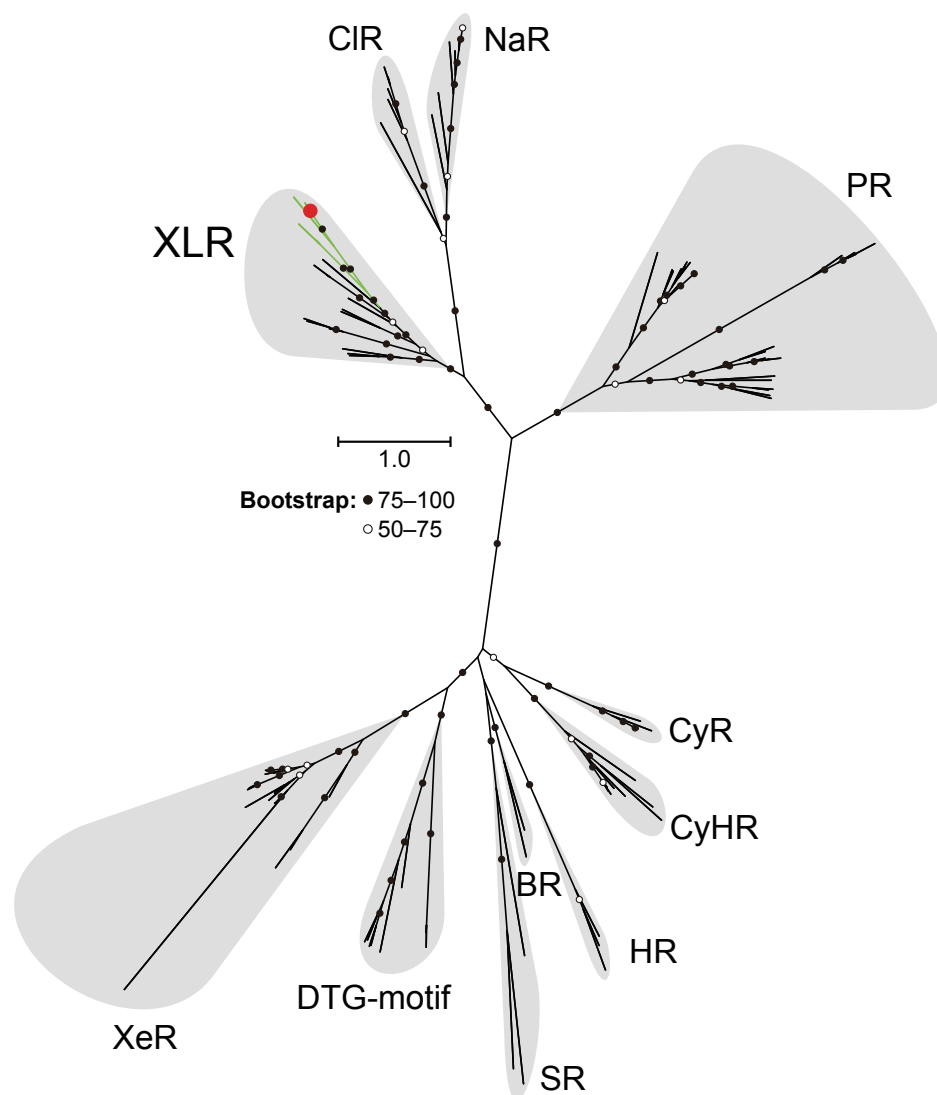


Figure 1. Yoshizawa et al.

Figure 1. Phylogenetic position of diatom rhodopsin. Maximum likelihood tree of amino acid sequences of microbial rhodopsins. Diatom rhodopsin (PngR) is indicated by a red circle and bootstrap probabilities ($\geq 50\%$) are indicated by black and white circles. Green branches indicate eukaryotic rhodopsins used in this analysis, and black branches indicate others. Rhodopsin clades are as follows: XLR (Xanthorhodopsin-like rhodopsin), CIR (Cl⁻-pumping rhodopsin), NaR (Na⁺-pumping rhodopsin), PR (proteorhodopsin), XeR (xenorhodopsin), DTG-motif rhodopsin, SR (sensory rhodopsin-I and sensory rhodopsin-II), BR (bacteriorhodopsin), HR (halorhodopsin), CyHR (cyanobacterial halorhodopsin), and CyR (cyanorhodopsin).

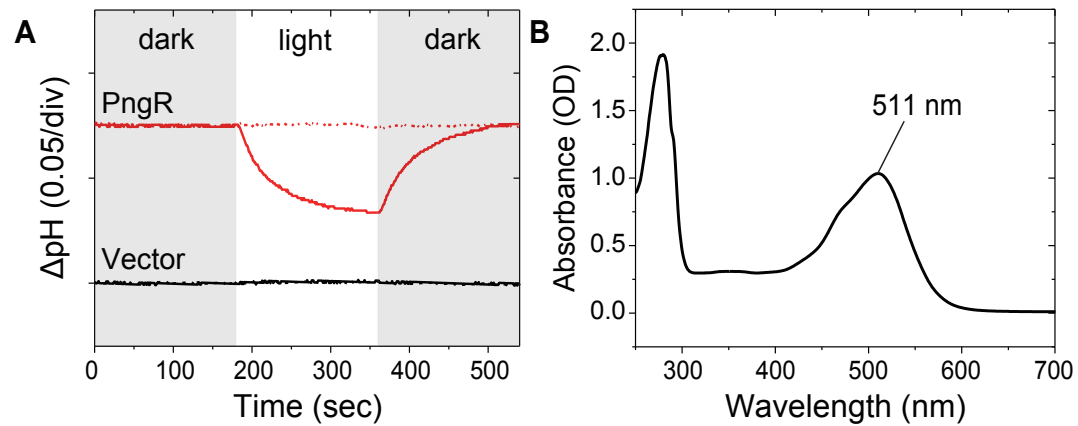


Figure 2. Yoshizawa et al.

Figure 2. Light-induced pH changes and absorption spectrum of PngR. (A) Outward proton pump activity of PngR in *E. coli* cells. Light-induced pH changes of solutions containing *E. coli* cells with the expression plasmid for PngR (upper panel) and the empty vector pET21a (lower panel) in the presence (red dashed line) or absence (red solid line) of CCCP. The white-filled region indicates the period of illumination. (B) Absorption spectrum of purified PngR in Buffer A (50 mM Tris-HCl, pH 7.0, 1 M NaCl and 0.05% (w/v) DDM).

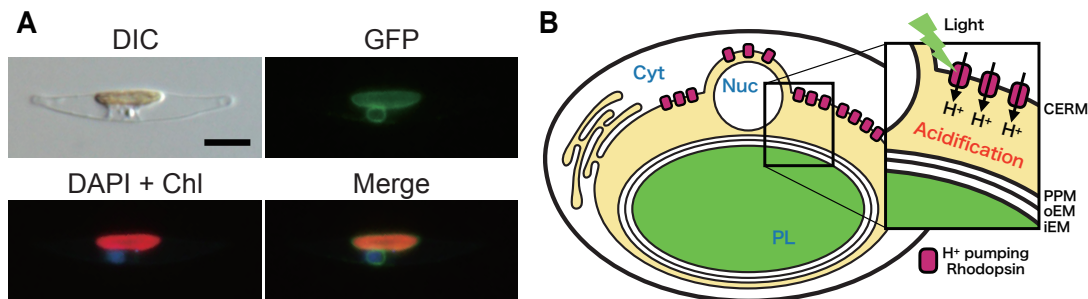


Figure 3. Yoshizawa et al.

Figure 3. Subcellular localization of rhodopsins in diatom cell. (A) A transformed diatom cell was observed with DIC (Differential Interface Contrast) (Upper left). Green fluorescence from the recombinant PngR (GFP) (Upper right). The nuclear DNA stained with DAPI and the chlorophyll autofluorescence (DAPI + Chl) and merged image (Merge) are shown in bottom left and bottom right, respectively. Scale bar indicates 5 μ m. (B) A mode for subcellular localization of PngR. The proton transport of PngR acidifies the region (the middle space) surrounded by the membrane of CERM and PPM. Abbreviations are as follows: Cyt (Cytosol), Nuc (Nucleus), PL (Plastid), CERM (Chloroplast endoplasmic reticulum membrane), PPM (Periplastidial membrane), oEM (Outer plastid envelop membrane) and iEM (Internal plastid envelop membrane).

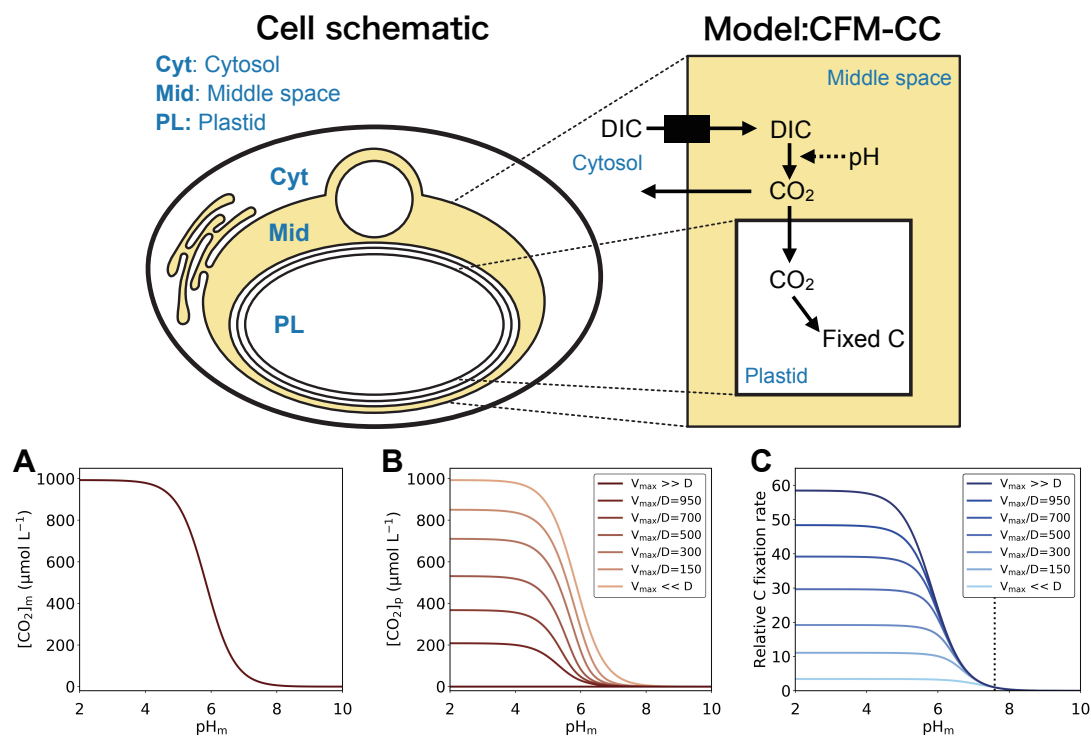


Figure 4. Yoshizawa et al.

Figure 4. A quantitative model of carbon concentration in diatoms. (Upper panel) Schematic of Cell Flux Model of C Concentration (CFM-CC). The left panel represents the actual cell and the right panel represents the model. Solid arrows show net flux of C and dashed arrow indicates the influence of pH. (Bottom panel) The influence of pH in the middle space on CO₂ concentrations and the photosynthesis rate. (A) CO₂ concentrations in the middle space [CO₂]_m. (B) CO₂ concentrations in the plastid [CO₂]_p. (C) C fixation rate relative to that with pH in the middle space of 7.59, the only value we found for intracellular pH in a diatom²³. (B) and (C) are plotted for various V_{max}/D . The solution for [CO₂]_m in (A) is independent of V_{max}/D .

Methods

Rhodopsin sequences and phylogenetic analysis

The rhodopsin sequence of the diatom *Pseudo-nitzschia granii* was previously reported³. All other rhodopsin sequences used in the phylogenetic analysis were collected from the National Center for Biotechnology Information. Detailed information on the strains used in this analysis is given in Extended Data Fig. 1. The sequences were aligned using MAFFT version 7.453 with options '--genafpair' and '--maxiterate 1000'. The phylogenetic tree was inferred using RAxML (v.8.2.12) with the 'PROTGAMMALGF' model using 1000 rapid bootstrap searches. Model selection was performed with the ProteinModelSelection.pl script in the RAxML package.

The search for eukaryotic rhodopsins that belong to the XLR clade was performed among protein sequences of Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP)³³. Namely, the phylogenetic placement of rhodopsin proteins from MMETSP using pplacer (v1.1.alpha19)³⁴ was performed on a prebuilt large-scale phylogenetic tree of rhodopsins and extracted placements on the XLR clade using gappa (v0.6.0)³⁵.

Gene preparation, protein expression and ion transport measurements of *E. coli* cells.

The full-length cDNA for PngR, whose codons were optimized for *E. coli*, were chemically synthesized by Eurofins Genomics and inserted into the NdeI-XhoI site of the pET21a(+) vector as previously described³⁶. A hexa-histidine-tag was fused at the C-terminus of PngR, which was utilized for purification of the expressed protein. The heterologous protein expression method is the same as previously reported³⁷. *E. coli* BL21(DE3) cells harboring the cognate plasmid were grown at 37 °C in LB medium supplemented with ampicillin (final concentration = 50 µg mL⁻¹). Protein expression was induced at an optical density at 600 nm of 0.7–1.2 with 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and 10 µM all-trans retinal, after which the cells were

incubated at 37 °C for 3 h. The proton transport activity of PngR was measured as light-induced pH changes of suspensions of *E. coli* cells as previously described³⁷. In short, cells expressing PngR were washed more than three times in 150 mM NaCl and were then resuspended in the same solution for measurements. Each cell suspension was placed in the dark for several min and then illuminated using a 300 W Xenon lamp (ca. 30 mW cm⁻², MAX-303, Asahi spectra, Japan) through a > 460 nm long-pass filter (Y48, HOYA, Japan) for 3 min. Measurements were repeated under the same conditions after the addition of the protonophore carbonyl cyanide m-chlorophenylhydrazone (CCCP) (final concentration = 10 µM). Light-induced pH changes were monitored using a Horiba F-72 pH meter. All measurements were conducted at 25 °C using a thermostat (Eyela NCB-1200, Tokyo Rikakikai Co. Ltd, Japan).

Purification of PngR from *E. coli* cells and spectroscopic measurements of the purified rhodopsin.

E. coli cells expressing PngR were disrupted by sonication for 30 min in ice-cold water in a buffer containing 50 mM Tris–HCl (pH 7.0) and 300 mM NaCl. The crude membrane fraction was collected by ultracentrifugation and solubilized with 1.0% (w/v) n-dodecyl-β-d-maltoside (DDM, DOJINDO Laboratories, Japan). The solubilized fraction was purified by Ni²⁺ affinity column chromatography with a linear gradient of imidazole as described previously³⁸. The purified protein was concentrated by centrifugation using an Amicon Ultra filter (30,000 M_w cut-off; Millipore, USA). The sample media was then replaced with Buffer A (50 mM Tris–HCl, pH 7.0, 1 M NaCl and 0.05% (w/v) DDM) by ultrafiltration for 3-times.

Absorption spectra of purified proteins were recorded using a UV-2450 spectrophotometer (Shimadzu, Japan) at room temperature in Buffer A. The retinal composition in PngR was analyzed by high-performance liquid chromatography (HPLC) as described previously³⁹. For dark-adaptation, the samples were kept in the dark condition for more than 72 hrs at 4 °C. For light-adaptation, the samples were

illuminated for 3 min at 520 ± 10 nm, where the light power was adjusted to approximately 10 mW cm^{-2} . The molar compositions of the retinal isomers were calculated from the areas of the peaks in HPLC patterns monitored at 360 nm using the extinction coefficients of retinal oxime isomers as described previously³⁹. For pH titration experiments, the samples were suspended in Buffer A. The pH values of the samples were adjusted to the desired acidic values by adding HCl, after which the absorption spectra were measured at each pH value. All measurements were conducted at room temperature (approx. 25°C) under room light. After the measurements, the reversibility of the spectral changes was checked to confirm that the sample was not denatured during the measurements. The absorption changes at specific wavelengths were plotted against pH values and the plots were fitted to the Henderson–Hasselbalch equation assuming single pK_a value as previously described³⁷.

Transient time-resolved absorption spectra of the purified proteins from 380 to 700 nm at 5 nm intervals were obtained using a homemade computer-controlled flash photolysis system equipped with an Nd: YAG laser as an actinic light source. Using an optical parametric oscillator, the wavelength of the actinic pulse was tuned at 510 nm for PngR. The pulse intensity was adjusted to 2 mJ per pulse. All data were averaged to improve the signal-to-noise ratio ($n = 30$). All measurements were conducted at 25°C . For these experiments, the samples were suspended in Buffer A. After the measurements, the reproducibility of the data was checked to confirm that the sample was not denatured during the measurements. To investigate proton uptake and release during the photocycle, we used the pH indicator pyranine (final concentration = $100 \mu\text{M}$, Tokyo Chemical Industry Co., Ltd, Japan), which has been extensively used to monitor light-induced pH changes in various rhodopsins. The pH changes in the bulk environment were measured as the absorption changes of pyranine at 450 nm. The absorption changes of pyranine were obtained by subtracting the absorption changes of samples without pyranine from those of samples with pyranine. The experiments using pyranine were performed in an unbuffered solution containing 1 M NaCl and 0.05%

(w/v) DDM (pH 7.0) to enhance the signals. The results of 1000-traces were averaged to improve the signal-to-noise ratio.

Subcellular localization of PngR in the model diatom.

The PngR:eGFP recombinant gene, coding the full length of PngR C-terminally tagged with eGFP, was cloned into the expression vector for the model diatom *Phaeodactylum tricornutum*, pPha-NR⁴⁰, by CloneEZ (GenScript) following the manufacturer's instruction. The plasmid was electroporated into *P. tricornutum* cells with NEPA21 Super Electroporator (NEPAGENE), and transformed cells were selected with a Zeocin-based antibiotic treatment as described previously^{17,41}. Selected clones were observed under an Olympus BX51 fluorescent microscope (Olympus) equipped with an Olympus DP72 CCD color camera (Olympus). The nucleus stained with DAPI and chlorophyll autofluorescence from the plastid were observed with a 420 nm filter by 330 to 385 nm excitation. GFP fluorescence was detected with a 510 to 550 nm filter by 470 to 495 nm excitation.

Quantitative model of C concentration in a diatom: CFM-CC

Membrane transport model. In this model, we combined membrane transport of CO₂ and C fixation (Fig. 4). Parameter definitions, units and values are provided in Supporting Information Table S2 and S3, respectively. The key model equation is the balance of CO₂ concentration in cytosol $[CO_2]_p$:

$$\frac{d[CO_2]_p}{dt} = D([CO_2]_m - [CO_2]_p) - V_{Cfix} \quad [\text{eq. 1}]$$

where t is time (units), D is diffusion coefficient, $[CO_2]_m$ is CO₂ concentration (units) in the middle space. Here the first term represents the diffusion of CO₂ from the middle space to cytosol and the second term V_{Cfix} represents the C fixation rate following Michaelis–Menten kinetics^{22,42}:

$$V_{Cfix} = V_{max} \frac{[CO_2]_p}{[CO_2]_p + K} \quad [\text{eq. 2}]$$

where V_{max} is maximum CO₂ fixation rate (units), and K is a half saturation constant (units). $[CO_2]_m$ is obtained based on the carbonate chemistry in the middle space (see

below). Under the steady-state, [eq. 1] with [eq. 2] becomes the following quadratic relationship for $[CO_2]_p$:

$$[CO_2]_p^2 + \left(\frac{V_{max}}{D} + K - [CO_2]_m\right) [CO_2]_p - K[CO_2]_m = 0 \quad [\text{eq. 3}]$$

Solving this equation for $[CO_2]_p^2$ leads to:

$$[CO_2]_p = \frac{-\left(\frac{V_{max}}{D} + K - [CO_2]_m\right) + \sqrt{\left(\frac{V_{max}}{D} + K - [CO_2]_m\right)^2 + 4K}}{2} \quad [\text{eq. 4}]$$

Note that the other solution for the negative root is unrealistic since it may lead to the overall negative value of $[CO_2]_p$. Once we obtain $[CO_2]_p$, we can then calculate the rate of C fixation V_{Cfix} with [eq. 2].

Also, from [eq. 4], we can get two extreme solutions. First, when $V_{max} \ll D$ (i.e., when the CO_2 uptake capacity is small relative to the speed of CO_2 diffusion) [eq. 4] leads to

$$[CO_2]_p \sim [CO_2]_m \quad [\text{eq. 5}]$$

With this relationship and [eq. 2], V_{Cfix} is computed as follows:

$$V_{Cfix} \sim V_{max} \frac{[CO_2]_m}{[CO_2]_m + K} \quad [\text{eq. 6}]$$

Second, when $V_{max} \gg D$ (i.e., when the CO_2 uptake capacity is high relative to the CO_2 diffusion across the membrane) [eq. 4] becomes

$$[CO_2]_p \sim 0 \quad [\text{eq. 7}]$$

Under the steady state, [eq. 1] becomes

$$V_{Cfix} = D([CO_2]_m - [CO_2]_p) \quad [\text{eq. 8}]$$

and plugging [eq. 7] into [eq. 8] leads to

$$V_{Cfix} \sim D[CO_2]_m \quad [\text{eq. 9}]$$

and V_{Cfix} is calculated. We note that $[CO_2]_p > [CO_2]_m$ could occur if there were membrane-bound transporters for HCO_3^- located on each membrane between the middle space and plastid²². However, such a set of transporters have not been discovered²⁰. Thus, our model is fidelity to the current state of knowledge. Even if $[CO_2]_p > [CO_2]_m$ was the case, moderately decreased pH_m and thus increased $[CO_2]_m$ could be useful since they would decrease the gradient of CO_2 across the membranes (i.e., $[CO_2]_p$ vs $[CO_2]_m$), mitigating the diffusive loss of CO_2 from the plastid.

Carbonate chemistry in the middle space. The above equations can be solved once we

obtain $[CO_2]_m$. The model uses a given DIC (dissolved inorganic C) concentration in the middle space $[DIC]_m$ to calculate $[CO_2]_m$ following the established equations for carbon chemistry⁴³.

$$[CO_2]_m = \frac{[DIC]_m}{1 + \frac{K_1}{[H^+]_m} + \frac{K_1 K_2}{[H^+]_m^2}} \quad [\text{eq. 10}]$$

where $[H^+]_m$ is H^+ concentration (thus 10^{-pH} mol l^{-1}) in the middle space and K_1 and K_2 are temperature and salinity dependent parameters^{43,44}:

$$K_1 = 10^{-pK_1} \quad [\text{eq. 11}]$$

$$K_2 = 10^{-pK_2} \quad [\text{eq. 12}]$$

where

$$pK_1 = \frac{3633.86}{T} - 61.2172 + 9.6777 \ln(T) - 0.011555S + 0.0001152S^2 \quad [\text{eq. 13}]$$

$$pK_2 = \frac{471.78}{T} + 25.9290 - 3.16967 \ln(T) - 0.01781S + 0.0001122S^2 \quad [\text{eq. 14}]$$

Code availability

Code for CFM-CC is freely available from GitHub/Zenodo at <https://zenodo.org/record/5182712> (DOI: 10.5281/zenodo.5182712).

References

- 33 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).
- 34 Matsen, F. A., Kodner, R. B. & Armbrust, E. V. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC bioinformatics* **11**, 1-16 (2010).
- 35 Czech, L., Barbera, P. & Stamatakis, A. Genesis and Gappa: processing, analyzing and visualizing phylogenetic (placement) data. *Bioinformatics* **36**, 3263-3265 (2020).
- 36 Hasegawa, M. *et al.* A unique clade of light-driven proton-pumping rhodopsins evolved in the cyanobacterial lineage. *Sci. Rep.* **10**:16752 (2020).
- 37 Inoue, S. *et al.* Spectroscopic characteristics of Rubricoccus marinus xenorhodopsin (RmXeR) and a putative model for its inward H⁺ transport mechanism. *Phys. Chem. Chem. Phys.* **20**, 3172-3183 (2018).
- 38 Kojima, K. *et al.* Vectorial Proton Transport Mechanism of RxR, a Phylogenetically Distinct and Thermally Stable Microbial Rhodopsin. *Sci. Rep.* **10**:282 (2020).
- 39 Kojima, K. *et al.* Green-sensitive, long-lived, step-functional anion channelrhodopsin-2 variant as a high-potential neural silencing tool. *J. Phys. Chem.* **11**, 6214-6218 (2020).
- 40 Stork, S. *et al.* Distribution of the SELMA Translocon in Secondary Plastids of Red Algal Origin and Predicted Uncoupling of Ubiquitin-Dependent Translocation from Degradation. *Eukaryot. Cell* **11**, 1472-1481 (2012).
- 41 Miyahara, M., Aoi, M., Inoue-Kashino, N., Kashino, Y. & Ifuku, K. Highly Efficient Transformation of the Diatom *Phaeodactylum tricornutum* by Multi-Pulse Electroporation. *Biosci. Biotechnol. Biochem.* **77**, 874-876 (2013).
- 42 Berg, J., Tymoczko, J. & Stryer, L. *Biochemistry, ed 7th edition.* (New York, WH Freeman, 2010).
- 43 Emerson, S. & Hedges, J. *Chemical oceanography and the marine carbon cycle.* (Cambridge University Press, 2008).
- 44 Lueker, T. J., Dickson, A. G. & Keeling, C. D. Ocean pCO₂ calculated from dissolved inorganic carbon, alkalinity, and equations for K₁ and K₂: validation based on laboratory measurements of CO₂ in gas and seawater at equilibrium. *Mar. Chem.* **70**, 105-119 (2000).

Acknowledgements

We thank Yusuke Matsuda for useful discussion. This work was supported by JSPS KAKENHI Grant Numbers 18K19224 and 18H04136 to S.Y., 19H04727, 21H00404 and 21H02446 to Y.S., and 19H03274 to R.K., and NSF grants OPP1745036 to A.M. This research was partially supported by the Interdisciplinary Collaborative Research Program of the Atmosphere and Ocean Research Institute, the University of Tokyo.

Author contributions

S.Y., A.M., Y.S., and R.K. designed the research; S.Y., T.A., K.K., K.Inomura, M.H., Y.N., M.K., K.Ifuku, and R.K. performed the research; S.Y., K.K., K.Inomura., Y.N., M.K., G.A., H.M., T.Y., H.F., A.M., Y.S. and R.K. analyzed the data; and S.Y., K.Inomura., Y.S., and R.K. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper.

Correspondence and requests for materials should be addressed to S.Y.