Science Advances
AAAS

Supplementary Materials for

Mixotrophic growth of a ubiquitous marine diatom

Manish Kumar *et al.*

Corresponding author: Karsten Zengler, kzengler@ucsd.edu

Sci. Adv. **10**, eado2623 (2024)
DOI: 10.1126/sciadv.ado2623

The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

Supplementary Text

Reconstruction method used to build *i*MK1961

The genome-scale metabolic model of *C. clostridium* was reconstructed using a multi-step pipeline. First, bi-directional BLASTP *(111)* was employed to compare protein sequences of *C. clostridium* with those of template organisms. Based on homology between proteins of *C. clostridium* and template organisms, a draft metabolic network was constructed. This draft reconstruction was extended by incorporating annotated genes with EC numbers in the *C. clostridium* genome. The reconstruction was further refined by assigning compartments to reactions, adding biomass *(68)*, transport and exchange reactions, and balancing mass and charge in reactions, as well as correcting gene-reaction associations. To incorporate any missing information in the model, a variety of resources such as KEGG *(112)*, modelSEED *(113)*, MetaCyc *(114)*, UniProt *(115)*, BRENDA *(116)*, IntEnz *(117)*, TPDB *(118)*, and TransportDB *(119)* were utilized.

In addition, organism-specific pathways (e.g., succinate pathway) were added based on experimental evidence, and the model's predicted capabilities were compared with experimental measurements. In cases where the model's predictions did not match experimental data, the model was iteratively refined by filling gaps in relevant pathways. To reconstruct and analyze the model, three types of computational tools were used: alignment tools (such as BLASTP), model reconstructing and analyzing tools (including RAVEN Toolbox, COBRA Toolbox, and COBRApy), and subcellular localization tools (including SignalP *(120)*, HECTAR *(121)*, Mitoprot *(122)*, predictNLS *(123)*, and TargetP *(124)*). Finally, experimentally confirmed phenotypic capabilities of *C. clostridium (103, 105, 106, 108-110, 125)* were used to gap-fill the model (*i*MK1961) by adding missing pathways/reactions to enable the related metabolic functions.

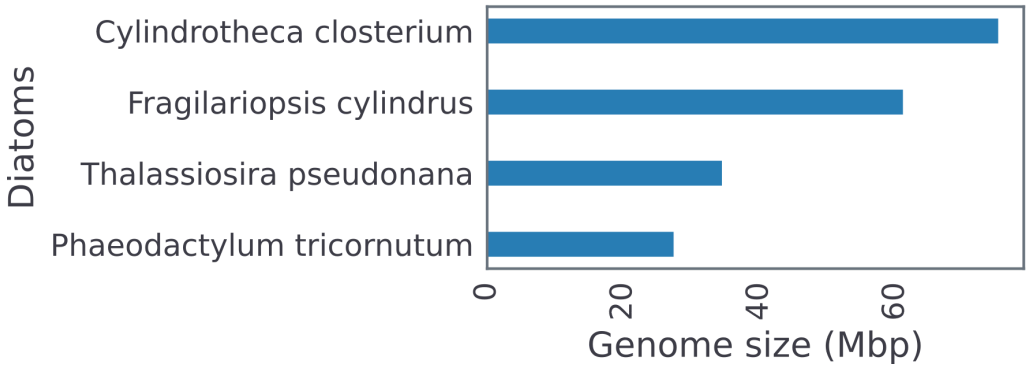

fig. S1: Comparison between four diatoms based on their genome size. In addition to *Cylindrotheca closterium*, this comparison includes three other diatoms with published metabolic models.

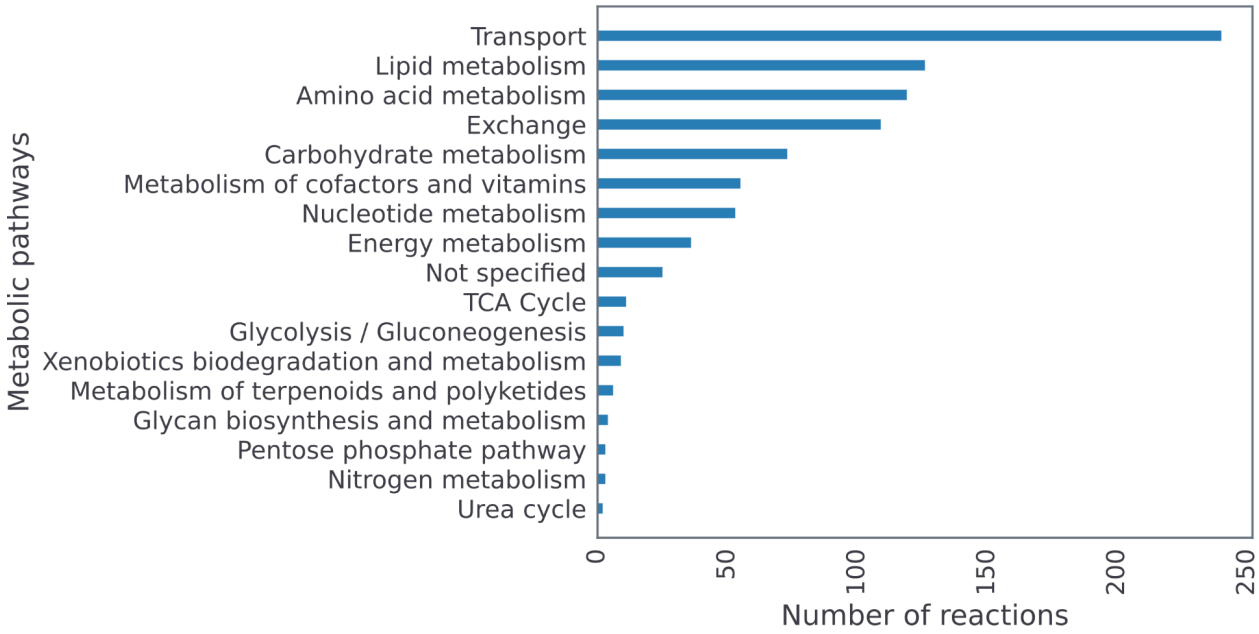

fig. S2: Distribution of unique reactions in the *C. closterium* model among metabolic pathways compared to other diatom models. The features of *i*MK1961 were compared with three other diatom models (*T. pseudonana*, *P. tricornutum,* and *F. cylindrus*).

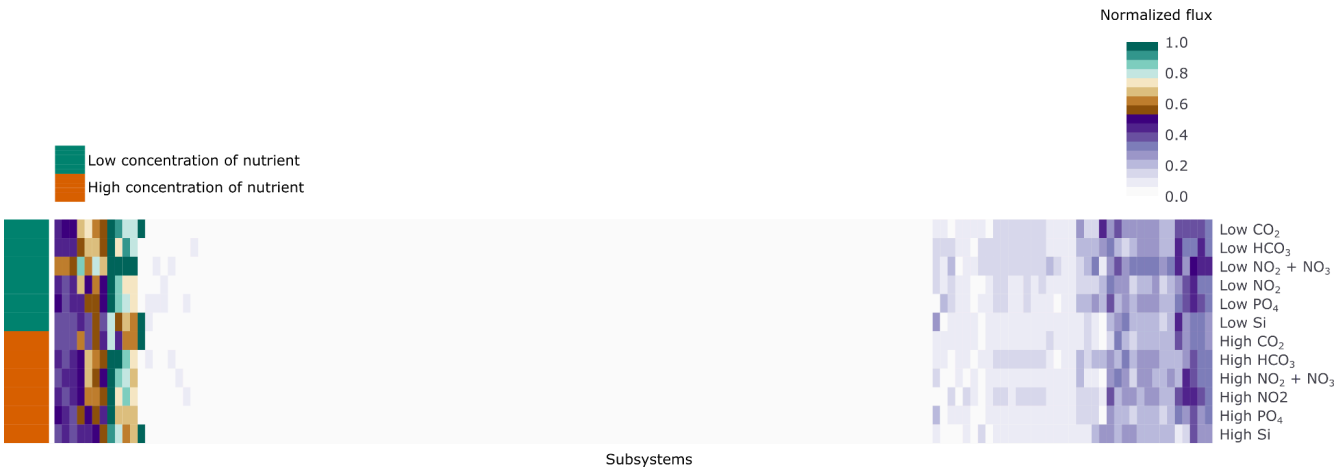

fig. S3: Model-predicted metabolic flux distributions through subsystems under low and elevated concentrations of nutrients such as CO_2, HCO_3, NO_2, NO_2 + NO_3, PO_4, and Si. Flux values were normalized between 0 and 1. table S25 contains the list of subsystems in the same sequence it is plotted in the heatmap.

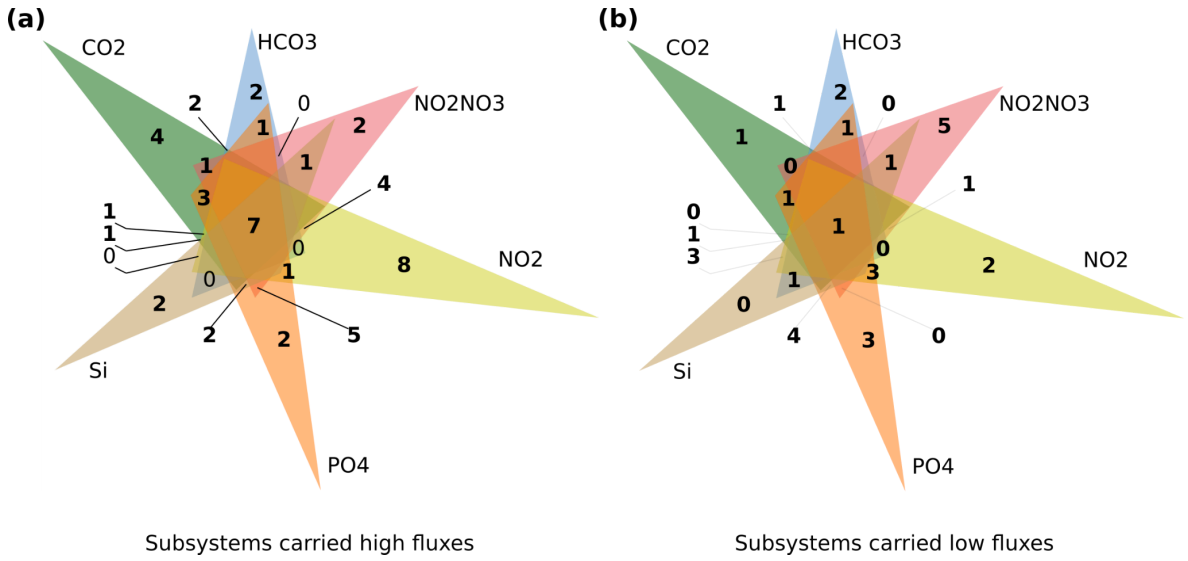

fig. S4: Comparing subsystems with differential flux among six growth environments CO_2, HCO_3, NO_2, NO_2 + NO_3, PO_4, and Si. (a) Seven subsystems (starch and sucrose metabolism, methionine metabolism, nucleotide sugar metabolism, aminoacyl-tRNA biosynthesis, ascorbate metabolism, fructose and mannose metabolism, porphyrin and chlorophyll metabolism) with high and **(b)** three subsystems (oxidative phosphorylation, arginine and proline metabolism, biotin biosynthesis) with low flux distributions were found common between all conditions.

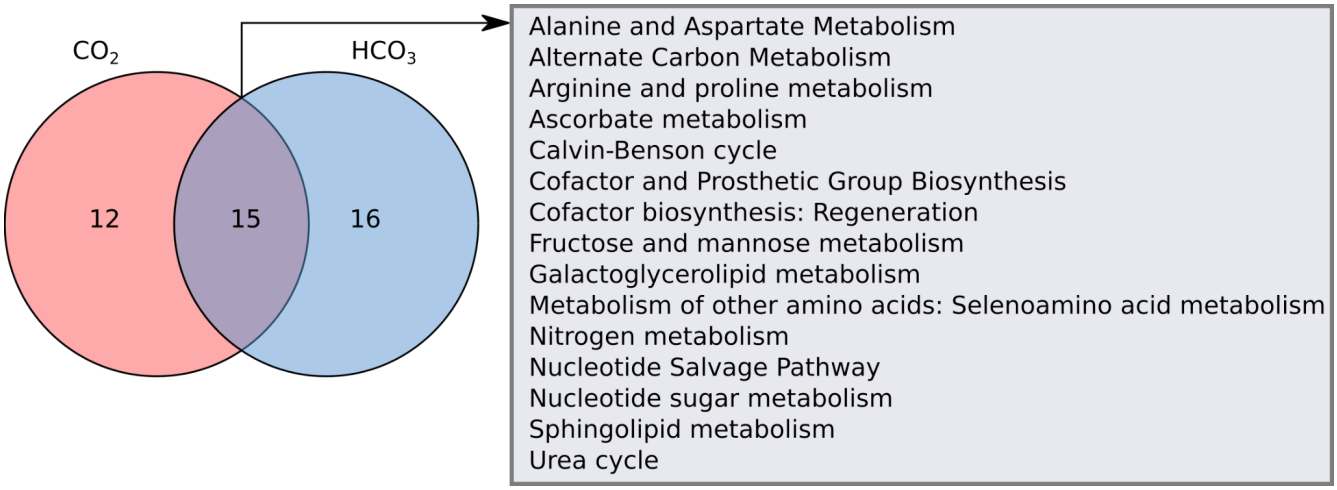

fig. S5: Subsystems with low flux due to elevated concentration of carbon sources. This Venn diagram represents shared and unique subsystems of *C. closterium* with low flux distributions under the elevated concentration of two different carbon sources (CO2 and HCO3). table S17 can be referred for more details.

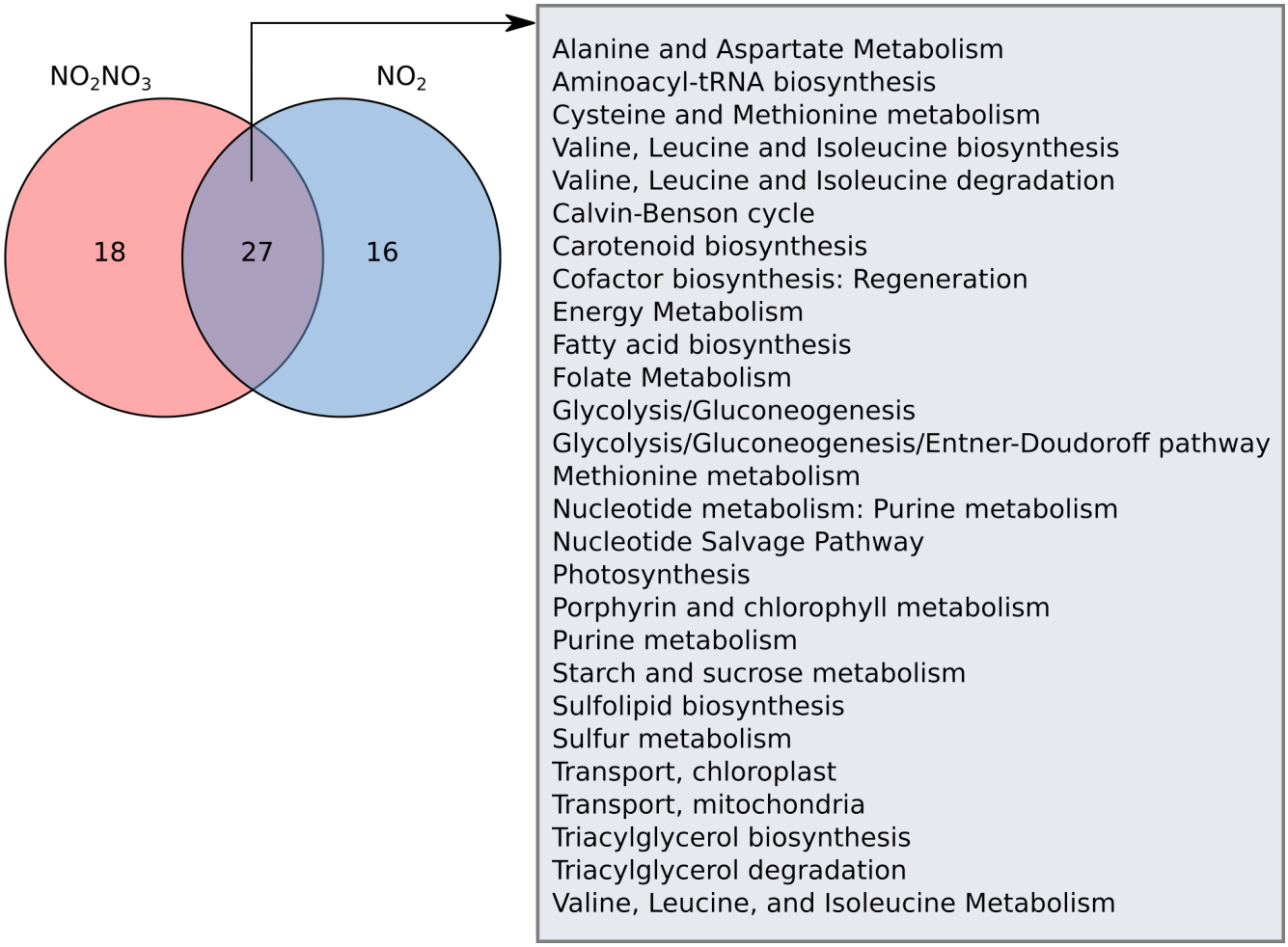

fig. S6: Subsystems with high flux due to elevated concentration of nitrogen sources. This Venn diagram represents shared and unique subsystems of *C. closterium* with high flux distributions under the elevated concentration of two different nitrogen sources (NO_2NO_3 and NO_2). table S18 can be referred for more details.

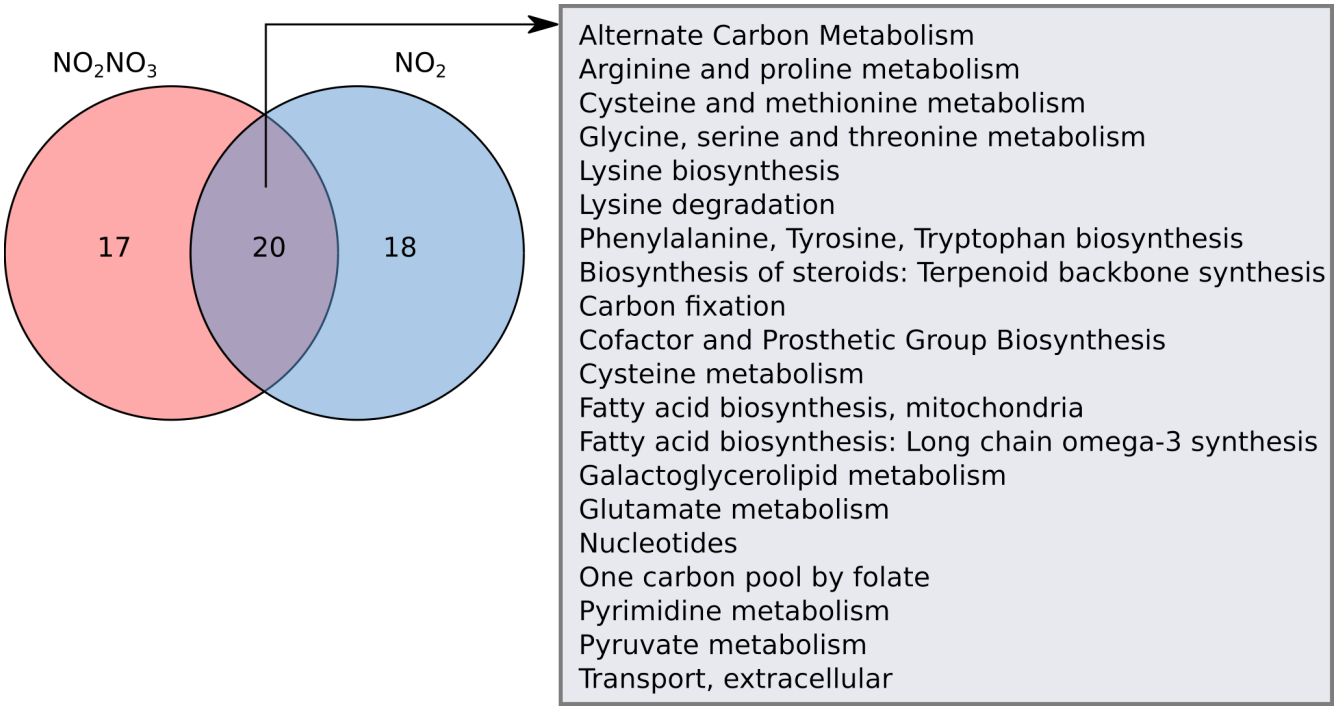

fig. S7: Subsystems with low flux due to elevated concentration of nitrogen sources. This Venn diagram represents shared and unique subsystems of *C. closterium* with low flux distributions under the elevated concentration of two different nitrogen sources (NO_2NO_3 and NO_2). Supplementary Table 17 can be referred for more details.

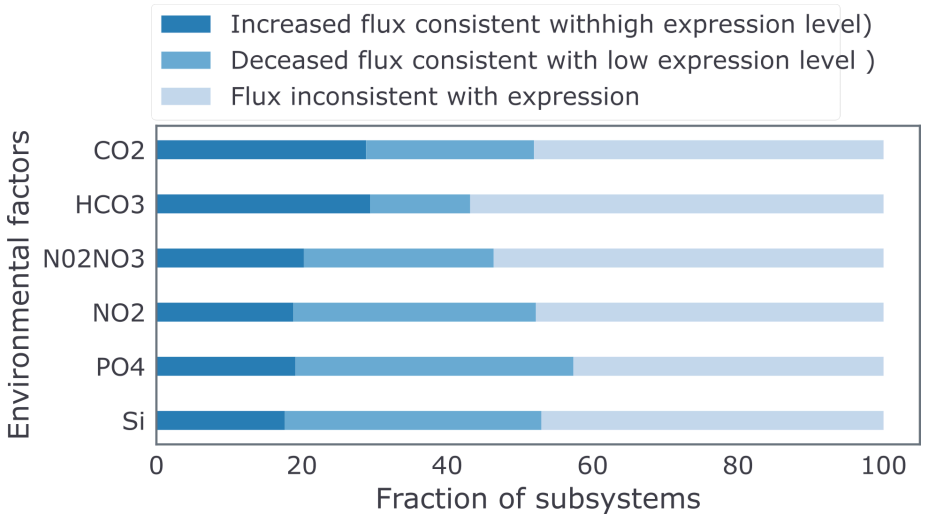

fig. S8: Intersection between differential model-predicted fluxes and gene expressions. Barplot represents the percentage of overlapped and inconsistent subsystems that carry differential (increased and decreased) fluxes and are associated with differentially expressed genes under elevated concentrations of nutrients such as CO_2, HCO_3, NO_2, $NO_2 + NO_3$, PO_4, and Si. While analyzing differential flux distributions and gene expression, we noticed that several gene/reaction pairs do not correlate. However, it was not surprising because it has been demonstrated that gene expression level can be a weak measure of enzyme activity in a given reaction (https://doi.org/10.1038/s41598-021-88129-3; http://dx.doi.org/10.1016/j.cels.2016.08.013; https://doi.org/10.1371/journal.pcbi.1003580). Moreover, these discrepancies can be because of the involvement of different strains of *C. closterium* in metabolic model generation and metatranscriptomics data analysis of ocean samples. Despite these challenges, model-predicted fluxes and gene expression levels were consistent for more than 50% of metabolic reactions, which was able to explain the detailed responses of *C. closterium* metabolism to drastically different nutrition conditions in the marine environment.

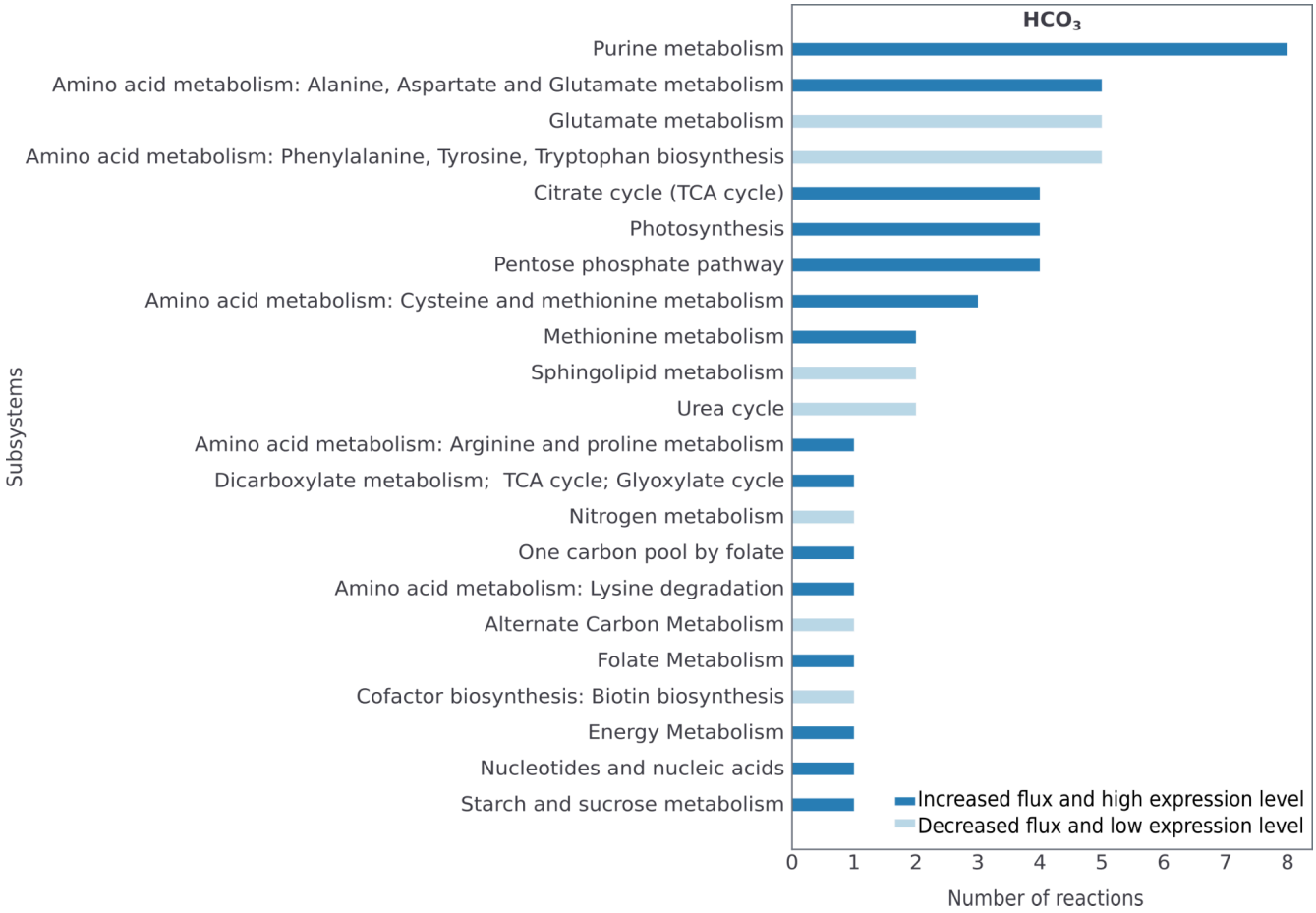

fig. S9: Comparison between reaction fluxes and gene expressions under increased HCO_3 concentration. This bar plot represents number of reactions in subsystems that showed consistency between differential (increased and decreased) model-predicted fluxes of reactions and differentially expressed genes under elevated HCO_3 condition.

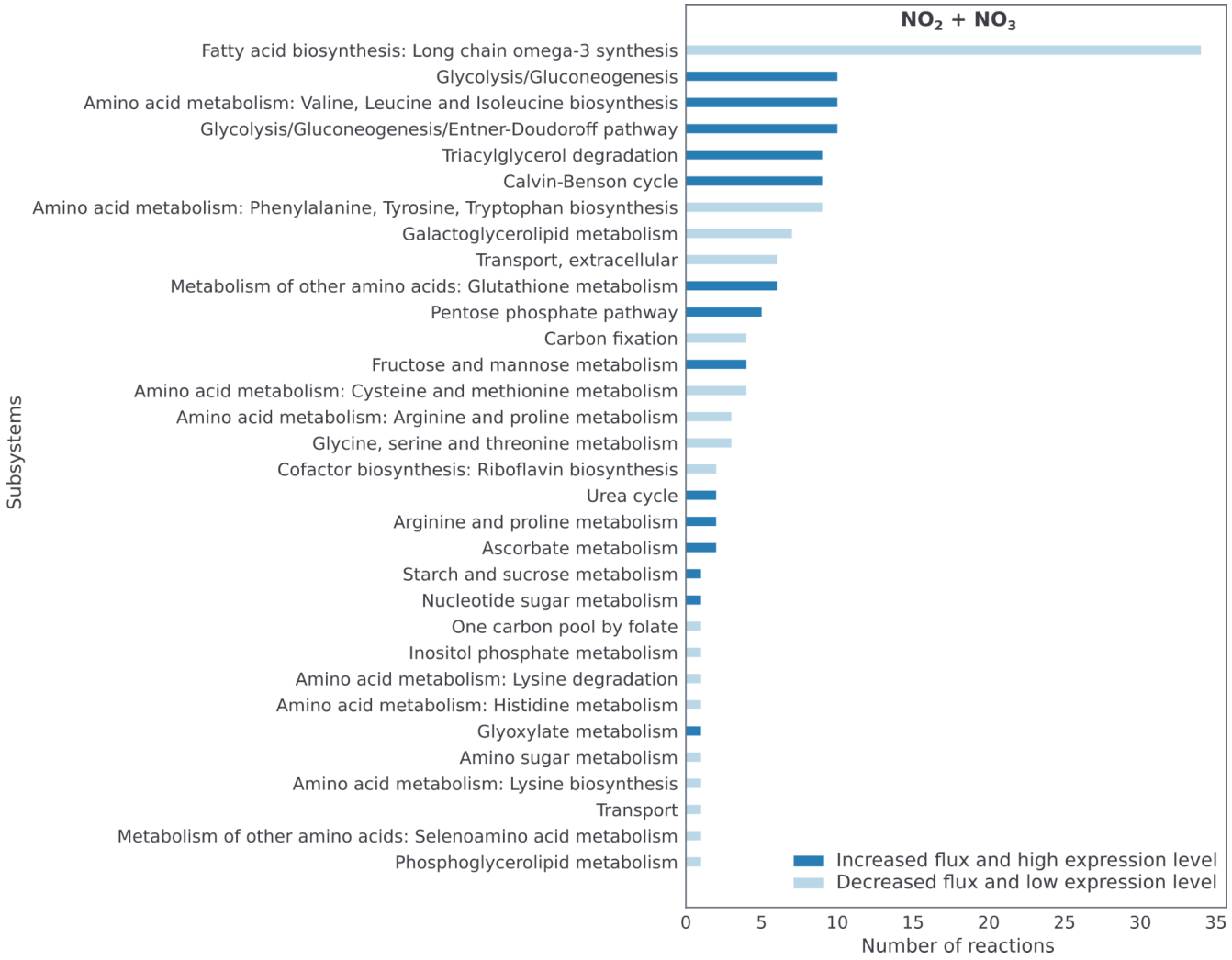

fig. S10: Comparison between reaction fluxes and gene expressions under increased $NO_2 + NO_3$ concentration. This bar plot represents number of reactions in subsystems that showed consistency between differential (increased and decreased) model-predicted fluxes of reactions and differentially expressed genes under elevated $NO_2 + NO_3$ condition.

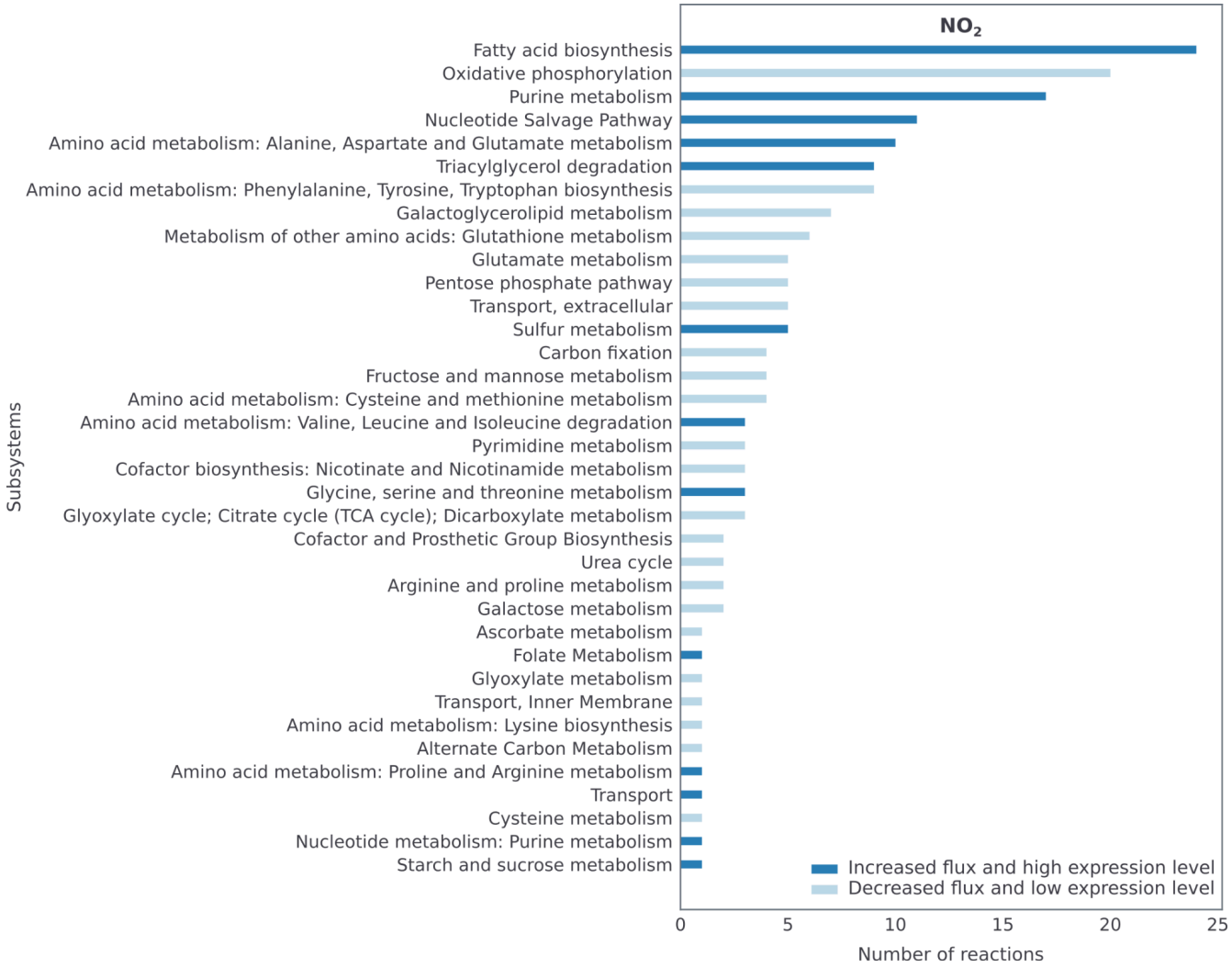

fig. S11: Comparison between reaction fluxes and gene expressions under increased NO_2 concentration. This bar plot number of reactions in subsystems that showed consistency between differential (increased and decreased) model-predicted fluxes of reactions and differentially expressed genes under elevated NO_2 condition.

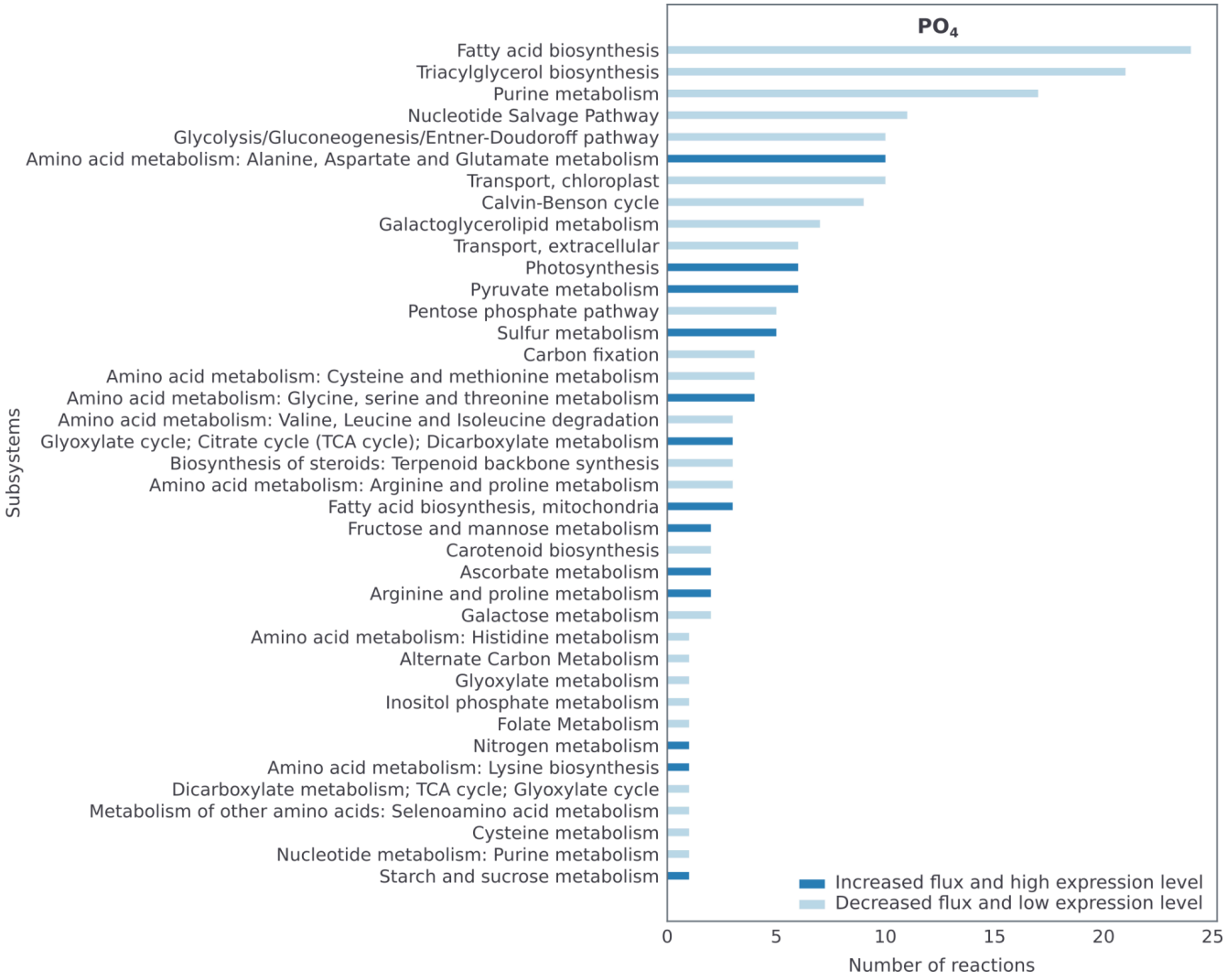

fig. S12: Comparison between reaction fluxes and gene expressions under increased PO_4 concentration. This bar plot number of reactions in subsystems that showed consistency between differential (increased and decreased) model-predicted fluxes of reactions and differentially expressed genes under elevated PO_4 condition.

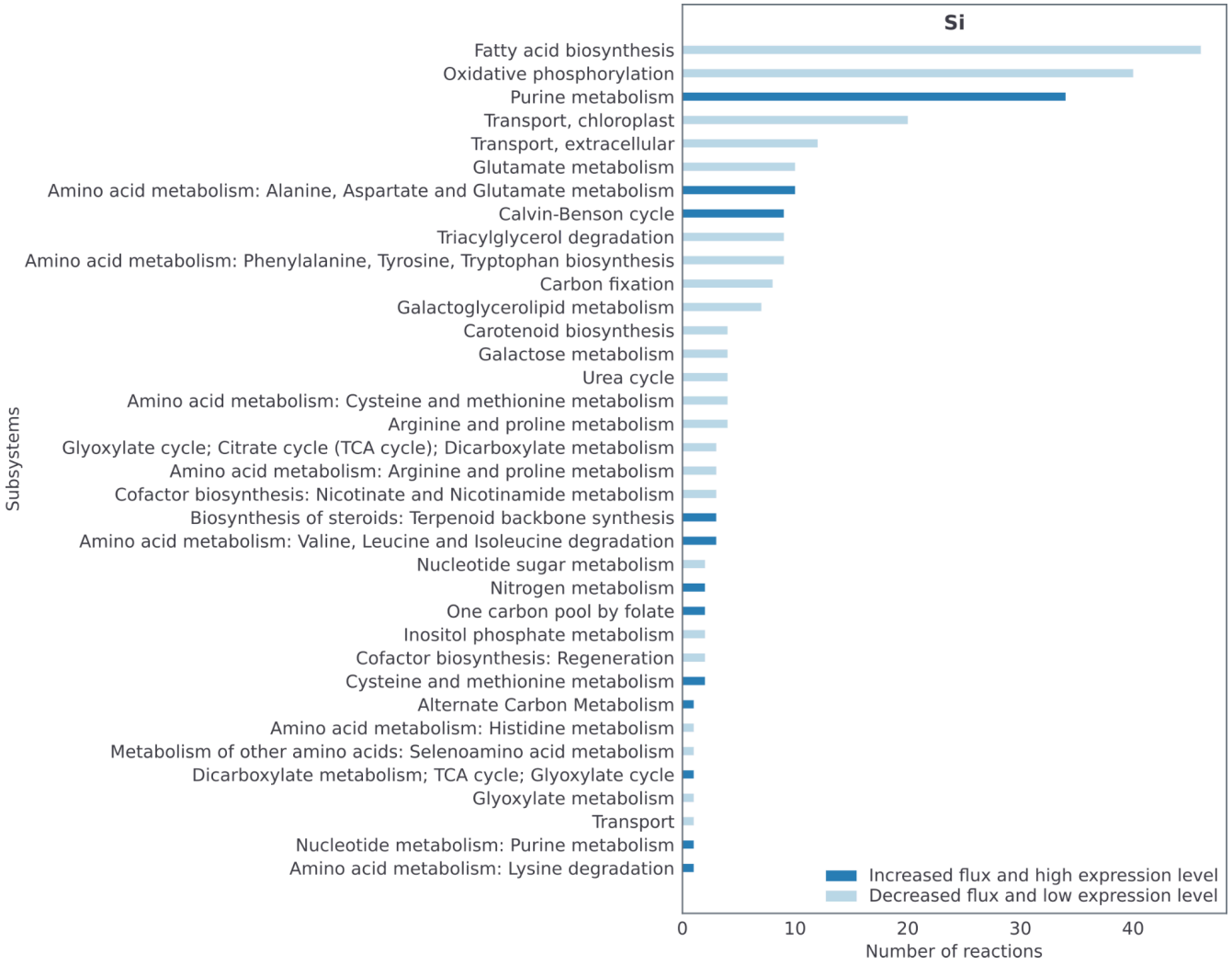

fig. S13: Comparison between reaction fluxes and gene expressions under increased Si concentration. This bar plot number of reactions in subsystems that showed consistency between differential (increased and decreased) model-predicted fluxes of reactions and differentially expressed genes under elevated Si condition.

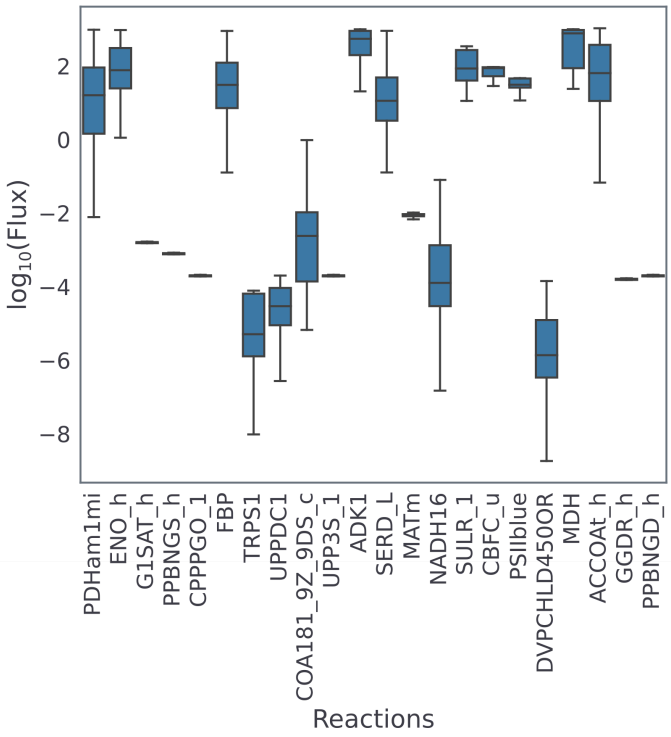

fig. S14: Photoautotrophy-specific reactions. The unique set of active reactions when iMK1961 was simulated under photoautotrophic condition. table S21 can be seen for more details about these reactions.

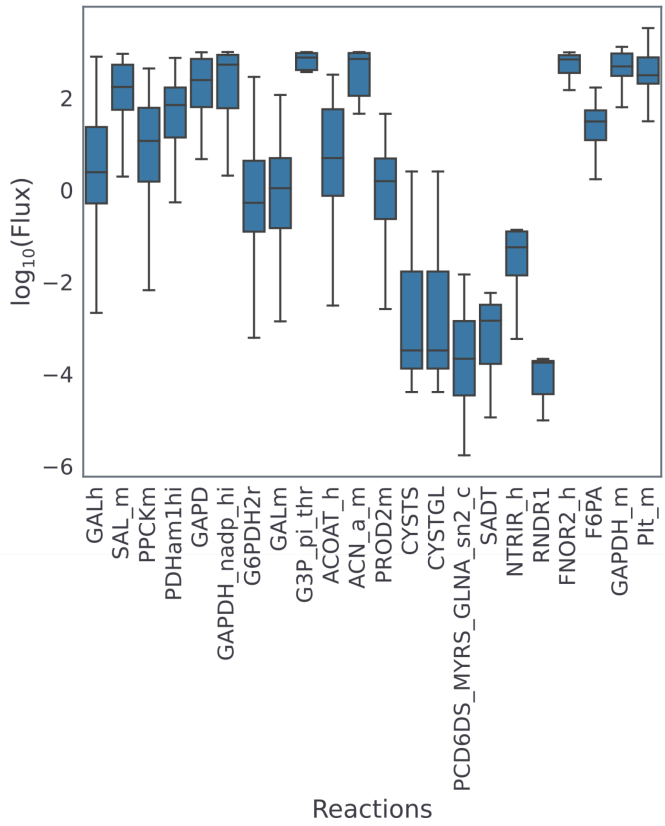

fig. S15: Heterotrophy-specific reactions. The unique set of active reactions when iMK1961 was simulated under heterotrophic condition. table S21 can be seen for more details about these reactions.

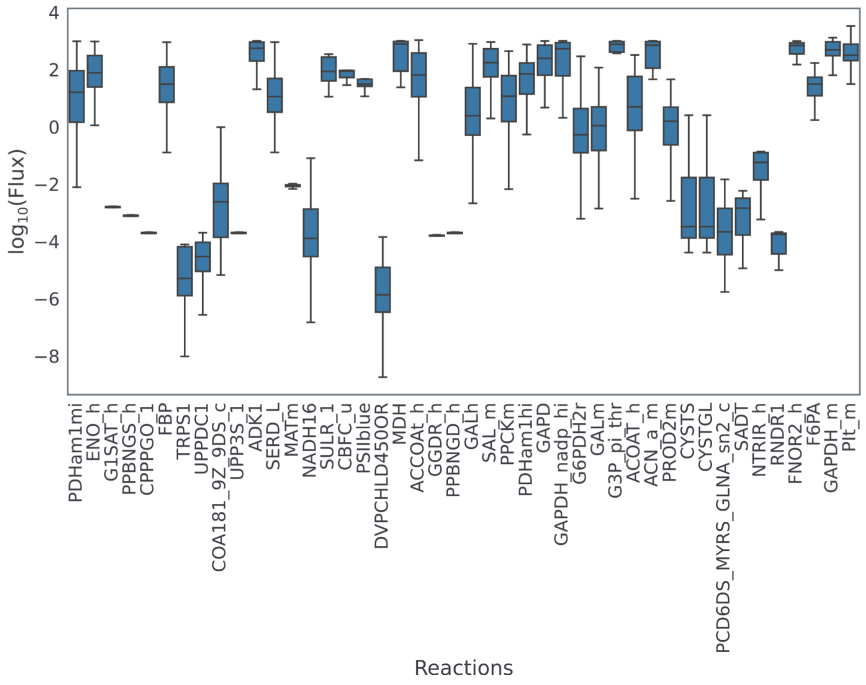

fig. S16: Mixotrophy-specific reactions. The unique set of reactions under photoautotrophic and heterotrophic conditions were found active together when iMK1961 was simulated under a mixotrophic condition. table S21 can be seen for more details about these reactions.

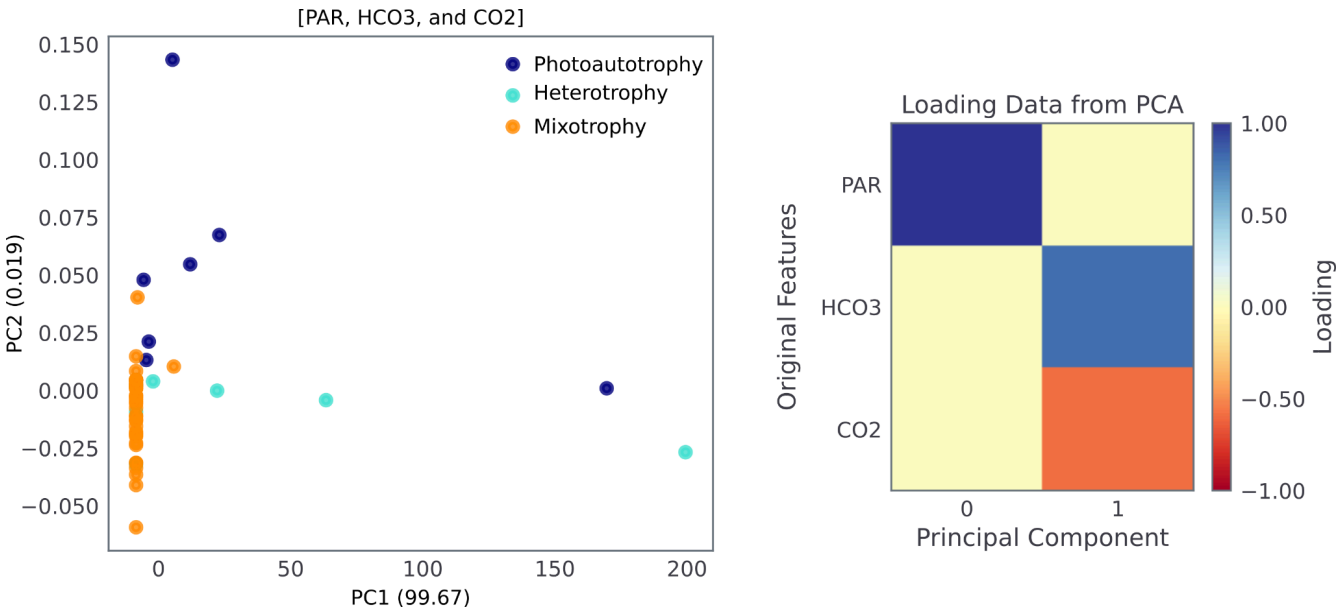

fig. S17: Principal Component Analysis (PCA) based on the CO_2/HCO_3/PAR (Photosynthetically active radiation) levels in global oceans. Each point represents a *Tara* Oceans sample. The values of environmental factors were used from a publication by Salazar et al. *(20)*. This figure is related to Figure 6c. The heatmap illustrates the PCA loadings of the environmental factors for the main principal components, indicating how these factors contribute to the clustering observed in the data.

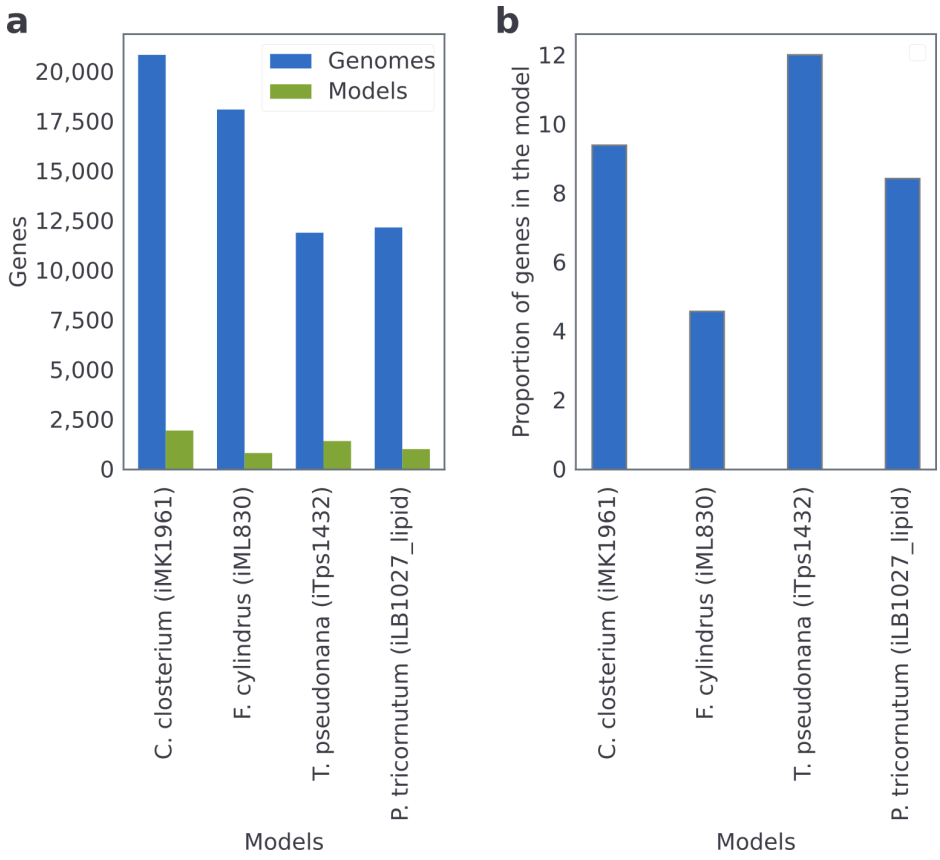

fig. S18: Comparison of gene representation in diatom models and their genome. (a) Total number of genes in diatom models and their respective genomes. (b) Proportion of genes represented in diatom models relative to the total genes in their genomes. iMK1961 exhibits a higher proportion of genes (9.4%) compared to the average proportion (8.6%) across all models. Among all diatom models, iTps1432, which contains fewer genes and reactions than iMK1961, demonstrates the highest proportion (12%) of genes represented in its genome.

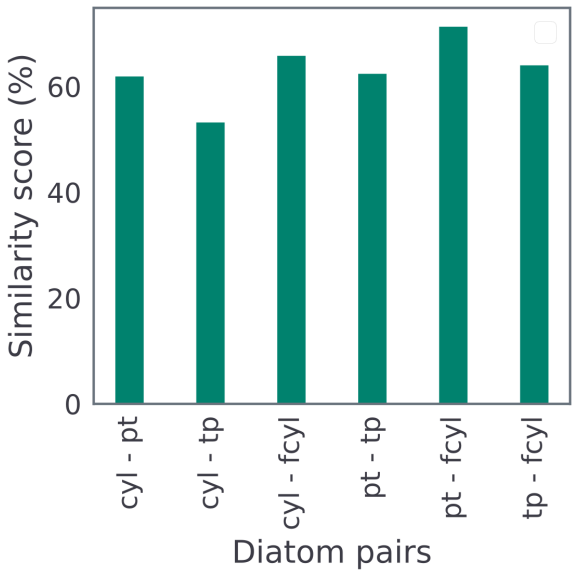

fig. S19: Homology analysis based on four diatom genomes. Similarity score (%) was obtain using BLAST (E-value < 1e-6) for protein sequences of diatom pairs: *C. closterium* - *P. tricornutum* (cyl - pt), *C. closterium* - *T. pseudonana* (cyl - tp), *C. closterium* - *F. cylindrus* (cyl - Fcyl), *P. tricornutum* - *T. pseudonana* (pt - tp), *P. tricornutum* - *F. cylindrus* (pt - fcyl), and *T. pseudonana* - *F. cylindrus* (tp - fcyl). The homology between these four diatoms ranges from 53% to 71%.

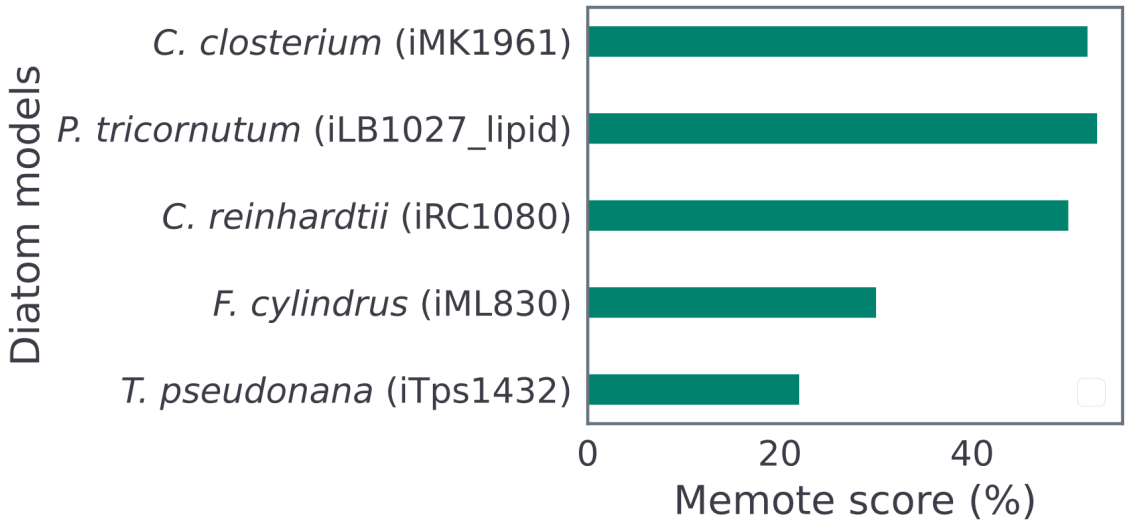

fig. S20: Model quality assessment using the Memote Test Suite. The Memote score of iMK1961 (C. closterium) were compared with three diatoms (*P. tricornutum*, *F. cylindrus*, and *T. pseudonana*) and a green microalgae (*C. reinhardtii*) for model quality evaluation. Among the compared models, iMK1961 demonstrated an equivalent Memote score to two models (iLB1027_lipid and iRC1080) obtained from BiGG models database. Additionally, iMK1961 exhibited a higher score compared to the other two diatoms (iML830 and iTps1432).

table S1: List of the articles used in the model reconstruction and validation processes

table S2: Precursors involved in biomass production

table S3: In silico media representing autotrophic growth condition

table S4: In silico media representing heterotrophic growth condition

table S5: In silico media representing mixotrophic growth condition

table S6: Reactions added to the model during the gap-filling process

table S7: Nutrients (carbon, nitrogen, phosphorus, and sulfur sources) used to test the prediction capabilities of the model. The references provided indicate whether the growth data was sourced from previously published reports or generated during the course of this study.

table S8: Comparing iMK1961 predictions with experimental measurements. This data is related to Fig. 3.

table S9: Uptake fluxes of nutrients used to constrain iMK1961 to simulate the growth of *C. closterium* under different growth conditions

table S10: Measurements of environmental factors global Tara Oceans metatranscriptomics data

table S11: Uptake flux of nutrients used to mimic the variable environmental conditions for simulating the *C. closterium* model

table S12: Reactions added to the model based on EC numbers

table S13: Model features (reactions)

table S14: Model features (metabolites)

table S15: Unique metabolites in the model

table S16: Subsystems with differential flux under elevated concentrations of nutrients (CO2, HCO3, NO2, NO2 + NO3, PO4, and Si)

table S17: Shared and common subsystems with low flux compared between different conditions. This table represents the subsystems while comparing all conditions as well as based on pairwise comparisons.

table S18: Shared and common subsystems with high flux compared between different conditions. This table represents the subsystems while comparing all conditions as well as based on pairwise comparisons.

table S19: Binary representation of subsystems with differential flux under elevated concentrations of nutrients (CO2, HCO3, NO2, NO2 + NO3, PO4, and Si)

table S20: Reactions associated with up and downregulated subsystems due to elevated CO_2 levels

table S21: Condition-specific active reactions in photoautotrophic, heterotrophic, and mixotrophic growth of *C. closterium*. Here, active reactions were defined based on non-zero values of fluxes of reactions and expression levels of associated genes under different conditions.

table S22: Model-predicted photoautotrophy-specific reactions based on only non-zero flux values.

table S23: Model-predicted heterotrophy-specific reactions based on only non-zero flux values

table S24: Model-predicted mixotrophy-specific reactions based on only non-zero flux values

table S25: Matrix used to plot fig. S3

table S26: Differentially abundant marine prokaryotes between different trophic modes using Tara Oceans metagenomic data. This data is related to Fig. 6E.

Source data: This file contains the source data for all figures, including both the main and supplementary figures.

REFERENCES AND NOTES

1. Y. Matsuda, P. G. Kroth, “Carbon fixation in diatoms” in *The Structural Basis of Biological Energy Generation* (Springer, 2014; https://link.springer.com/10.1007/978-94-017-8742-0_18), pp. 335–362.

2. C. Büchel, How diatoms harvest light. *Science* **365**, 447–448 (2019).

3. M. D. Guiry, How many species of algae are there? *J. Phycol.* **48**, 1057–1063 (2012).

4. E. De Tommasi, J. Gielis, A. Rogato, Diatom frustule morphogenesis and function: A multidisciplinary survey. *Mar. Genomics* **35**, 1–18 (2017).

5. S. Wang, S. K. Verma, I. Hakeem Said, L. Thomsen, M. S. Ullrich, N. Kuhnert, Changes in the fucoxanthin production and protein profiles in *Cylindrotheca closterium* in response to blue light-emitting diode light. *Microb. Cell Fact.* **17**, 110 (2018).

6. N. R. Cohen, Mixotrophic plankton foraging behaviour linked to carbon export. *Nat. Commun.* **13**, 1302 (2022).

7. B. A. Ward, M. J. Follows, Marine mixotrophy increases trophic transfer efficiency, mean organism size, and vertical carbon flux. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2958–2963 (2016).

8. S. Malviya, E. Scalco, S. Audic, F. Vincent, A. Veluchamy, J. Poulain, P. Wincker, D. Iudicone, C. de Vargas, L. Bittner, A. Zingone, C. Bowler, Insights into global diatom distribution and diversity in the world’s ocean. *Proc. Natl. Acad. Sci. U. S.A.* **113**, E1516–E1525 (2016).

9. H. Wang, J. L. Robinson, P. Kocabas, J. Gustafsson, M. Anton, P.-E. Cholley, S. Huang, J. Gobom, T. Svensson, M. Uhlen, H. Zetterberg, J. Nielsen, Genome-scale metabolic network reconstruction of model animals as a platform for translational research. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2102344118 (2021).

10. J. E. Yang, S. J. Park, W. J. Kim, H. J. Kim, B. J. Kim, H. Lee, J. Shin, S. Y. Lee, One-step fermentative production of aromatic polyesters from glucose by metabolically engineered *Escherichia coli* strains. *Nat. Commun.* **9**, 79 (2018).

11. J. M. Monk, C. J. Lloyd, E. Brunk, N. Mih, A. Sastry, Z. King, R. Takeuchi, W. Nomura, Z. Zhang, H. Mori, A. M. Feist, B. O. Palsson, iML1515, a knowledgebase that computes *Escherichia coli* traits. *Nat. Biotechnol.* **35**, 904–908 (2017).

12. J. M. Monk, P. Charusanti, R. K. Aziz, J. A. Lerman, N. Premyodhin, J. D. Orth, A. M. Feist, B. Ø. Palsson, Genome-scale metabolic reconstructions of multiple *Escherichia coli* strains highlight strain-specific adaptations to nutritional environments. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20338–20343 (2013).

13. M. Kumar, B. Ji, K. Zengler, J. Nielsen, Modelling approaches for studying the microbiome. *Nat. Microbiol.* **4**, 1253–1267 (2019).

14. M. Kumar, B. Ji, P. Babaei, P. Das, D. Lappa, G. Ramakrishnan, T. E. Fox, R. Haque, W. A. Petri, F. Bäckhed, J. Nielsen, Gut microbiota dysbiosis is associated with malnutrition and reduced plasma amino acid levels: Lessons from genome-scale metabolic modeling. *Metab. Eng.* **49**, 128–142 (2018).

15. C. Zuñiga, T. Li, M. T. Guarnieri, J. P. Jenkins, C.-T. Li, K. Bingol, Y.-M. Kim, M. J. Betenbaugh, K. Zengler, Synthetic microbial communities of heterotrophs and phototrophs facilitate sustainable growth. *Nat. Commun.* **11**, 3803 (2020).

16. J. Levering, J. Broddrick, C. L. Dupont, G. Peers, K. Beeri, J. Mayers, A. A. Gallina, A. E. Allen, B. O. Palsson, K. Zengler, Genome-scale model reveals metabolic basis of biomass partitioning in a model diatom. *PLOS ONE* **11**, e0155038 (2016).

17. M. Lavoie, B. Saint-Béat, J. Strauss, S. Guérin, A. Allard, S. V. Hardy, A. Falciatore, J. Lavaud, Genome-scale metabolic reconstruction and in silico perturbation analysis of the polar diatom *Fragilariopsis cylindrus* predicts high metabolic robustness. *Biology (Basel).* **9**, 30 (2020).

18. H. M. van Tol, E. V. Armbrust, Genome-scale metabolic model of the diatom *Thalassiosira pseudonana* highlights the importance of nitrogen and sulfur metabolism in redox balance. *PLOS ONE* **16**, e0241960 (2021).

19. A. Ahmad, A. Tiwari, S. Srivastava, A genome-scale metabolic model of *Thalassiosira pseudonana* CCMP 1335 for a systems-level understanding of its metabolism and biotechnological potential. *Microorganisms* **8**, 1396 (2020).

20. G. Salazar, L. Paoli, A. Alberti, J. Huerta-Cepas, H.-J. Ruscheweyh, M. Cuenca, C. M. Field, L. P. Coelho, C. Cruaud, S. Engelen, A. C. Gregory, K. Labadie, C. Marec, E. Pelletier, M. Royo-Llonch, S. Roux, P. Sánchez, H. Uehara, A. A. Zayed, G. Zeller, M. Carmichael, C. Dimier, J. Ferland, S. Kandels, M. Picheral, S. Pisarev, J. Poulain, S. G. Acinas, M. Babin, P. Bork, C. Bowler, C. de Vargas, L. Guidi, P. Hingamp, D. Iudicone, L. Karp-Boss, E. Karsenti, H. Ogata, S. Pesant, S. Speich, M. B. Sullivan, P. Wincker, S. Sunagawa, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels-Lewis, L. Karp-Boss, E. Karsenti, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, P. Wincker, Gene expression changes and community turnover differentially shape the global ocean metatranscriptome. *Cell* **179**, 1068–1083.e21 (2019).

21. C. de Vargas, S. Audic, N. Henry, J. Decelle, F. Mahé, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horák, O. Jaillon, G. Lima-Mendez, J. Lukeš, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, S. G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M. E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, E. Karsenti, E. Boss, M. Follows, L. Karp-Boss, U. Krzic, E. G. Reynaud, C. Sardet, M. B. Sullivan, D. Velayoudon, Ocean plankton. Eukaryotic plankton diversity in the sunlit ocean. *Science* **348**, 1261605 (2015).

22. R. F. Strzepek, B. L. Nunn, L. T. Bach, J. A. Berges, E. B. Young, P. W. Boyd, The ongoing need for rates: Can physiology and omics come together to co-design the measurements needed to understand complex ocean biogeochemistry? *J. Plankton Res.* **44**, 485–495 (2022).

23. C. Zuñiga, C.-T. Li, G. Yu, M. M. Al-Bassam, T. Li, L. Jiang, L. S. Zaramela, M. Guarnieri, M. J. Betenbaugh, K. Zengler, Environmental stimuli drive a transition from cooperation to competition in synthetic phototrophic communities. *Nat. Microbiol.* **4**, 2184–2191 (2019).

24. H. Wang, S. Marcišauskas, B. J. Sánchez, I. Domenzain, D. Hermansson, R. Agren, J. Nielsen, E. J. Kerkhoven, RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. *PLoS Comput. Biol.* **14**, e1006541 (2018).

25. A. Ebrahim, J. A. Lerman, B. O. Palsson, D. R. Hyduke, COBRApy: COnstraints-based reconstruction and analysis for Python. *BMC Syst. Biol.* **7**, 74 (2013).

26. L. Heirendt, S. Arreckx, T. Pfau, S. N. Mendoza, A. Richelle, A. Heinken, H. S. Haraldsdóttir, J. Wachowiak, S. M. Keating, V. Vlasov, S. Magnusdóttir, C. Y. Ng, G. Preciat, A. Žagare, S. H. J. Chan, M. K. Aurich, C. M. Clancy, J. Modamio, J. T. Sauls, A. Noronha, A. Bordbar, B. Cousins, D. C. El Assal, L. V. Valcarcel, I. Apaolaza, S. Ghaderi, M. Ahookhosh, M. Ben Guebila, A. Kostromins, N. Sompairac, H. M. Le, D. Ma, Y. Sun, L. Wang, J. T. Yurkovich, M. A. P. Oliveira, P. T. Vuong, L. P. El Assal, I. Kuperstein, A. Zinovyev, H. S. Hinton, W. A. Bryant, F. J. Aragón Artacho, F. J. Planes, E. Stalidzans, A. Maass, S. Vempala, M. Hucka, M. A. Saunders, C. D. Maranas, N. E. Lewis, T. Sauter, B. Ø. Palsson, I. Thiele, R. M. T. Fleming, Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat. Protoc.* **14**, 639–702 (2019).

27. R. L. Chang, L. Ghamsari, A. Manichaikul, E. F. Y. Hom, S. Balaji, W. Fu, Y. Shen, T. Hao, B. Ø. Palsson, K. Salehi-Ashtiani, J. A. Papin, Metabolic network reconstruction of *Chlamydomonas* offers insight into light-driven algal metabolism. *Mol. Syst. Biol.* **7**, 518 (2011).

28. Z. A. King, J. Lu, A. Dräger, P. Miller, S. Federowicz, J. A. Lerman, A. Ebrahim, B. O. Palsson, N. E. Lewis, BiGG models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Res.* **44**, D515–D522 (2016).

29. S. Moretti, V. D. T. Tran, F. Mehl, M. Ibberson, M. Pagni, MetaNetX/MNXref: Unified namespace for metabolites and biochemical reactions in the context of metabolic models. *Nucleic Acids Res.* **49**, D570–D574 (2021).

30. A. M. Elagoz, L. Ambrosino, C. Lauritano, *De novo* transcriptome of the diatom *Cylindrotheca closterium* identifies genes involved in the metabolism of anti-inflammatory compounds. *Sci. Rep.* **10**, 4138 (2020).

31. W. Megchelenbrink, M. Huynen, E. Marchiori, optGpSampler: An improved tool for uniformly sampling the solution-space of genome-scale metabolic networks. *PLOS ONE* **9**, e86587 (2014).

32. S. Sunagawa, S. G. Acinas, P. Bork, C. Bowler, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels, L. Karp-Boss, E. Karsenti, M. Lescot, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, M. Sieracki, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, P. Wincker, D. Eveillard, G. Gorsky, L. Guidi, D. Iudicone, E. Karsenti, F. Lombard, H. Ogata, S. Pesant, M. B. Sullivan, P. Wincker, C. de Vargas, Tara Oceans: Towards global ocean ecosystems biology. *Nat. Rev. Microbiol.* **18**, 428–445 (2020).

33. E. L. Jensen, K. Yangüez, F. Carrière, B. Gontero, Storage compound accumulation in diatoms as response to elevated CO_2 concentration. *Biology (Basel)* **9**, 5 (2019).

34. S. Wu, W. Gu, S. Jia, L. Wang, L. Wang, X. Liu, L. Zhou, A. Huang, G. Wang, Proteomic and biochemical responses to different concentrations of CO_2 suggest the existence of multiple carbon metabolism strategies *in Phaeodactylum tricornutum*. *Biotechnol. Biofuels* **14**, 235 (2021).

35. K. Nakajima, A. Tanaka, Y. Matsuda, SLC4 family transporters in a marine diatom directly pump bicarbonate from seawater. *Proc. Natl. Acad. Sci. U.S.A* **110**, 1767–1772 (2013).

36. N. Yodsuwan, S. Sawayama, S. Sirisansaneeyakul, Effect of nitrogen concentration on growth, lipid production and fatty acid profiles of the marine diatom *Phaeodactylum tricornutum. Agric. Nat. Resour.* **51**, 190–197 (2017).

37. Z.-K. Yang, Y.-H. Ma, J.-W. Zheng, W.-D. Yang, J.-S. Liu, H.-Y. Li, Proteomics to reveal metabolic network shifts towards lipid accumulation following nitrogen deprivation in the diatom *Phaeodactylum tricornutum*. *J. Appl. Phycol.* **26**, 73–82 (2014).

38. Y. Zhang, H. Wu, M. Sun, Q. Peng, A. Li, Photosynthetic physiological performance and proteomic profiling of the oleaginous algae *Scenedesmus acuminatus* reveal the mechanism of lipid accumulation under low and high nitrogen supplies. *Photosynth. Res.* **138**, 73–102 (2018).

39. N. L. Hockin, T. Mock, F. Mulholland, S. Kopriva, G. Malin, The response of diatom central carbon metabolism to nitrogen starvation is different from that of green algae and higher plants. *Plant Physiol.* **158**, 299–312 (2012).

40. J. Jian, D. Zeng, W. Wei, H. Lin, P. Li, W. Liu, The combination of RNA and protein profiling reveals the response to nitrogen depletion in *Thalassiosira pseudonana. Sci. Rep.* **7**, 8989 (2017).

41. J. Msanne, D. Xu, A. R. Konda, J. A. Casas-Mollano, T. Awada, E. B. Cahoon, H. Cerutti, Metabolic and gene expression changes triggered by nitrogen deprivation in the photoautotrophically grown microalgae *Chlamydomonas reinhardtii* and *Coccomyxa* sp. C-169. *Phytochemistry* **75**, 50–59 (2012).

42. L. Recht, A. Zarka, S. Boussiba, Patterns of carbohydrate and fatty acid changes under nitrogen starvation in the microalgae *Haematococcus pluvialis* and *Nannochloropsis* sp. *Appl. Microbiol. Biotechnol.* **94**, 1495–1503 (2012).

43. A. P. Dean, D. C. Sigee, B. Estrada, J. K. Pittman, Using FTIR spectroscopy for rapid determination of lipid accumulation in response to nitrogen limitation in freshwater microalgae. *Bioresour. Technol.* **101**, 4499–4507 (2010).

44. H. A. Qari, M. Oves, Fatty acid synthesis by *Chlamydomonas reinhardtii* in phosphorus limitation. *J. Bioenerg. Biomembr.* **52**, 27–38 (2020).

45. M. Chen, J. Li, X. Dai, Y. Sun, F. Chen, Effect of phosphorus and temperature on chlorophyll a contents and cell sizes of *Scenedesmus obliquus* and *Microcystis aeruginosa*. *Limnology* **12**, 187–192 (2011).

46. B. Zhou, J. Ma, F. Chen, Y. Zou, Y. Wei, H. Zhong, K. Pan, Mechanisms underlying silicon-dependent metal tolerance in the marine diatom *Phaeodactylum tricornutum*. *Environ. Pollut.* **262**, 114331 (2020).

47. J. Cheng, J. Feng, J. Sun, Y. Huang, J. Zhou, K. Cen, Enhancing the lipid content of the diatom *Nitzschia* sp. by 60 Co-γ irradiation mutation and high-salinity domestication. *Energy* **78**, 9–15 (2014).

48. G. d'Ippolito, A. Sardo, D. Paris, F. M. Vella, M. G. Adelfi, P. Botte, C. Gallo, A. Fontana, Potential of lipid metabolism in marine diatoms for biofuel production. *Biotechnol. Biofuels* **8**, 28 (2015).

49. E. Armstrong, A. Rogerson, J. W. Leftley, The abundance of heterotrophic protists associated with intertidal seaweeds. *Estuar. Coast. Shelf Sci.* **50**, 415–424 (2000).

50. V. Villanova, C. Spetea, Mixotrophy in diatoms: Molecular mechanism and industrial potential. *Physiol. Plant.* **173**, 603–611 (2021).

51. B. Bailleul, N. Berne, O. Murik, D. Petroutsos, J. Prihoda, A. Tanaka, V. Villanova, R. Bligny, S. Flori, D. Falconet, A. Krieger-Liszkay, S. Santabarbara, F. Rappaport, P. Joliot, L. Tirichine, P. G. Falkowski, P. Cardol, C. Bowler, G. Finazzi, Energetic coupling between plastids and mitochondria drives CO_2 assimilation in diatoms. *Nature* **524**, 366–369 (2015).

52. D. K. Stoecker, P. J. Hansen, D. A. Caron, A. Mitra, Mixotrophy in the marine plankton. *Ann. Rev. Mar. Sci.* **9**, 311–335 (2017).

53. N. Meyer, A. Rydzyk, G. Pohnert, Pronounced uptake and metabolism of organic substrates by diatoms revealed by pulse-labeling metabolomics. *Front. Mar. Sci.* **9**, 821167 (2022).

54. D. A. Caron, Mixotrophy stirs up our understanding of marine food webs. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2806–2808 (2016).

55. S. A. Amin, M. S. Parker, E. V. Armbrust, Interactions between diatoms and bacteria. *Microbiol. Mol. Biol. Rev.* **76**, 667–684 (2012).

56. J. Liu, C.-X. Xue, H. Sun, Y. Zheng, Z. Meng, X.-H. Zhang, Carbohydrate catabolic capability of a *Flavobacteriia* bacterium isolated from hadal water. *Syst. Appl. Microbiol.* **42**, 263–274 (2019).

57. N. Kallscheuer, S. Wiegand, T. Kohn, C. Boedeker, O. Jeske, P. Rast, R.-W. Müller, F. Brümmer, A. Heuer, M. S. M. Jetten, M. Rohde, M. Jogler, C. Jogler, Cultivation-independent analysis of the bacterial community associated with the calcareous sponge *Clathrina clathrus* and isolation of *Poriferisphaera corsica* Gen. Nov., Sp. Nov., belonging to the barely studied class *Phycisphaerae* in the Phylum *Planctomycetes*. *Front. Microbiol.* **11**, 602250 (2020).

58. Y. Fukunaga, M. Kurahashi, Y. Sakiyama, M. Ohuchi, A. Yokota, S. Harayama, *Phycisphaera mikurensis* gen. nov., sp. nov., isolated from a marine alga, and proposal of *Phycisphaeraceae* fam. nov., *Phycisphaerales* ord. nov. and *Phycisphaerae* classis nov. in the phylum *Planctomycetes. J. Gen. Appl. Microbiol.* **55**, 267–275 (2009).

59. T. Priest, A. Heins, J. Harder, R. Amann, B. M. Fuchs, Niche partitioning of the ubiquitous and ecologically relevant NS5 marine group. *ISME J.* **16**, 1570–1582 (2022).

60. A. Buchan, G. R. LeCleir, C. A. Gulvik, J. M. González, Master recyclers: Features and functions of bacteria associated with phytoplankton blooms. *Nat. Rev. Microbiol.* **12**, 686–698 (2014).

61. I. Thiele, B. Ø. Palsson, A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* **5**, 93–121 (2010).

62. M. Kumar, C. Zuniga, J. D. Tibocha-Bonilla, S. R. Smith, J. Coker, A. E. Allen, K. Zengler, "Constraint-based modeling of diatoms metabolism and quantitative biology approaches" in *The Molecular Life of Diatoms* (Springer International Publishing, 2022; https://link.springer.com/10.1007/978-3-030-92499-7_26), pp. 775–808.

63. W. Stock, B. Vanelslander, F. Rüdiger, K. Sabbe, W. Vyverman, U. Karsten, Thermal niche differentiation in the benthic diatom *Cylindrotheca closterium* (Bacillariophyceae) complex. *Front. Microbiol.* **10**, 1395 (2019).

64. D. A. Hutchins, F. Fu, Microorganisms and ocean global change. *Nat. Microbiol.* **2**, 17058 (2017).

65. S. L. Pahl, D. M. Lewis, F. Chen, K. D. King, Growth dynamics and the proximate biochemical composition and fatty acid profile of the heterotrophically grown diatom *Cyclotella cryptica*. *J. Appl. Phycol.* **22**, 165–171 (2010).

66. E. Armstrong, A. Rogerson, J. W. Leftley, Utilisation of seaweed carbon by three surface-associated heterotrophic protists, *Stereomyxa ramosa*, *Nitzschia alba* and *Labyrinthula* sp. *Aquat. Microb. Ecol.* **21**, 49–57 (2000).

67. M. A. Brzezinski, The Si:C:N ratio of marine diatoms: Interspecific variability and the effect of some environmental variables. *J. Phycol.* **21**, 347–357 (2004).

68. J.-C. Lachance, C. J. Lloyd, J. M. Monk, L. Yang, A. V. Sastry, Y. Seif, B. O. Palsson, S. Rodrigue, A. M. Feist, Z. A. King, P.-É. Jacques, BOFdat: Generating biomass objective functions for genome-scale metabolic models from experimental data. *PLoS Comput. Biol.* **15**, e1006971 (2019).

69. D. Jallet, M. A. Caballero, A. A. Gallina, M. Youngblood, G. Peers, Photosynthetic physiology and biomass partitioning in the model diatom *Phaeodactylum tricornutum* grown in a sinusoidal light regime. *Algal Res.* **18**, 51–60 (2016).

70. E. Cunha, D. Lagoa, J. P. Faria, F. Liu, C. S. Henry, O. Dias, TranSyT, an innovative framework for identifying transport systems. *Bioinformatics* **39**, btad466 (2023).

71. C. Lieven, M. E. Beber, B. G. Olivier, F. T. Bergmann, M. Ataman, P. Babaei, J. A. Bartell, L. M. Blank, S. Chauhan, K. Correia, C. Diener, A. Dräger, B. E. Ebert, J. N. Edirisinghe, J. P. Faria, A. M. Feist, G. Fengos, R. M. T. Fleming, B. García-Jiménez, V. Hatzimanikatis, W. van Helvoirt, C. S. Henry, H. Hermjakob, M. J. Herrgård, A. Kaafarani, H. U. Kim, Z. King, S. Klamt, E. Klipp, J. J. Koehorst, M. König, M. Lakshmanan, D.-Y. Lee, S. Y. Lee, S. Lee, N. E. Lewis, F. Liu, H. Ma, D. Machado, R. Mahadevan, P. Maia, A. Mardinoglu, G. L. Medlock, J. M. Monk, J. Nielsen, L. K. Nielsen, J. Nogales, I. Nookaew, B. O. Palsson, J. A. Papin, K. R. Patil, M. Poolman, N. D. Price, O. Resendis-Antonio, A. Richelle, I. Rocha, B. J. Sánchez, P. J. Schaap, R. S. M. Sheriff, S. Shoaie, N. Sonnenschein, B. Teusink, P. Vilaça, J. O. Vik, J. A. H. Wodke, J. C. Xavier, Q. Yuan, M. Zakhartsev, C. Zhang, MEMOTE for standardized genome-scale metabolic model testing. *Nat. Biotechnol.* **38**, 272–276 (2020).

72. J. D. Orth, I. Thiele, B. Ø. Palsson, What is flux balance analysis? *Nat. Biotechnol.* **28**, 245–248 (2010).

73. F. M. M. Morel, J. G. Rueter, D. M. Anderson, R. R. L. Guillard, AQUIL: A chemically defined phytoplankton culture medium for trace metal studies. *J. Phycol.* **15**, 135–141 (1979).

74. D. Kim, B. Langmead, S. L. Salzberg, HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

75. H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

76. J. Caballero, A. F. A. Smit, L. Hood, G. Glusman, Realistic artificial DNA sequences as negative controls for computational genomics. *Nucleic Acids Res.* **42**, e99–e99 (2014).

77. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

78. B. G. Galuzzi, L. Milazzo, C. Damiani, *Best Practices in Flux Sampling of Constrained-Based Models* (Springer, 2023; https://link.springer.com/10.1007/978-3-031-25891-6_18), pp. 234–248.

79. N. Segata, J. Izard, L. Waldron, D. Gevers, L. Miropolsky, W. S. Garrett, C. Huttenhower, Metagenomic biomarker discovery and explanation. *Genome Biol.* **12**, R60 (2011).

80. J. Monk, J. Nogales, B. O. Palsson, Optimizing genome-scale network reconstructions. *Nat. Biotechnol.* **32**, 447–452 (2014).

81. E. P. Morris, J. C. Kromkamp, Influence of temperature on the relationship between oxygen- and fluorescence-based estimates of photosynthetic parameters in a marine benthic diatom *(Cylindrotheca closterium). Eur. J. Phycol.* **38**, 133–142 (2003).

82. T. Alcoverro, E. Conte, L. Mazzella, Production of mucilage by the adriatic epipelic diatom *Cylindrotheca closterium* (Bacillariophyceae) under nutrient limitation. *J. Phycol.* **36**, 1087–1095 (2000).

83. J. W. Rijstenbil, Effects of UVB radiation and salt stress on growth, pigments and antioxidative defence of the marine diatom *Cylindrotheca closterium*. *Mar. Ecol. Prog. Ser.* **254**, 37–48 (2003).

84. J. W. Rijstenbil, UV- and salinity-induced oxidative effects in the marine diatom *Cylindrotheca closterium* during simulated emersion. *Mar. Biol.* **147**, 1063–1073 (2005).

85. N. Staats, L. J. Stal, L. R. Mur, Exopolysaccharide production by the epipelic diatom *Cylindrotheca closterium:* Effects of nutrient conditions. *J. Exp. Mar. Bio. Ecol.* **249**, 13–27 (2000).

86. S. Wang, D. Sirbu, L. Thomsen, N. Kuhnert, M. S. Ullrich, C. Thomsen, Comparative lipidomic studies of *Scenedesmus* sp. (Chlorophyceae) and *Cylindrotheca closterium* (Bacillariophyceae) reveal their differences in lipid production under nitrogen starvation. *J. Phycol.* **55**, 1246–1257 (2019).

87. C. V. M. Araújo, F. R. Diz, L. M. Lubián, J. Blasco, I. Moreno-Garrido, Sensitivity of *Cylindrotheca closterium* to copper: Influence of three test endpoints and two test methods. *Sci. Total Environ.* **408**, 3696–3703 (2010).

88. M. D. Almeyda, P. G. Scodelaro Bilbao, C. A. Popovich, D. Constenla, P. I. Leonardi, Enhancement of polyunsaturated fatty acid production under low-temperature stress in *Cylindrotheca closterium*. *J. Appl. Phycol.* **32**, 989–1001 (2020).

89. S. A. Van Bergeijk, C. Van der Zee, L. J. Stal, Uptake and excretion of dimethylsulphoniopropionate is driven by salinity changes in the marine benthic diatom *Cylindrotheca closterium*. *Eur. J. Phycol.* **38**, 341–349 (2003).

90. A. Erdogan, Z. Demirel, M. C. Dalay, A. E. Eroglu, Fucoxanthin content of *Cylindrotheca closterium* and its oxidative stress mediated enhancement. *Turkish J. Fish. Aquat. Sci.* **16**, 499–506 (2016).

91. D. M. Orcutt, G. W. Patterson, Effect of light intensity upon lipid composition of *Nitzschia closterium (Cylindrotheca fusiformis). Lipids* **9**, 1000–1003 (1974).

92. A. Affan, S. Heo, Y. Jeon, J. Lee, Optimal growth conditions and antioxidative activities of *Cylindrotheca closterium* (Bacillariophyceae). *J. Phycol.* **45**, 1405–1415 (2009).

93. T. Brembu, A. Mühlroth, L. Alipanah, A. M. Bones, The effects of phosphorus limitation on carbon metabolism in diatoms. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160406 (2017).

94. N. Ruocco, G. Nuzzo, G. D'Ippolito, E. Manzo, A. Sardo, A. Ianora, G. Romano, A. Iuliano, V. Zupo, M. Costantini, A. Fontana, Lipoxygenase pathways in diatoms: Occurrence and correlation with grazer toxicity in four benthic species. *Mar. Drugs* **18**, 66 (2020).

95. V. I. Ryabushko, S. N. Zheleznova, M. V. Nekhoroshev, Effect of nitrogen on fucoxanthin accumulation in the diatom *Cylindrotheca closterium* (Ehrenb.) Reimann et Lewin. *Int. J. Algae* **19**, 79–84 (2017).

96. T. Han, P. Song, R. Shi, Z. Qi, J. Li, H. Huang, Optimal nutrient availability could alleviate diatom *Cylindrotheca closterium* fouling during seedling cultivation of *Sargassum hemiphyllum*. *Aquaculture* **552**, 738020 (2022).

97. A. E. Becker, D. Copplestone, Cadmium uptake from sediment by *Cylindrotheca closterium* and the effect of diatom presence on partitioning of cadmium between sediment and water: A laboratory study. *Limnol. Oceanogr.* **64**, 2550–2568 (2019).

98. G. Humphrey, D. Subba Rao, Photosynthetic rate of the marine diatom *Cylindrotheca closterium*. *Mar. Freshw. Res.* **18**, 123 (1967).

99. F. Zhang, J. Chi, Influences of nutritional conditions on degradation of dibutyl phthalate in coastal sediments with *Cylindrotheca closterium. Mar. Pollut. Bull.* **153**, 111021 (2020).

100. F. Roncarati, J. W. Rijstenbil, R. Pistocchi, Photosynthetic performance, oxidative damage and antioxidants in *Cylindrotheca closterium* in response to high irradiance, UVB radiation and salinity. *Mar. Biol.* **153**, 965–973 (2008).

101. K. Wolfstein, J. de Brouwer, L. Stal, Biochemical partitioning of photosynthetically fixed carbon by benthic diatoms during short-term incubations at different irradiances. *Mar. Ecol. Prog. Ser.* **245**, 21–31 (2002).

102. K. Wolfstein, L. Stal, Production of extracellular polymeric substances (EPS) by benthic diatoms: Effect of irradiance and temperature. *Mar. Ecol. Prog. Ser.* **236**, 13–22 (2002).

103. K. Suman, T. Kiran, U. K. Devi, N. S. Sarma, Culture medium optimization and lipid profiling of *Cylindrotheca,* a lipid- and polyunsaturated fatty acid-rich pennate diatom and potential source of eicosapentaenoic acid. *Bot. Mar.* **55**, 289–299 (2012).

104. G. J. C. Underwood, M. Boulcott, C. A. Raines, K. Waldron, Environmental effects on exopolymer production by marine benthic diatoms: Dynamics, changes in composition, and pathways of production. *J. Phycol.* **40**, 293–304 (2004).

105. W. Admiraal, R. Laane, H. Peletier, Participation of diatoms in the amino acid cycle of coastal waters; uptake and excretion in cultures. *Mar. Ecol. Prog. Ser.* **15**, 303–306 (1984).

106. J. F. C. De Brouwer, L. J. Stal, Daily fluctuations of exopolymers in cultures of the benthic diatoms *Cylindrotheca closterium* and *Nitzschia* sp. (Bacillariophyceae). *J. Phycol.* **38**, 464–472 (2002).

107. S. N. Zheleznova, Production characteristics of the diatom *Cylindrotheca closterium* (Ehrenb.) Reimann et Lewin grown in an intensive culture at various nitrogen sources in the medium. *Mar. Biol. J.* **4**, 33–44 (2019).

108. C. Nilsson, K. Sundback, K. Sundback, Amino acid uptake in natural microphytobenthic assemblages studied by microautoradiography. *Hydrobiologia* **332**, 119–129 (1996).

109. N. Staats, B. De Winder, L. R. Mur, L. J. Stal, Isolation and characterization of extracellular polysaccharides from the epipelic diatoms *Cylindrotheca closterium* and *Navicula salinarum. Eur. J. Phycol.* **34**, 161–169 (1999).

110. R. Urbani, E. Magaletti, P. Sist, A. M. Cicero, Extracellular carbohydrates released by the marine diatoms *Cylindrotheca closterium, Thalassiosira pseudonana* and *Skeletonema costatum:* Effect of P-depletion and growth status. *Sci. Total Environ.* **353**, 300–306 (2005).

111. Y. I. Wolf, E. V. Koonin, A tight link between orthologs and bidirectional best hits in bacterial and archaeal genomes. *Genome Biol. Evol.* **4**, 1286–1294 (2012).

112. M. Kanehisa, Y. Sato, M. Kawashima, KEGG mapping tools for uncovering hidden features in biological data. *Protein Sci.* **31**, 47–53 (2022).

113. S. M. D. Seaver, F. Liu, Q. Zhang, J. Jeffryes, J. P. Faria, J. N. Edirisinghe, M. Mundy, N. Chia, E. Noor, M. E. Beber, A. A. Best, M. DeJongh, J. A. Kimbrel, P. D'haeseleer, S. R. McCorkle, J. R. Bolton, E. Pearson, S. Canon, E. M. Wood-Charlson, R. W. Cottingham, A. P. Arkin, C. S. Henry, The ModelSEED Biochemistry Database for the integration of metabolic annotations and the reconstruction, comparison and analysis of metabolic models for plants, fungi and microbes. *Nucleic Acids Res.* **49**, D575–D588 (2021).

114. R. Caspi, R. Billington, I. M. Keseler, A. Kothari, M. Krummenacker, P. E. Midford, W. K. Ong, S. Paley, P. Subhraveti, P. D. Karp, The MetaCyc database of metabolic pathways and enzymes - A 2019 update. *Nucleic Acids Res.* **48**, D445–D453 (2020).

115. A. Bateman, M.-J. Martin, S. Orchard, M. Magrane, R. Agivetova, S. Ahmad, E. Alpi, E. H. Bowler-Barnett, R. Britto, B. Bursteinas, H. Bye-A-Jee, R. Coetzee, A. Cukura, A. Da Silva, P. Denny, T. Dogan, T. Ebenezer, J. Fan, L. G. Castro, P. Garmiri, G. Georghiou, L. Gonzales, E. Hatton-Ellis, A. Hussein, A. Ignatchenko, G. Insana, R. Ishtiaq, P. Jokinen, V. Joshi, D. Jyothi, A. Lock, R. Lopez, A. Luciani, J. Luo, Y. Lussi, A. MacDougall, F. Madeira, M. Mahmoudy, M. Menchi, A. Mishra, K. Moulang, A. Nightingale, C. S. Oliveira, S. Pundir, G. Qi, S. Raj, D. Rice, M. R. Lopez, R. Saidi, J. Sampson, T. Sawford, E. Speretta, E. Turner, N. Tyagi, P. Vasudev, V. Volynkin, K. Warner, X. Watkins, R. Zaru, H. Zellner, A. Bridge, S. Poux, N. Redaschi, L. Aimo, G. Argoud-Puy, A. Auchincloss, K. Axelsen, P. Bansal, D. Baratin, M.-C. Blatter, J. Bolleman, E. Boutet, L. Breuza, C. Casals-Casas, E. de Castro, K. C. Echioukh, E. Coudert, B. Cuche, M. Doche, D. Dornevil, A. Estreicher, M. L. Famiglietti, M. Feuermann, E. Gasteiger, S. Gehant, V. Gerritsen, A. Gos, N. Gruaz-Gumowski, U. Hinz, C. Hulo, N. Hyka-Nouspikel, F. Jungo, G. Keller, A. Kerhornou, V. Lara, P. Le Mercier, D. Lieberherr, T. Lombardot, X. Martin, P. Masson, A. Morgat, T. B. Neto, S. Paesano, I. Pedruzzi, S. Pilbout, L. Pourcel, M. Pozzato, M. Pruess, C. Rivoire, C. Sigrist, K. Sonesson, A. Stutz, S. Sundaram, M. Tognolli, L. Verbregue, C. H. Wu, C. N. Arighi, L. Arminski, C. Chen, Y. Chen, J. S. Garavelli, H. Huang, K. Laiho, P. McGarvey, D. A. Natale, K. Ross, C. R. Vinayaka, Q. Wang, Y. Wang, L.-S. Yeh, J. Zhang, P. Ruch, D. Teodoro, UniProt: The universal protein knowledgebase in 2021. *Nucleic Acids Res.* **49**, D480–D489 (2021).

116. A. Chang, L. Jeske, S. Ulbrich, J. Hofmann, J. Koblitz, I. Schomburg, M. Neumann-Schaal, D. Jahn, D. Schomburg, BRENDA, the ELIXIR core data resource in 2021: New developments and updates. *Nucleic Acids Res.* **49**, D498–D508 (2021).

117. A. Fleischmann, IntEnz, the integrated relational enzyme database. *Nucleic Acids Res.* **32**, 434D–437 (2004).

118. M. H. Saier, V. S. Reddy, G. Moreno-Hagelsieb, K. J. Hendargo, Y. Zhang, V. Iddamsetty, K. J. K. Lam, N. Tian, S. Russum, J. Wang, A. Medrano-Soto, The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res.* **49**, D461–D467 (2021).

119. L. D. H. Elbourne, S. G. Tetu, K. A. Hassan, I. T. Paulsen, TransportDB 2.0: A database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.* **45**, D320–D324 (2017).

120. T. N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786 (2011).

121. B. Gschloessl, Y. Guermeur, J. M. Cock, HECTAR: A method to predict subcellular targeting in heterokonts. *BMC Bioinformatics* **9**, 393 (2008).

122. M. G. Claros, MitoProt, a Macintosh application for studying mitochondrial proteins. *Bioinformatics* **11**, 441–447 (1995).

123. M. Cokol, R. Nair, B. Rost, Finding nuclear localization signals. *EMBO Rep.* **1**, 411–415 (2000).

124. O. Emanuelsson, S. Brunak, G. von Heijne, H. Nielsen, Locating proteins in the cell using TargetP, SignalP and related tools. *Nat. Protoc.* **2**, 953–971 (2007).

125. B. Grant, J. Madgwick, P. G. Dal, Growth of *Cylindrotheca closterium* var. californica (Mereschk.) Reimann & Lewin on nitrate, ammonia, and urea. *Mar. Freshw. Res.* **18**, 129 (1967).