The ongoing need for rates: can physiology and omics come together to co-design the measurements needed to understand complex ocean biogeochemistry?

Robert Strzepek[1], Brook Nunn[2], Lennart Bach[1], John Berges[3], Erica Young[3], and Philip Boyd[1]

[1]University of Tasmania
[2]University of Washington
[3]University of Wisconsin-Milwaukee

November 22, 2022

Abstract

The necessity to understand the influence of global ocean change on biota has exposed wide-ranging gaps in our knowledge of the fundamental principles that underpin marine life. Concurrently, physiological research has stagnated, in part driven by the advent and rapid evolution of molecular biological techniques, such that they now influence all lines of enquiry in biological and microbial oceanography. This dominance has led to an implicit assumption that physiology is outmoded, and advocacy that ecological and biogeochemical models can be directly informed by omics. However, the main modelling currencies continue to be biological rates and biogeochemical fluxes. Here we ask: how do we translate the wealth of information on physiological potential from omics-based studies to quantifiable physiological rates and, ultimately, to biogeochemical fluxes? Based on the trajectory of the state-of-the-art in biomedical sciences, along with case-studies from ocean sciences, we conclude that it is unlikely that omics can provide such rates in the coming decade. Thus, while physiological rates will continue to be central to providing projections of global change biology, we must revisit the metrics we rely upon. We advocate for the co-design of a new generation of rate measurements that better link the benefits of omics and physiology.

OXFORD UNIVERSITY PRESS | Journal of Plankton Research

The ongoing need for rates: can physiology and omics come together to co-design the measurements needed to understand complex ocean biogeochemistry?

Journal:	*Journal of Plankton Research*
Manuscript ID	Draft
Manuscript Type:	Horizons
Date Submitted by the Author:	n/a
Complete List of Authors:	Strzepek, Robert; University of Tasmania, Australian Antarctic Program Partnership (AAPP), Institute for Marine and Antarctic Studies Nunn, Brook; University of Washington, Department of Genome Sciences Bach, Lennart; University of Tasmania, Institute for Marine and Antarctic Studies Berges, John; University of Wisconsin-Milwaukee, Department of Biological Sciences Young, Erica; University of Wisconsin-Milwaukee, Department of Biological Sciences Boyd, Philip; University of Tasmania, Australian Antarctic Program Partnership (AAPP), Institute for Marine and Antarctic Studies
Keywords:	physiology, omics, co-design, marine biogeochemistry, ocean change

The ongoing need for rates: can physiology and omics come together to co-design the measurements needed to understand complex ocean biogeochemistry?

Robert F. Strzepek[1*], Brook L. Nunn[2], Lennart T. Bach[3], John A. Berges[4], Erica B. Young[4], Philip W. Boyd[1]

Affiliations

1 Australian Antarctic Program Partnership (AAPP), Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia

2 Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

3 Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, TAS, Australia

4 Department of Biological Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA

*Corresponding author: robert.strzepek@utas.edu.au

Keywords: physiology, omics, co-design, marine biogeochemistry, ocean change

8 December 2021

Abstract

The necessity to understand the influence of global ocean change on biota has exposed wide-ranging gaps in our knowledge of the fundamental principles that underpin marine life. Concurrently, physiological research has stagnated, in part driven by the advent and rapid evolution of molecular biological techniques, such that they now influence all lines of enquiry in biological and microbial oceanography. This dominance has led to an implicit assumption that physiology is outmoded, and advocacy that ecological and biogeochemical models can be directly informed by omics. However, the main modelling currencies continue to be biological rates and biogeochemical fluxes. Here we ask: how do we translate the wealth of information on physiological potential from omics-based studies to quantifiable physiological rates and, ultimately, to biogeochemical fluxes? Based on the trajectory of the state-of-the-art in biomedical sciences, along with case-studies from ocean sciences, we conclude that it is unlikely that omics can provide such rates in the coming decade. Thus, while physiological rates will continue to be central to providing projections of global change biology, we must revisit the metrics we rely upon. We advocate for the co-design of a new generation of rate measurements that better link the benefits of omics and physiology.

Introduction

A major challenge for ocean scientists is to address key questions on future ecosystem services. For example, how will global climate change alter low latitude primary productivity and hence food security? A powerful tool to address these global-scale questions is Earth system models, such as those within the Coupled Model Intercomparison Project (CMIP6) (Kwiatkowski et al., 2020). The CMIP currencies are mainly the rates at which metabolism occurs in living organisms (i.e., physiological rates) and the biogeochemical fluxes of bioactive elements. It is unlikely that these currencies will change in the coming decade, for example when CMIP7 is developed. At present, the accuracy of the model projections is hindered by two issues: 1) computational limitations to developing more complex parameterisations for processes such as nitrogen (N) fixation (Kwiatkowski et al., 2020) and 2) our inability to untangle how marine life responds to complex ocean change. For the latter, we need to decipher the fundamental physiological rules that govern biological responses to ocean change. These include the metabolic co-dependencies in response to multiple stressors, and strategies to buffer responses to rapid change, such as phenotypic plasticity and microevolution.

The physiological metrics used to quantify biological rates that are the cornerstones of Earth system models, such as primary productivity, have not fundamentally changed in decades. In contrast, omics techniques have evolved rapidly this century and have superseded physiological metrics as the main approach to study the fundamental principles driving marine life. With this dominance has come an implicit assumption by many that measuring physiological rates directly is obsolete, as they can be inferred from omics (Hellweger, 2020; McCain et al., 2021). However, omics provides a surfeit of data, at a level of detail that is often difficult to relate to the information provided by physiological rate measurements and the current needs of Earth system models. This growing mismatch between the currencies of global-scale models (rates and fluxes) and the aspirations of omics (coupling cellular potential via omics to Earth system model projections) must be addressed urgently.

Here we ask: How do we translate the wealth of information on physiological *potential* from omics-based studies to quantifiable physiological rates and, ultimately, to biogeochemical processes and their representation in Earth system models? We employ three approaches to address this question. First, we examine the evolution of research into ocean N_2 fixation from the perspective of advances in physiology and omics (Fig. 1). Second, we examine the recent trajectory of biomedical research to forecast how ocean sciences might evolve in the next decade. Third, we broaden our view by examining insights that can be gained for understanding the ocean phosphorus (P) and iron (Fe) cycles by better linking omics and physiology. We conclude with advocacy for the co-design of better physiological tools.

Lessons from marine diazotrophy

Here, we use the history of N_2 fixation (diazotrophy) research to reveal the benefits and limitations of physiological rate measurements, and how these measurements are complemented by more recent omics approaches (Fig. 1).

The contribution of diazotrophy to the supply of new N is central to understanding ocean N cycling (Fogg, 1942; Dugdale et al., 1961). Physiological studies played an important early role by quantifying rates of diazotrophy (e.g., Dilworth, 1966). These measurements provided the integrated rates necessary to estimate global biogeochemical fluxes of N (Karl et al., 2002), and to identify the environmental drivers of N_2 fixation (see Carpenter and Capone, 2008), including how climate changes may affect future diazotrophy (Garcia et al., 2011; Hutchins et al., 2013). However, imbalances in these N fluxes have uncovered unidentified N sources, and the subsequent application

of genetic tools has identified additional diazotrophic taxa that contribute to ocean N_2 fixation (Zehr and Capone, 2020).

Nitrogen fixation provides clear examples of both the limitations and benefits of non-targeted omics-based discoveries (Fig. 1). Nitrogenase (*nif*) genes can be used to detect N_2 fixation potential, and their expression is used as an index of N_2 fixation activity (Zehr et al., 1996; Zehr and Montoya, 2007). Omics has revealed diverse N_2 fixers including the unicellular cyanobacteria *Crocosphaera* and UCYN-A, and endosymbiotic and heterotrophic diazotrophs (Mehta et al., 2003; Church et al., 2005; Martinez-Perez et al., 2016). However, *nif* gene abundance does not directly equate to N_2 fixation rates (Turk-Kubo et al., 2013). Transcriptomics and proteomics targeting *nif* genes provide more relevant information about nitrogenase *activity* than genomics. However, taxon-specific dynamics can complicate estimates of community N_2 fixation rates (Church et al., 2005), and measurements of *nif* expression are not well correlated with ^{15}N-based rates of N_2 fixation (Turk et al., 2011).

Thus, despite the insights gained from omics, critical gaps remain in our understanding of the phylogenies, distribution, and physiology of marine N_2 fixers, and accurate global estimates of N_2 fixation remain elusive (Zehr and Capone, 2020). Measuring N_2 fixation remains critical to estimate the biogeochemical processing and ecological fates of new N. However, N_2 fixation is not included in the CMIP6 models, which presently project declining productivity in low latitude oceans in coming decades (Kwiatkowski et al., 2020). So, both rates and omics will be needed increasingly to reveal and quantify currently unknown (but biogeochemically important) pathways for the turnover of N (Fig. 1) to improve global models.

Resolving these unknowns will require combined measurements of *nif* gene expression with rate measurements based on nitrogenase enzyme activity (e.g., Turk et al., 2011). Broader application of flow-through high-throughput rate measurements can improve the spatial and temporal coverage of N_2 fixation (Cassar et al., 2018). Rates, when coupled with omics approaches to N_2 fixation research (Tang et al., 2020), will continue to expand our understanding of diazotroph diversity and could help focus N_2 fixation rate measurements on these emerging diazotrophic groups (Zehr and Capone, 2020). Mechanistic controls on diazotrophy can be revealed through variations in *nif* gene expression (Church et al., 2005), supporting prior conclusions that local environmental conditions influence N_2 fixation rates (Carpenter and Capone, 2008; Capone, 1993). Such environmental controls could be further explored using targeted proteomics analyses (e.g., Saito et al., 2011).

The status of omics

Both marine and biomedical sciences focus on the genome, transcriptome, proteome, and metabolome, with most research on the first three. In the field of meta-omics, marine metagenomics has set the pace, and is directly influencing research into the human microbiome (Poceviciute and Ismagilo, 2019). Here, we focus on genomics through to proteomics at the cellular level where, in contrast to meta-omics, biomedical research has led the way (Okada and Kuroda 2019). Genomics demonstrates the breadth of possible gene functions, but only catalogues the functional potential of an organism (Sunagawa et al., 2015). Transcriptomics is a popular approach to explore how organisms respond to environmental change by characterizing shifts in mRNA abundance (Evans, 2015). Feder and Walser (2005) offered a pointed description of the major issues facing the use of transcriptomics in finding the genes that matter for environmental adaptation. Their critique focused on three major issues: (1) genes with large impacts on fitness are rare and therefore unlikely to be identified with transcriptomics, (2) the relationship between gene expression and fitness is unreliable, and (3) fitness is primarily determined by proteins, and mRNA abundance is a poor proxy for protein abundance. Proteomics, on the other hand, provides taxonomically specific information on structural and metabolic enzymes with tighter correlation to functional activity. Proteomics has advanced methodologically, with more accurate standardized quantitative analyses (Collins et al., 2017; Pino et al., 2020) and protein identifications that allow metabolic profiling (Nunn et al., 2013; Mikan et al., 2020).

Numerous efforts have been made to identify correlations between omics layers. However, evidence from both marine and biomedical science reveals that making these linkages is not straightforward. For example, in marine sciences it is well recognized that the amplitude and timing of the mRNA pool does not align with protein expression. This misalignment was illustrated in Waldbauer et al. (2012) while tracking diel changes in the transcriptome and proteome within a single cyanobacteria species (Fig. 2). Subsequent research on the model diatom *Phaeodactylum tricornutum* used multiple omics layers to explore the regulation of N limitation and again reported mismatches between transcript, protein, and metabolite abundance (Remmers et al., 2018). In the further advanced biomedical field, it remains difficult to obtain mechanistic and functional insights by simply integrating multiomics data (Okada and Kuroda, 2019). As far back as the late eighties, Kurland and Ehrenberg (1987) discussed the challenges of linking cellular design and molecular design (such as via enzyme expression) in the context of physiology. More recently, Lalanne et al. (2018) uncovered post-transcriptional controls that ensure the maintenance of the protein stoichiometries required for specific biological pathways. This compensatory mechanism rectifies

divergences in regulation driven by changes of internal promoters and terminators. Hence, even in advanced biomedical research there are confounding issues, driven by post-transcriptional and post-translational modifications to enzymes, in deriving metabolic rates from omics.

In the marine context, omics has clearly demonstrated large scale patterns in microbial diversity across oceanic provinces and provided insights into which metabolic pathways are active (Fig. 1). However, omics-based approaches provide static 'snap-shots' of physiological potential, and we need to improve our quantitative, process-level understanding of the roles of marine microbes in biogeochemical cycles. Indeed, it is physiological activity or realized potential – the chemical fluxes generated by cellular metabolism as modified by biological species differences, external environmental drivers, and the interactions between the two – that drives biogeochemical cycles.

Linking physiology and omics: the need for co-design

We propose that physiological rates can bridge biogeochemistry and omics. Physiological rates quantify the integrated activity of proteins that drive marine biogeochemical cycles in units that modellers can use (Fig. 3). Research into the ocean's N cycle reveals the potential of using the joint expertise of the physiology and omics communities (i.e., co-design) to guide future research (Fig. 1). We can extend this complementary approach to use omics datasets to develop new targeted physiological metrics that improve the parameterisation of biogeochemical processes. Here, we explore the feasibility of co-design using case studies of the ocean P and Fe cycles that illustrate how physiological metrics may act as a 'currency converter' to link omics datasets and biogeochemical models.

In the case of P, a lab study used proteomics and physiological metrics to explore the cumulative effect of five climate-change stressors on a subpolar diatom (Boyd et al., 2015). A central finding was that the effect of decreased nutrient supply in a future ocean was offset by warming. Proteomics revealed that a decreased need for P was driven by the under-expression of P-containing proteins associated with translation (Fig. 3). Physiological metrics corroborated this finding, with lower cellular P quotas under warming. Hence, P quotas acted as a currency converter between protein synthesis and the biogeochemical cycle of P. They showed serendipitously a link between protein synthesis and P quotas. In the future, we must actively seek conceptual linkages, rather than uncovering them by chance. Better links from omics via physiology to biogeochemistry would benefit from input from the modelling and biogeochemical research communities.

Physiology was established earlier than omics or biogeochemistry and so many of the conventional metrics used preceded developments in these disciplines. This begs the question: are we currently measuring the best physiological metrics to mesh omics with biogeochemistry? Two examples that begin to straddle the gaps between omics and physiology come from Saito et al. (2011) and Wu et al. (2019). The former revealed diel changes in the proteome, including Fe-metalloproteins involved in N_2 fixation and photosynthesis of *Crocosphaera watsonii* resulting in more efficient use of Fe, which is essential for N_2 fixation. In the latter case, protein expression and physiological metrics were coupled to examine the influence of Fe and manganese on *Phaeocystis antarctica*.

Although our current choice of physiological metrics needs urgent attention, there is compelling evidence of the utility of long-established assays, such as those used to determine the macromolecular P content of cells from Liefer et al. (2019), for more innovative phytoplankton cellular P models (Inomura et al., 2020). But can we be inventive, and use omics to interpret P physiology in a more holistic manner (Fig. 4)? Physiology can provide valuable insights, even when considering only a few components of the cellular P cycle. Imagine the progress if we developed better metrics jointly with omics (Feng et al., 2014; Lin et al. 2016). So, the way ahead may be to use molecular biology to 'reverse engineer' the most pertinent physiological metrics (Fig. 4). For example, a useful point of departure would be to select processes in which protein abundance correlates with quantifiable metabolic activity. Such co-design, in our opinion, will further facilitate the transition from lab- to field-based omics and will lead in the coming decades to incorporation of omics into biogeochemical models.

The transition to field studies will face additional challenges that centre on how marine biota integrate environmental history (i.e., cellular status imposed by conditions encountered prior to sampling; Fig. 2) (Prairie et al., 2012; Deutschmann et al., 2021). This requires a multi-stranded approach. First, placing the sampling locale in a wider environmental context (Figure 5A). For example, profiling robotic floats with multiple sensors are providing synoptic snapshots of spatial variability in ocean properties along with the prior seasonal dynamics of key resources such as nutrients (Claustre et al., 2021). Second, how do such prior oceanic conditions set cellular status, for example the degree of Fe stress (Fig. 5B)? An open question is whether the relationship between environmental forcing and cellular status is instantaneous or lagged (Fig. 2). Will such co-designed metrics reconcile a biological product with a chemical residual since different physiological metrics display a range of response times (Boyd et al., 2005; Baker et al., 2018), as do different omics layers (Waldbauer et al., 2012)? One promising approach to probe environmental history and cellular status is physiological titration. For example, by manipulating Fe availability to contextualize cellular Fe status (Fig. 5B).

Towards the future

We conclude with recent field-leading examples from ocean sciences, that seek to derive metabolic rates from omics, explored through the lens of biomedical sciences. Saito et al. (2020) conducted metaproteomic analysis on subsurface biota in the Tropical North Pacific to pinpoint commonly occurring enzymes. They reported that nitrite oxidoreductase associated with the bacterium *Nitrospina* was abundant in this stratum and explored whether they could estimate rates of nitrite oxidation using wide-ranging methods, including biochemistry (specific activity), physiology (Michaelis-Menten kinetics), and omics. Despite employing this innovative suite of approaches derived rates ranged >200-fold, pointing to the need to develop targeted physiological assays (c.f. Fig. 4). There are also promising initial developments from the emergence of phenomenological models based on simple geochemical/taxonomic principles that yield phytoplankton growth rates assuming steady-state growth (McCain et al., 2021).

The latest developments in biomedical and model-system omics suggest obtaining rates from omics is still under development. First, holistic investigations of well-characterized model organisms have tracked every metabolite and protein to generate enzyme-directed functional rates in the bacterium *Escherichia coli* (Taniguchi et al., 2010) and the yeast *Saccharomyces cerevisiae* (Ho et al., 2018), but this approach is restricted to the organisms for which the function of every gene and protein is known. Second, expression-fitness landscapes (linking enzyme expression with growth rate) have revealed that enzyme expression can have a 'ripple' effect across layers of biological organisation ranging from mechanistic, regulatory to systemic (Lalanne et al., 2021), which adds further complexity to deriving growth rates from enzymatic fluxes. Third, sophisticated microbiome studies (from cheese to the human gut) (Poceviciute and Ismagilo, 2019), which are more akin to oceanic microbial systems, reveal that there are still a high number of metabolic functions that remain uncharacterized (Price et al., 2018). Fourth, progress in tackling cell regulatory mechanisms using multiomic modelling has been made but requires complex computing using deep neural networks such as GEMS (Genome-scale metabolic models) (Okada and Kuroda 2019).

These four categories of advanced well-resourced research point to challenges yet to be surmounted in obtaining physiological rates from omics for biomedical sciences. But, they also provide cautionary lessons for ocean sciences. In our opinion, it may be more straight-forward to co-design targeted physiological metrics that better link omics with marine biogeochemistry.

Funding

This work was supported by the Australian Antarctic Program partnership [ASCI000002 to R.F.S. and P.W.B.], by a Wisconsin Sea Grant [R/HCE-33] and a RACAS grant from the University of Wisconsin-Milwaukee to E. B. Y. and J. A. B., and by a University of Tasmania Visiting Scholarship to J. A. B.

References

Baker, K. G., Radford, D. T., Evenhuis, C., Kuzhiumparam, U., Ralph, P. J., and Doblin, M. A. (2018) Thermal niche evolution of functional traits in a tropical marine phototroph. J. Phycol. 54:799-810. https://doi.org/10.1111/jpy.12759

Boyd, P. W., Strzepek, R., Takeda, S., Jackson, G., Wong, C. S., McKay, R. M. et al. (2005). The evolution and termination of an iron-induced mesoscale bloom in the northeast subarctic Pacific. Limnology and Oceanography 50: 1872-1886

Boyd, P., Dillingham, P., McGraw, C., Armstrong, E. A., Cornwall, C. E., Feng, Y.-Y., et al. (2015). Physiological responses of a Southern Ocean diatom to complex future ocean conditions. Nat. Clim. Change 6: 207–213. Doi: 10.1038/nclimate2811

Carpenter, E. J., and Capone, D. G. (2008) Nitrogen fixation in the marine environment. In: Capone, D. G., Bronk, D. A., Mulholland, M. R., and Carpenter, E. J. (eds) Nitrogen in the Marine Environment (Second Edition), Academic Press, p 141-198, https://doi.org/10.1016/B978-0-12-372522-6.00004-9.

Capone, D. G. (1993) Determination of nitrogenase activity in aquatic samples using the acetylene reduction procedure. In Kemp P. F., Sherr, B. F., Sherr, E. B., and Cole, J. J. (eds) Handbook of methods in aquatic microbial ecology. Lewis Publishers, p 621–631.

Cassar, N., Tang, W., Gabathuler, H., and Huang, K. (2018). Method for high frequency underway N_2 fixation measurements: Flow-through incubation acetylene reduction assays by cavity ring down laser absorption spectroscopy (FARACAS). Analytical Chemistry 90: 2839–2851. https://doi.org/10.1021/acs.analchem.7b04977

Church, M. J., Short, C. M., Jenkins, B. D., Karl, D. M., and Zehr, J. P. (2005) Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. Applied and Environmental Microbiology 71: 5362–5370

Claustre, H., Legendre, L., Boyd, P. W., and Levy, M. (in press) The oceans' biological carbon pumps: framework for a research observational community approach. Frontiers Marine Science.

Collins, B. C., Hunter, C. L., Liu, Y., Schilling, B., Rosenberger, G., Bader, S. L., et al. (2017) Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry. Nat. Commun. 8: No. e1002165.

Darzi, Y., Letunic, I., Bork, P., and Yamada, T. (2018) iPath3.0: interactive pathways explorer v3. Nucleic Acids Res. 46: W510–W513. Doi: 10.1093/nar/gky299

Deutschmann, I. M., Krabberød, A. K., Benites, L. F., Latorre, F., Delage, E., Marrasé, C. V. et al. (2021) Disentangling temporal associations in marine microbial networks. Research Square preprint doi: https://doi.org/10.1101/2021.07.13.452187

Dilworth, M. J. (1966) Acetylene reduction by nitrogen-fixing preparations from *Clostridium pasteurianum*. Biochim. Biophys. Acta. 127:285–294

Dugdale, R. C., Goering, J. J., and Ryther, J. H. (1961). Nitrogen fixation in the Sargasso Sea. Deep Sea Res. 7, 298–300.

Dugdale, V. A., and Dugdale, R. C. (1962) Nitrogen metabolism in lakes II. Role of nitrogen fixation in Sanctuary Lake, Pennsylvania. Limnol. Oceanogr. 7: 170-177. https://doi.org/10.4319/lo.1962.7.2.0170

Evans, T. G. (2015) Considerations for the use of transcriptomics in identifying the 'genes that matter' for environmental adaptation. J Exp Biol. 218:1925-35. doi: 10.1242/jeb.114306

Feder, M. E., and Walser, J.-C. (2005) The biological limitations of transcriptomics in elucidating stress and stress responses. J. Evol. Biol. 18: 901-910. https://doi.org/10.1111/j.1420-9101.2005.00921.x

Feng, T-Y., Yang, Z. -K., Zheng, J. -W., Xie, Y., Li, D. -W., Murugan, S. B. et al. (2014) Examination of metabolic responses to phosphorus limitation via proteomic analyses in the marine diatom *Phaeodactylum tricornutum*. Scientific Reports 5:10373 DOI: 10.1038/srep10373

Fogg, G. E. (1942) Studies on nitrogen fixation by blue-green algae: I. Nitrogen fixation by *Anabaena cylindrica* Lemm. J. Exp. Biol. 19: 78-87. https://doi.org/10.1242/jeb.19.1.78

Garcia, N., Fu, F., Sedwick, P., and Hutchins, D.A. (2015) Iron deficiency increases growth and nitrogen-fixation rates of phosphorus-deficient marine cyanobacteria. ISME J 9: 238–245. https://doi.org/10.1038/ismej.2014.104

Held, N. A., Webb, E. A., McIlvin, M. M., Hutchins, D. A., Cohen, N. R., Moran, D. M., et al. (2020) Co-occurrence of Fe and P stress in natural populations of the marine diazotroph *Trichodesmium*. Biogeosciences 17: 2537–2551 https://doi.org/10.5194/bg-17-2537-2020

Hellweger, F. (2020) Combining molecular observations and microbial ecosystem modeling: A practical guide. Annu. Rev. Mar. Sci. 12:267-289. https://doi.org/10.1146/annurev-marine-010419-010829

Ho, B., Baryshnikova, A., and Brown, G. W. (2018). Unification of protein abundance datasets yields a quantitative *Saccharomyces cerevisiae* proteome. Cell systems, 6: 192-205.

Inomura K., Omta, A. W., Talmy, D., Bragg, J., Deutsch, C., and Follows, M. J. (2020) A mechanistic model of macromolecular allocation, elemental stoichiometry, and growth rate in phytoplankton. Front. Microbiol. 11:86. Doi: 10.3389/fmicb.2020.00086

Karl D. et al. (2002) Dinitrogen fixation in the world's oceans. In: Boyer E.W., Howarth R.W. (eds) The Nitrogen Cycle at Regional to Global Scales. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3405-9_2

Kriest I., Khatiwala, S., and Oschlies, A. (2010) Towards an assessment of simple global marine biogeochemical models of different complexity. Progress in Oceanography 86: 337-360, https://doi.org/10.1016/j.pocean.2010.05.002

Kurland, C. G., and Ehrenberg, M. (1987) Growth-optimizing accuracy of gene expression. Annu. Rev. Biophys. Biophys. Chem. 16: 291 – 317.

Kwiatkowski, L., Torres, O., Bopp, L., Aumont, O., Chamberlain, M., Christian, J. R., et al. (2020) Twenty-first century ocean warming, acidification, deoxygenation, and upper ocean nutrient and primary production decline from CMIP6 model projections. Biogeosciences 17: 3439–3470. https://doi.org/10.5194/bg-17-3439-2020

Lalanne, J. B., Taggart, J. C., Guo, M. S., Herzel, L., Schieler, A., and Li, G. W. (2018) Evolutionary convergence of pathway-specific enzyme expression stoichiometry. Cell 173: 749 – 761

Lalanne, J. B., Parker, D. J., and Li, G. W. (2021) Spurious regulatory connections dictate the expression-fitness landscape of translation factors. Molecular Systems Biology 17: e10302. DOI 10.15252/msb.202110302

Letunic, I., Yamada, T., Kanehisa, M., and Bork, P. (2008) iPath: interactive exploration of biochemical pathways and networks. Trends Biochem Sci 33:101−103

Liefer, J. D., Garg, A., Fyfe, M. H., Irwin, A. J., Benner, I., Brown, C. M., Follows, M. J., Omta, A. W., and Finkel, Z. V. (2019) The macromolecular basis of phytoplankton C:N:P under nitrogen starvation. Front. Microbiol. 10:763. Doi: 10.3389/fmicb.2019.00763

Lin, S., Litaker, R. W, and Sunda, W. G. (2016) Phosphorus physiological ecology and molecular mechanisms in marine phytoplankton**.** J. Phycol. 52: 10-36. https://doi.org/10.1111/jpy.12365

McCain, J. S. P., Tagliabue, A., Susko, E., Achterberg, E. P., Allen, A. E., and Bertrand, E. M. (2021) Cellular costs underpin micronutrient limitation in phytoplankton. Sci. Adv. 7: eabg6501

Martínez-Pérez, C., Mohr, W., Löscher, C. R., Dekaezemacker, J., Littmann, S., Yilmaz, P., Lehnen, N., Fuchs, B. M., Lavik, G., Schmitz, R. A., LaRoche, J., and Kuypers, M. M. (2016) The small unicellular diazotrophic symbiont, UCYN-A, is a key player in the marine nitrogen cycle. Nat Microbiol. 1:16163. doi: 10.1038/nmicrobiol.2016.163.

Mehta, M. P., Butterfield, D. A., and Baross, J. A. (2003) Phylogenetic diversity of nitrogenase (*nifH*) genes in deep-sea and hydrothermal vent environments of the Juan de Fuca Ridge. Applied and Environmental Microbiology 69: 960–970. DOI: 10.1128/AEM.69.2.960–970.2003

Mikan, M. P., Harvey, H. R., Timmins-Schiffman, E., Riffle, M., May, D. H., Salter, I., et al. (2020) Metaproteomics reveal that rapid perturbations in organic matter prioritize functional restructuring over taxonomy in western Arctic Ocean microbiomes. The ISME journal 14: 39-52.

Nunn, B. L., Faux, J. F., Hippmann, A. A., Maldonado, M. T., Harvey, H. R., Goodlett, D. R., et al. (2013) Diatom proteomics reveals unique acclimation strategies to mitigate Fe limitation. PloS ONE 8: e75653.

Okada, M. and Kuroda, S. (2019) Editorial overview: regulatory mechanism from multi-omic data Current Opinion in Systems Biology, 15:iv-vi https://doi.org/10.1016/j.coisb.2019.06.002

Pino, L. K., Searle, B. C., Yang, H-Y, Hoofnagle, A. N., Noble, W. S., and MacCoss., M. J. (2020) Matrix-matched calibration curves for assessing analytical figures of merit in quantitative proteomics. Journal of Proteome Research 19: 1147-1153.

Poceviciute, R., and Ismagilov, R. F. (2019) Human-gut-microbiome on a chip. Nature biomedical engineering 3: 500-501.

Prairie, J. C., Sutherland, K. R., Nickols, K. J. and Kaltenberg, A. M. (2012) Biophysical interactions in the plankton: A cross-scale review. Limnology and Oceanography: Fluids and Environments 2: 121–145. https://doi.org/10.1215/21573689-1964713

Price, M. N., Wetmore, K. M., Waters, R. J., Callaghan, M., Ray, J., Liu, H., Kuehl, J.V. et al. (2018). Mutant phenotypes for thousands of bacterial genes of unknown function. Nature 557: 503-509. https://doi.org/10.1038/s41586-018-0124-0

Remmers, I. M., D'Adamo, S., Martens, D. E., de Vos, R. C. H., Mumm, R., America, A. H. P., et al. (2018) Orchestration of transcriptome, proteome and metabolome in the diatom *Phaeodactylum tricornutum* during nitrogen limitation. Algal Research 35: 33-49. https://doi.org/10.1016/j.algal.2018.08.012

Saito, M. A., Bertrand, E. M., Dutkiewicz, S., Bulygin, V. V., Moran, D. M., Monteiro, F. M., et al. (2011) Iron conservation by reduction of metalloenzyme inventories in the marine diazotroph *Crocosphaera watsonii*. Proc. Natl. Acad. Sci. 108: 2184–2189. https://doi.org/10.1073/pnas.1006943108

Saito, M. A., McIlvin, M. R., Moran, D. M., Santoro, A. E., Dupont, C. L., Rafter, P. A., et al. (2020) Abundant nitrite-oxidizing metalloenzymes in the mesopelagic zone of the tropical Pacific Ocean. Nature Geoscience 13: 355–362. https://doi.org/10.1038/s41561-020-0565-6

Sedwick, P. N., Bowie, A. R., Church, T. M., Cullen, J. T., Johnson, R. J., Lohan, M. C., et al. (2020) Dissolved iron in the Bermuda region of the subtropical North Atlantic Ocean: Seasonal dynamics, mesoscale variability, and physicochemical speciation. Mar. Chem. 219, 103748. https://doi.org/10.1016/j.marchem.2019.103748

Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G. et al. (2015) Structure and function of the global ocean microbiome. Science 348: 1261359.

Tang, W., Cerdán-García, E., Berthelot, H., Polyviou, D., Wang, S., Baylay, A., et al. (2020) New insights into the distributions of nitrogen fixation and diazotrophs revealed by high-resolution sensing and sampling methods. ISME J 14: 2514–2526. https://doi.org/10.1038/s41396-020-0703-6

Taniguchi, Y., Choi, P. J., Li, G. -W., Chen, H., Babu, M., Hearn, J., et al. (2010) Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. Science 329: 533-538.

411 Toseland, A., Daines, S. J., Clark, J. R., Kirkham, A., Strauss, J., Uhlig, C., et al. (2013). The
412 impact of temperature on marine phytoplankton resource allocation and metabolism. Nat. Clim.
413 Chang. 3, 979–984.

414 Turk, K. A., Rees, A. P., Zehr, J. P., Pereira, N., Swift, P., Shelley, R., Lohan, M., Woodward, E.
415 M. S., and Gilbert, J. (2011) Nitrogen fixation and nitrogenase (nifH) expression in tropical waters
416 of the eastern North Atlantic. The ISME Journal 5: 1201–1212

417 Turk-Kubo, K. A., Achilles, K. M., Serros, T. R. C., Ochiai, M., Montoya, J. P., and Zehr, J. P.
418 (2013) Nitrogenase (nifH) gene expression in diazotrophic cyanobacteria in the Tropical North
419 Atlantic in response to nutrient amendments. Frontiers in Microbiology 3: 386

420 Waldbauer, J. R., Rodrigue, S., Coleman, M. L., and Chisholm, S. W. (2012) Transcriptome and
421 proteome dynamics of a light-dark synchronized bacterial cell cycle. PLoS ONE 7(8): e43432.
422 doi:10.1371/journal.pone.0043432

423 Walworth N. G., Lee, M. D., Fu, F. X., Hutchins, D. A., and Webb, E. A. (2016) Molecular and
424 physiological evidence of genetic assimilation to high CO_2 in the marine nitrogen fixer
425 *Trichodesmium*. Proc Natl Acad Sci USA. 113(47): E7367-E7374. doi: 10.1073/pnas.1605202113

426 Wilhelm, S. W., King, A. L., Twining, B. S., LeCleir, G. R., DeBruyn, J. M., Strzepek, R. F., et al.
427 (2013) Elemental quotas and physiology of a southwestern Pacific Ocean plankton community as a
428 function of iron availability. Aquatic Microbial Ecology 68:185-194

429 Wu, M., McCain, J. S. P., Rowland, E., Middag, R., Sandgren, M., Allen, A. E., et al. (2019)
430 Manganese and iron deficiency in Southern Ocean *Phaeocystis antarctica* populations revealed
431 through taxon-specific protein indicators. Nat Commun 10, 3582). https://doi.org/10.1038/s41467-
432 019-11426-z

433 Zehr, J. P., Braun, S., Chen, Y. B., and Mellon, M. (1996) Nitrogen fixation in the marine
434 environment: relating genetic potential to nitrogenase activity. J. Exp. Mar. Biol. Ecol. 203:61-73.

435 Zehr, J. P., and Montoya, J. P. (2007) Measuring N_2 fixation in the field. In Bothe, H., Ferguson, S.
436 J., and Newton W. E. (eds) Biology of the nitrogen cycle. Elsevier, p. 193-205.
437 https://doi.org/10.1016/B978-044452857-5.50014-X

438 Zehr, J. P., and Capone, D. G. (2020) Changing perspectives in marine nitrogen fixation. Science
439 368:eaay9514. DOI: 10.1126/science.aay9514

440

441

442 Figure legends

443 Figure 1. The contributions of physiology and omics to understanding the role of diazotrophy in the
444 ocean N cycle (based on Zehr and Capone, 2020). Key events in the physiology timeline (top green
445 line) include estimation of N fluxes through nitrogenase (Dugdale and Dugdale, 1962), initial
446 estimates of global marine N_2 fixation rates (Capone et al., 1982), and the combining of lab and
447 field measurements to understand individual diazotrophs and community contributions and
448 constraints. Pivotal events in the omics timeline (lower green line) include problem solving (Zehr
449 and Montoya, 2007) and discovery of diazotroph diversity including in unicellular cyanobacteria
450 group A (UNCYN-A) and diverse uncultured heterotrophic bacteria (UHB) (Martinez-Perez et al.,
451 2016). Recent examples of more integrated physiological and omics co-designed studies (Walworth
452 et al., 2016; Held et al. 2020) offer an important way forward.

453

454 Figure 2. Examples of the potential for mismatches in transcriptomics versus proteomics in a pico-
455 prokaryote over the diurnal cycle. A) The diel cycling and amplitudes of transcripts and proteins in
456 *Prochlorococcus* for Ribonucleotide reductase (nrdJ), the large sub-unit of Rubisco (rbcL), and
457 Geranylgeranyl diphosphate reductase (chlP). B) Histogram of lag-times for proteins and their
458 transcripts for a 312 gene dataset. Antiphase refers to genes that are offset by ~12 h (i.e., 50%) of
459 the diel cell cycle. Redrawn from Waldbauer et al. (2012).

460

461 Figure 3. An example illustrating the utility of physiological metrics as a 'currency converter' to
462 link omics and biogeochemical modelling. A) The under- (downward arrows) and over-expression
463 (upwards arrows) of proteins in 4 treatments within a climate change manipulation experiment
464 measured with proteomics (Boyd et al., 2015). Warming results in an under-expression of P-
465 containing proteins associated with translation. B) Corresponding changes to the cellular P quotas
466 of the study subject, a lab culture of a subantarctic diatom, across the treatments A-D. This
467 physiological metric reveals the causal link between under-expression of translation proteins and
468 decreased P quotas (as previously described by Toseland et al., 2013). C) A subset of global model
469 projections of upper ocean phosphate (PO_4^-) stocks across biogeochemical models of different
470 complexity (Kriest et al., 2010). The approaches employed in panels A and C can be linked using
471 the cellular P quotas obtained from panel B.

472

Figure 4. The potential of reverse-engineering physiological metrics to provide better linkages with molecular tools using the example of P. A) Findings of a physiological study (Leifer et al., 2019) using a cluster of long-established metrics (residual P pools/intracellular storage of inorganic P) to compare the P allocation strategies of a diatom (*Thalassiosira pseudonana*) and a prasinophyte (*Micromonas sp*.). B) Cartoon summarizing the known PO_4^- acquisition and metabolic pathways that may be present in most phytoplankton species (for details see Fig. 4 in Lin et al., 2016). C) A KEGG map from I-PATH (Letunic et al., 2008; Darzi et al., 2018) overlaid with the functional categories of differentially expressed proteins (mapped to KEGG pathways) involved in various biological processes for P limitation by a *Phaeocystis* species (Feng et al., 2014). I-PATH is a web-application for the visualization and analysis of cellular pathways from omics (e.g., see Nunn et al., 2013 for Fe replete versus Fe deplete proteomes).

Figure 5. Utility of environmental context to define the present physiological status of cells in relation to prior oceanic conditions. A) Dissolved Fe time series for the upper ocean in the subtropical Atlantic (BATS site) that reveals conspicuous aerosol Fe inputs (>0.5 nmol L^{-1}) along with the influence of eddy activity (<0.3 nmol L^{-1}) on dissolved Fe concentrations (Sedwick et al., 2020). B) Photosynthetic efficiency of PSII (F_v/F_m) measured in deckboard incubation experiments 'titrated' with dissolved Fe concentrations by either reducing bioavailable Fe using the fungal siderophore desferrioxamine B (DFB) or increasing it with chelated inorganic Fe addition (from Wilhelm et al., 2013). The circles denote putative linkages between chemical stocks and biological responses (red = high Fe; green = low Fe).

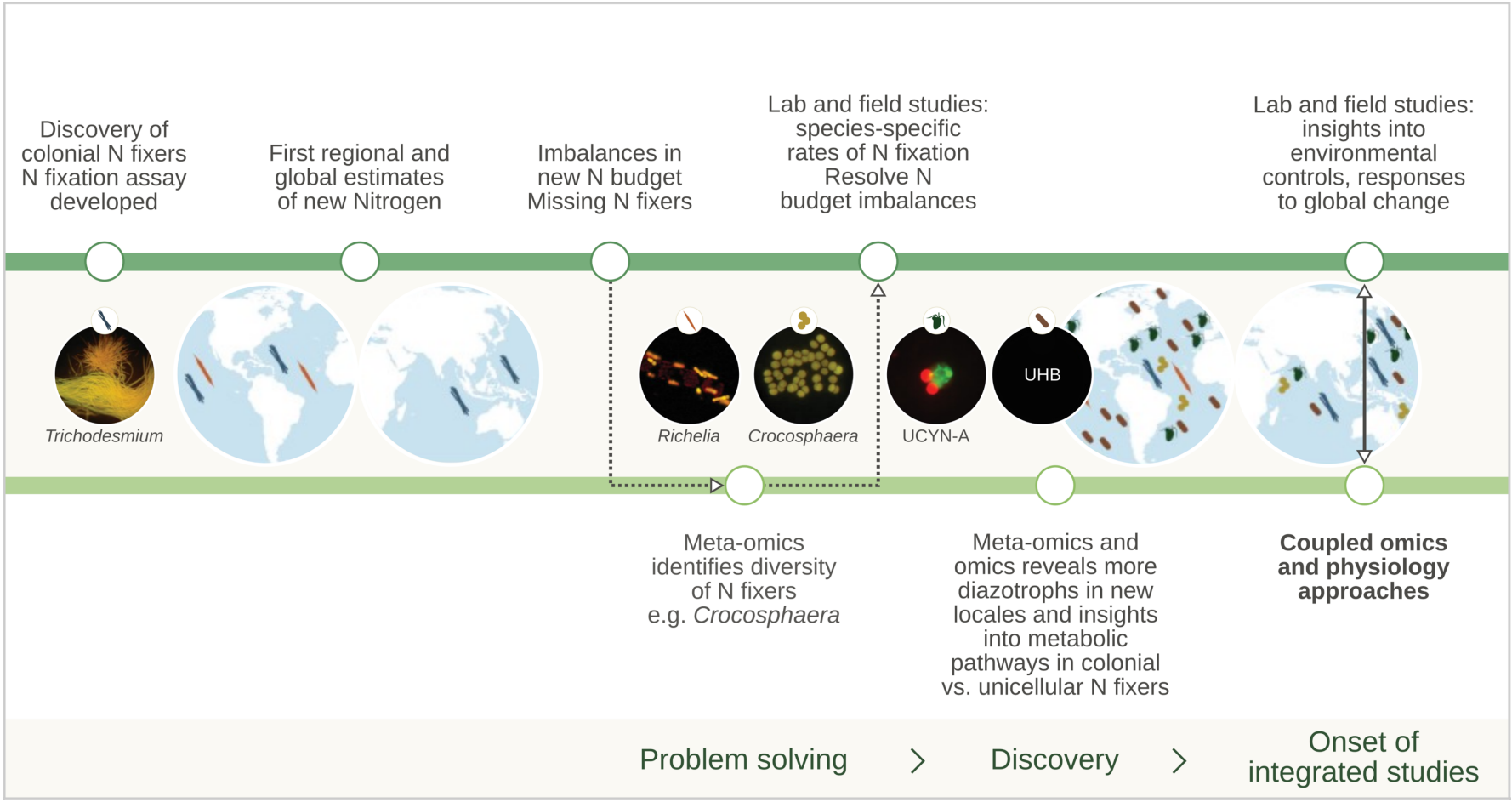
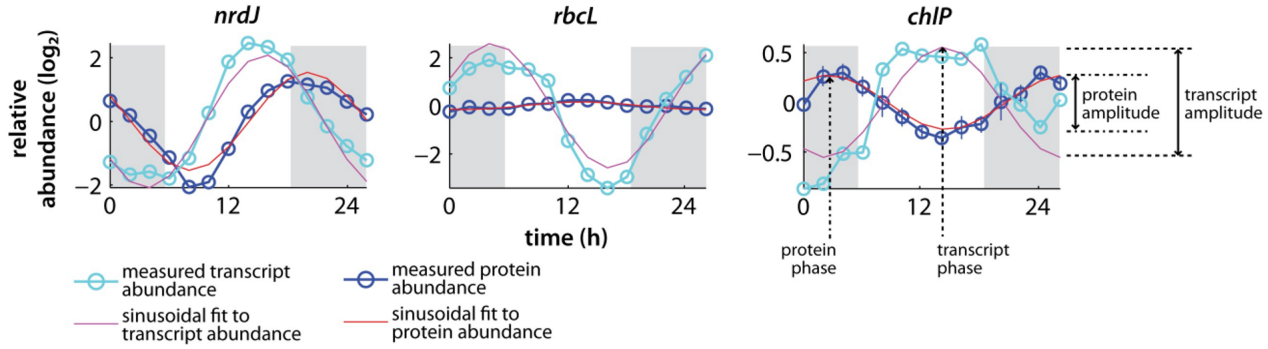

Discovery of colonial N fixers N fixation assay developed
First regional and global estimates of new Nitrogen
Imbalances in new N budget Missing N fixers
Lab and field studies: species-specific rates of N fixation Resolve N budget imbalances
Lab and field studies: insights into environmental controls, responses to global change
Trichodesmium
Richelia
Crocosphaera
UCYN-A
UHB
Meta-omics identifies diversity of N fixers e.g. Crocosphaera
Meta-omics and omics reveals more diazotrophs in new locales and insights into metabolic pathways in colonial vs. unicellular N fixers
Coupled omics and physiology approaches
Problem solving
Discovery
Onset of integrated studies

Figure 1

A


nrdJ
rbcL
chlP
relative abundance (log2)
time (h)
protein amplitude
transcript amplitude
protein phase
transcript phase
measured transcript abundance
measured protein abundance
sinusoidal fit to transcript abundance
sinusoidal fit to protein abundance

B

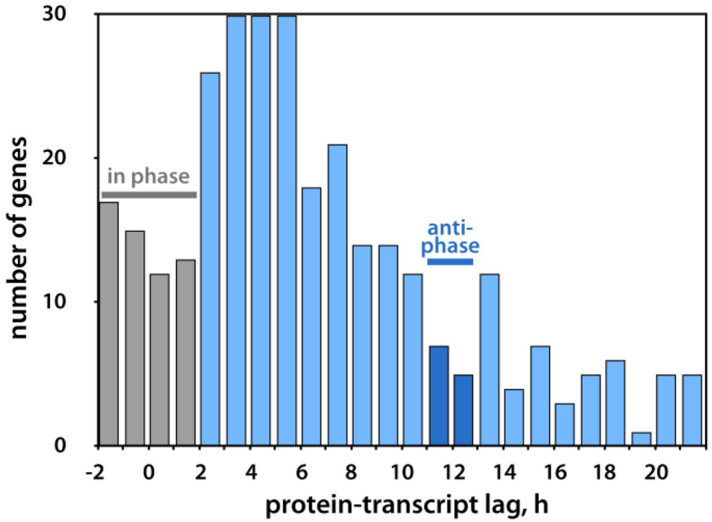

number of genes
in phase
anti-phase
protein-transcript lag, h

Figure 2

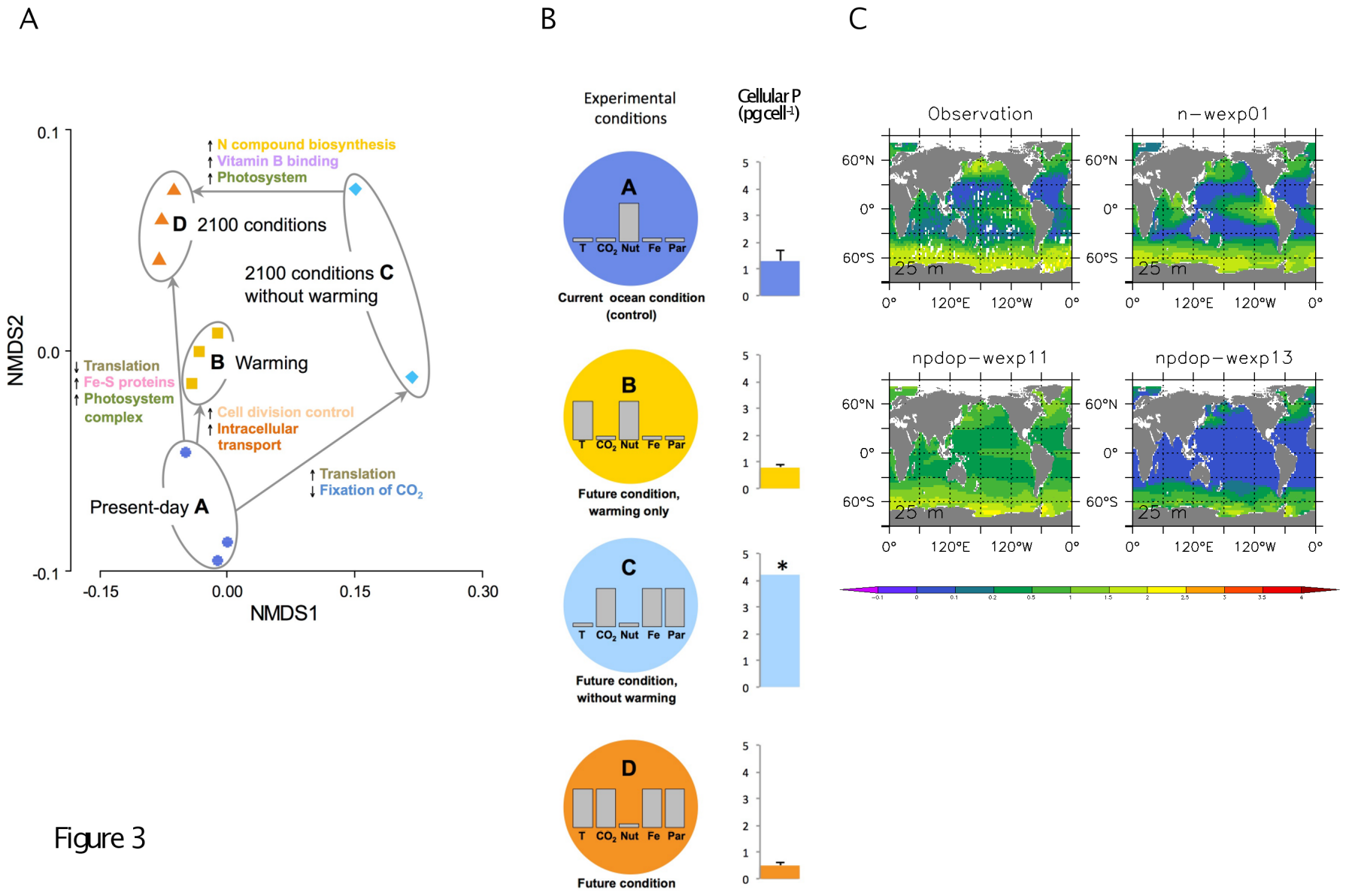
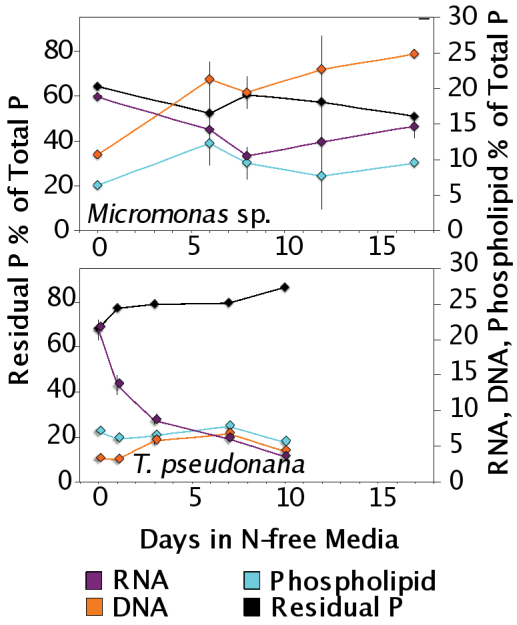
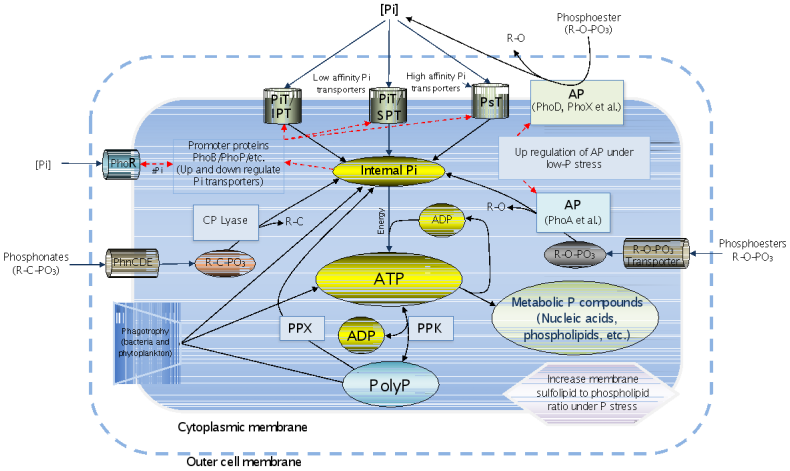

A
B
C
NMDS2
NMDS1
0.1
0.0
-0.1
-0.15
0.00
0.15
0.30
↑ N compound biosynthesis
↑ Vitamin B binding
↑ Photosystem
D 2100 conditions
2100 conditions C
without warming
B Warming
↓ Translation
↑ Fe-S proteins
↑ Photosystem
complex
↑ Cell division control
↑ Intracellular
transport
↑ Translation
↓ Fixation of CO2
Present-day A
Experimental
conditions
Cellular P
(pg cell-1)
A
T CO2 Nut Fe Par
Current ocean condition
(control)
B
Future condition,
warming only
C
Future condition,
without warming
D
Future condition
Observation
n−wexp01
npdop−wexp11
npdop−wexp13
60°N
0°
60°S
25 m
0°
120°E
120°W

Figure 3

A


Residual P % of Total P
RNA, DNA, Phospholipid % of Total P
Micromonas sp.
T. pseudonana
Days in N-free Media
RNA
DNA
Phospholipid
Residual P

B


[Pi]
Phosphoester (R-O-PO3)
R-O
Low affinity Pi transporters
High affinity Pi transporters
PiT/ IPT
PiT/ SPT
PstI
AP (PhoD, PhoX et al.)
Promoter proteins PhoB/PhoP/etc. (Up and down regulate Pi transporters)
PhoR
Internal Pi
Up regulation of AP under low-P stress
CP Lyase
R-C
Energy
ADP
AP (PhoA et al.)
Phosphonates (R-C-PO3)
PhnCDE
R-C-PO3
ATP
R-O-PO3 transporter
Phosphoesters R-O-PO3
Metabolic P compounds (Nucleic acids, phospholipids, etc.)
Phagotrophy (bacteria and phytoplankton)
PPX
ADP
PPK
PolyP
Increase membrane sulfolipid to phospholipid ratio under P stress
Cytoplasmic membrane
Outer cell membrane

C

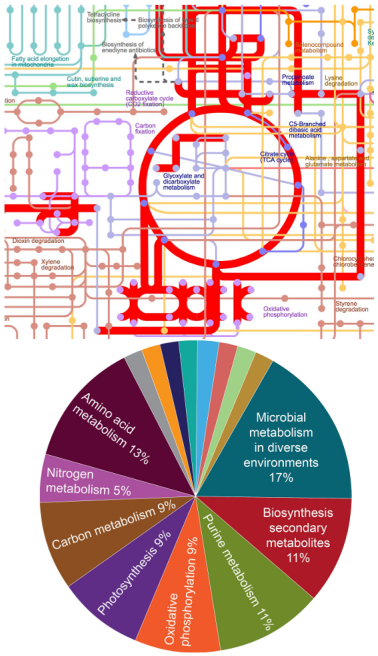

Amino acid metabolism 13%
Microbial metabolism in diverse environments 17%
Nitrogen metabolism 5%
Carbon metabolism 9%
Biosynthesis secondary metabolites 11%
Photosynthesis 9%
Oxidative phosphorylation 9%
Purine metabolism 11%

Figure 4

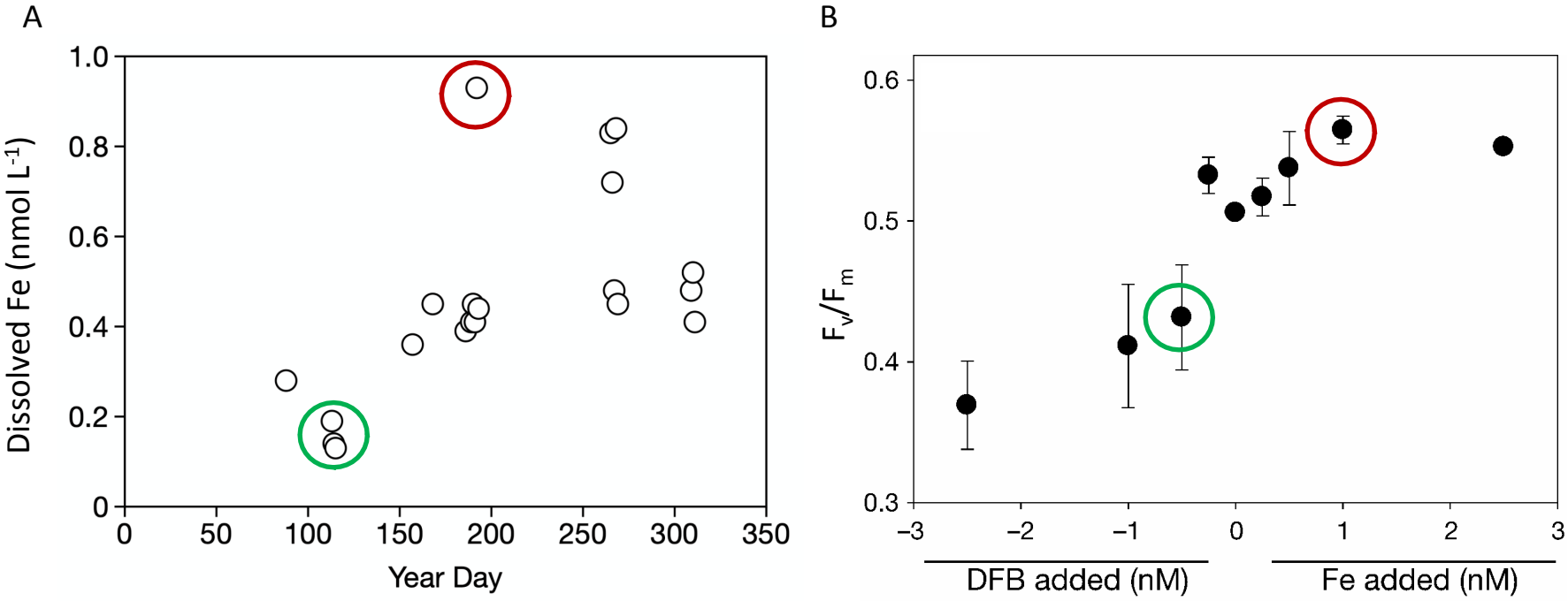

A
Dissolved Fe (nmol L-1)
1.0
0.8
0.6
0.4
0.2
0
0
50
100
150
200
250
300
350
Year Day
B
Fv/Fm
0.6
0.5
0.4
0.3
−3
−2
−1
0
1
2
3
DFB added (nM)
Fe added (nM)

Figure 5