Statement of authorship: All authors contributed to the ideas in this manuscript and to its writing.

Title: How mutualisms arise in phytoplankton communities: building eco-evolutionary principles for aquatic microbes

Authors: Elena Kazamia[1*], Katherine Emma Helliwell[1§], Saul Purton[2] and Alison Gail Smith[1]

Affiliations:

[1]Department of Plant Sciences, University of Cambridge, Downing Street, Cambridge, CB2 3EA, UK.

[2]Institute of Structural and Molecular Biology, University College London, Gower Street, London WC1E 6BT, UK.

*current address: Department of Civil and Environmental Engineering, MIT, 15 Vassar Street, Cambridge, MA 02139, USA

§current address: The Marine Biological Association of the UK, Citadel Hill, Plymouth, PL1 2PB, UK

Key words: co-cultures, eco-evolutionary dynamics, Foraging-to-Farming hypothesis, metabolite exchange, metagenomics, microbial communities, mutualism, phytoplankton, vitamins

Running title: Metabolic drivers of mutualisms between aquatic microorganisms

Article Type: Reviews and Syntheses

Number of words: abstract (198 words), main text (6630 words), Fig 1 (41 words) Fig 2 (124 words), Fig 3 (121 words), Fig 4 (93 words) Fig 5 (146 words)

Number of references: 100

Number of figures, tables, and text boxes: 5 figures

Corresponding author details:

Dr Elena Kazamia
MIT 48-209
15, Vassar Street
Cambridge, MA
02139
USA

Email: kazamia@mit.edu
Tel:+1(857)259-9457
Fax: +1(617) 324-0336

Abstract

Extensive sampling and metagenomics analyses of plankton communities across all aquatic environments are beginning to provide insights into the ecology of microbial communities. In particular the importance of metabolic exchanges that provide a foundation for ecological interactions between microorganisms has emerged as a key factor in forging the foundations of such communities. Here we describe the insights from this work, and show how both studies of environmental samples and physiological experimentation in the laboratory with defined microbial co-cultures are being used to decipher the metabolic and molecular underpinnings of such exchanges. In addition, we explain how metabolic modelling may be used to conduct investigations in reverse, deducing novel molecular exchanges from analysis of large-scale datasets, which can identify persistently co-occurring species. Finally, we consider how knowledge of microbial community ecology can be built into evolutionary theories tailored to these species' unique lifestyles. We propose a novel model for the evolution of metabolic auxotrophy in microorganisms that arises as a result of symbiosis, termed the Foraging-to-Farming hypothesis. The model has testable predictions, fits several known examples of mutualism in the aquatic world, and sheds light on how interactions, which cement dependencies within communities of microorganisms, might be initiated.

Introduction

Microorganisms are the "unseen majority" of life on Earth. As well as being numerically dominant, they also constitute the major phylogenetic diversity, even within the Eukaryotes where nearly every lineage is dominated by unicellular or microscopic species, and where multicellularity is the exception rather than the rule (Fig. 1). Moreover in a range of ecosystems including the soil and ocean biomes, microorganisms make the major impact on global processes such as the biogeochemical cycling of carbon, nitrogen and sulphur (Falkowski *et al.* 2008; van der Heijden *et al.* 2008). They are particularly important in the aquatic environment because here a subset of species, the phytoplankton, which comprise both eukaryotic microalgae and cyanobacteria, are responsible for primary production, contributing an estimated 50% of the total global carbon fixation (Field *et al.* 1998), and sustaining all other trophic levels. However despite their importance, our understanding of the ecology of the phytoplankton and the associated microorganisms in the photic zone is limited, due both to a lack of theoretical principles that are relevant to their unique lifestyles, as highlighted by Prosser *et al.* (2007), but also to the difficulty in studying many of these species in the laboratory. Aquatic microorganisms are notoriously difficult to isolate from the natural environment, as only an estimated 0.01 to 0.1% of oceanic marine bacterial cells produce colonies by standard diagnostic plating techniques (Conon & Giovannoni 2002), and most species lack morphological characters. One of the proposed explanations for the difficulty in isolating species is that community interactions, which are severed in axenic (single organism) laboratory cultures, are vital for the survival of microorganisms (e.g. Joint *et al*. 2010).

Photosynthetic algae are at the start of most aquatic food chains, and their consumption by zooplankton is the first link in a classic aquatic trophic cascade. However, many planktonic algae have more complex lifestyles (Fig. 2). Mixotrophy, where photosynthesis is combined with uptake of dissolved organic carbon, is common throughout the algal lineages. Alongside osmotrophic uptake of dissolved organic carbon (osmotrophy), representatives of several ecologically significant groups such as dinoflagellates, carry out phagotrophy by grazing on bacterioplankton prey (Jones 2000). Evidence from field studies in the North Atlantic ocean have demonstrated that photosynthetic mixotrophs can account for a staggering 40–95% of bacterivory in certain regions of the ocean (Hartmann *et al.* 2012). This complements laboratory-based studies, which have demonstrated phagotrophy in eukaryotic phytoplankton species typically considered strict autotrophs, such as the picoalga *Micromonas (*McKie-Krisberg & Sanders 2014), a particularly remarkable example given its tiny cell size (<2 µm). As such, mixotrophy is increasingly being recognised as a major contributor to plankton dynamics, challenging the classic distinction made between 'phototrophic' phytoplankton and heterotrophic zooplankton (Flynn *et al.* 2012).

Similarly, defining symbiotic interactions between aquatic microbes is equally complex, since these often defy strict categorisation (with only a subset of examples shown in Fig. 2). For example, corals are one of the better-studied aquatic associations, where a cnidarian exchanges metabolites with a dinoflagellate from the *Symbiodinium* genus. Traditionally, this has been considered a classic case of mutualism, as inorganic waste metabolites from the animal host are exchanged for organic nutrients fixed by dinoflagellate photosynthesis (Muscatine & Porter 1977), benefitting both partners. However, experimental evidence shows that there is considerable functional diversity within the *Symbiodinium* lineage, with certain clades interacting with hosts in a manner that is closer to parasitism, by fixing and releasing significantly less carbon than close relatives also capable of the association (Stat *et al.* 2008). Furthermore, some species of *Symbiodinium* have been shown to switch their behaviour from mutualist to parasitic depending on the manner of transmission. *S. microadriaticum*, which infects jellyfish as an endophotosymbiont, was experimentally enforced to infect its host through horizontal transmission (defined as infectious transfer among unrelated hosts) and this selected for an evolutionary shift to parasitism (Sachs and Wilcox, 2006). Similarly, it has been observed that within the dynamic environment of open oceans, mutualisms underpinning algal blooms may turn to parasitism as the bloom reaches its climax and fades, often accompanied by sudden and extensive viral lysis (e.g. Yager *et al.* 2001).

That microbial interactions in the aquatic realm do not fit known ecological modes of interaction is not surprising, since historically these have been formulated to explain life in terrestrial biomes. We argue that the best way of interpreting aquatic microbial interactions, on a fundamental level, is through the lens of characterising the metabolism of the constituent species. We propose that symbioses in this context should be classified as "*active metabolic associations between two or more organisms, with an implied ecology*" to distinguish from the passive interactions that can be a by-product of living in a shared environment. Passive interactions would include nutrient exchange through metabolic by-product exudates, or following virus-induced lysis of cells. These give rise to a dissolved pool of organic nutrients, made available to species either at a different depth or significantly later in time, in pace with circulation events such as seasonal upwelling along continental margins, in a process known as the 'microbial loop' (Azam *et al.* 1983). Similarly, trophic interactions would not be strictly symbiotic as they constitute an active behaviour on the part of only one of the species, rather than both. Following this definition, a symbiotic interaction would require that both species are alive during their association and affect each other's metabolism, cellular functions and lifestyle. Whilst physical associations are possible and frequent in the microbial world, in our view they are not a pre-requisite for symbiosis, as metabolic exchanges could happen at a distance. We discuss the challenges associated with studying the physical aspects of microbial interactions later in the manuscript. Furthermore, we consider that whilst not all symbiotic interactions are specific, this should be a pre-requisite for mutualism, which implies recurrent interactions and the potential for co-

evolution. Nonetheless, the mutualism need not be permanent, but might operate under particular environmental conditions, or stage of the life cycle.

An important development in recent years, which is revolutionising the study of microbial communities, is that of 'omics methodologies that can collect whole systems data from environmental samples, including aquatic ecosystems. Sequencing of marker genes such as 16S rRNA genes for prokaryotes and 18S rRNA for eukaryotes has allowed the 'metabarcoding' of the species diversity within these samples (Mende *et al*. 2013; De Vargas *et al.* 2015). This is a culture-free method that can track microbial species richness across global transects. Further, the sequencing of whole genomes from sampled communities, a practice referred to as metagenomics, has opened the possibility for comparisons of metabolic functions across samples, and metatranscriptomics and metaproteomics can confirm the expression of the genes and infer presence of function, which gives a snapshot of the ecological state of a community, allowing for comparisons in time or under different conditions (e.g. Moran *et al.* 2013). On the other hand, these whole state studies do not in themselves shed light on the specific interactions of the microbes within, which is only possible through physiological experimentation using defined model systems. Nonetheless, there are cross-links between the two. Co-occurrence within metabarcoding datasets can be indicators of active interactions. Similarly, it is possible to assess whether a model laboratory physiological system is widespread by looking for functional genes that reflect it in metagenomic or metatransciptomic compilations.

In the following review we synthesise recent knowledge from these two complementary fields in order to draw conclusions on what they can tell us about mutualisms in aquatic communities, with particular reference to microalgae. Further, we demonstrate how mechanistic understanding of a specific interaction can explain eco-evolutionary aspects, leading to establishment of theoretical principles in the nascent field of microbial ecology.

Whole systems approaches to studying interactions within microbial communities

The advent of affordable high-throughput genomic analyses has uncovered unexpected microbial diversity from a range of biomes, including soil and aquatic environments, and has expanded our horizons of the 'known unknowns' of all extant life on earth (e.g. Pace 1997). Once it was possible to catalogue the genetic diversity into datasets, this opened the door for investigations into the function of genes within organisms as well as between interacting species. Chaffron *et al*. (2010) presented one of the first studies into genomic correlates between diverse sampling sites. The purpose was to uncover consistent associations that could be indicative of ecological interactions. Using network

clustering and co-occurrence analyses the investigators were able to identify putative novel associations; for example, a lineage of cyanobacteria (belonging to the halophilic *Euhalothece*) was observed to be associated with an uncharacterized lineage having no cultivated or named representatives (a monophyletic sister group of the *Psychroflexus* lineage of bacteroidetes).

The Sorcerer Global Ocean Survey (GOS) was the first global scale effort to obtain metagenome DNA sequences from communities of marine microbes (Venter *et al.* 2004). It obtained 6.3 gigabases of DNA sequences from surface-water samples collected along a transect from the Northwest Atlantic to the Eastern Tropical Pacific. Information from the GOS together with 30 smaller scale independent projects contributed towards the Census of Marine Life and the International Census of Marine Microbes (ICoMM). Analysis of the data compiled by ICoMM uncovered the 'rare biosphere' of low-abundance microbial populations that account for most of the observed phylogenetic diversity in the deep ocean, and represent an inexhaustible source of genomic innovation (Sogin *et al*. 2006). Interrogation of microbial metagenomic sequence data collected as part of the Sorcerer II Global Ocean Expedition revealed a high abundance of viral sequences, representing approximately 3% of the total predicted proteins (Williamson *et al*. 2008). Whilst the GOS projects clearly represent an important advance in allowing microbial species composition and diversity to be catalogued on global scales, species composition analysis was confined to prokaryotic organisms, and there was little consideration of interactions between the species.

The Tara Oceans Expedition, which was unveiled in the spring of 2015, superseded all previous efforts in its magnitude, collecting ~35,000 samples across multiple depths at a global scale over a period of three years (Bork *et al.* 2015; de Vargas *et al.* 2015). The extent of the dataset has provided the means to determine the microbial 'interactome'. For example, Lima-Mendez *et al.* (2015) focussed on data from viruses to small metazoans collected at 68 ocean stations. Co-occurrence detection techniques were applied to datasets sub-classified into kingdoms (e.g. data exploring eukaryote diversity focused on the hyper variable V9 region of the 18S rRNA genes only) and also across kingdoms. Machine learning techniques applied to the integrated network predicted 81,590 interactions based on correlations, and found that the majority (~78%) of these were positive, that is to say the presence of one species provided a supporting role for another. Largely, these reflected the nature of trophic cascades, for example zooplankton were commonly associated with their preferred food, and parasites were associated with their hosts, but interestingly, there were also a range of positive interactions observed between planktonic microorganisms belonging to the same size category. The basis for these interactions is unknown, but it is possible that these represent mutualisms. Mining of the Tara interactome confirmed that previously known mutualisms were captured by the network analysis, such as the symbiosis between diatoms and flavobacteria (Jolley & Jones 1977), and dinoflagellate associations with members of Rhodobacterales (*Ruegeria* sp.) (Miller

& Belas 2004). Moreover, certain interactions predicted by the study could then be confirmed experimentally. For example, an endosymbiotic photosymbiotic interaction between an acoel flatworm (*Symsagittifera sp.*) and a green microalga (*Tetraselmis* sp.) was predicted by consistent V9-V9 co-occurrence in the Tara dataset. Fifteen separate collected acoel specimens were then used to validate the interaction, both by laser scanning confocal microscopy and by molecular biology, through single cell 18S rDNA sequencing.

One of the main limitations of metagenomics and metabarcoding however, is that it does not link metabolic capabilities directly to a particular species within the community. At best, analysis of sequence data provides correlations, and confirmation of direct interactions requires physiological and biochemical data, which are challenging to obtain with environmental samples. Metatranscriptomics and metaproteomics can arguably provide a bridge to link the two. Such analyses can demonstrate the executed functions within microbial communities at time of sampling, and often reveal unexpected information. For example whilst metagenomic analysis of biofilms on the hulls of naval vessels indicated a preponderance of bacteria, the majority of proteins identified by metaproteomics using liquid chromatography-tandem mass spectrometry were eukaryotic (Leary *et al.* 2014). Quantitative ^{18}O and iTRAQ analyses, coupled with measurement of photosynthetic pigments confirmed that the communities were dominated by photosynthetic organisms. Similar inferences about the functional dynamics of aquatic communities can be made from metatranscriptomic analyses, and have been used to good effect to monitor successions during algal blooms, to monitor the changing composition of the members of the community (Cooper *et al*. 2014), to establish the effect on the bacterial population associated with the bloom (Wemheuer *et al*. 2014), or to show how nutrients affect niche differentiation (Harke *et al*. 2016).

Studying the metabolic basis for mutualisms between microbes

Despite the considerable information that can be gained from studies of environmental samples, inferring and characterising interactions is at best correlative – a process once described as "*akin to boiling dinner leftovers in a pot for 24 h, pureeing heavily and then trying to attribute any spice or stew fragment back to the original dish or constituent from which it derived*" (Heidelberg *et al.*, 2010). To gain real mechanistic understanding of interactions it is necessary to study defined or model systems that can be co-cultured under controlled conditions, enabling identification of the metabolic foundations for the ecology, demonstrating directly the compounds that are exchanged, and studying the molecular machinery involved in the associations. Importantly, studies of model systems generate testable hypotheses of how the interactions might have evolved, which in turn will allow development of eco-evolutionary principles.

The lifestyle of a unicellular organism is constrained by its individual metabolism. Metabolic requirements that have to be fulfilled to ensure survival but cannot be satisfied by the abiotic environment create the ecological niche for mutualism. For example, microorganisms vary in their capacity to access essential elements from inorganic molecules, frequently requiring them in a specific "bioavailable" form that is dependent on the metabolic activity of other species.

Historically, the study of metabolic requirements in aquatic microorganisms has focused largely on the adaptations of individual species for acquiring nutrients, the latter classified as either macro- (C, N, O, H, P and S) or micro- depending on the quantity required. Micronutrients are required in much lower amounts, but which have essential functions. These include the mineral micronutrients (e.g. Ni, Mo, Zn, Cu, Mn, Fe and Co), involved in protein function, enzyme catalysis and electron transfer reactions, and a range of organic vitamins required for enzyme cofactors. A particularly important example is iron, which is limiting to algal growth in much of the world's oceans, including the equatorial Pacific and Southern Oceans (Behrenfeld *et al.* 1996; Behrenfeld & Kolber 1999). Many bacteria produce siderophores, organic chelators of both Fe^{II} and Fe^{III}, which allow them to sequester this scarce nutrient. Specific mutualistic interactions have been described in which microalgae engage with heterotrophic bacteria to access this source of chelated Fe. For example, several clades of *Marinobacter* found in close association with two major groups of marine algae, dinoflagellates and coccolithophores, produced an unusual siderophore, vibrioferrin (VF), which when chelated to iron undergoes photolysis at rates that are 10–20 times higher than siderophores produced by free-living marine bacteria (Amin *et al.* 2009). Experiments confirmed that photolysis of the chelates increased algal uptake of Fe-VF by >20-fold.

It is widely acknowledged that the two macronutrients most limiting to phytoplankton are nitrogen and phosphorus (Tyrrell 1999). Only certain prokaryotic species ('diazotrophs') are capable of fixing the inert dinitrogen gas into bioavailable forms. In the photic zones of the ocean, these diazotrophs are mainly cyanobacterial species of various morphologies and lifestyles (Zehr & Kudela 2011). A proportion of these exchange fixed nitrogen for photosynthate with diatoms, in a symbiosis referred to as diatom-diazotroph associations (DDAs). For example, *Richelia intracellularis* and *Calothrix rhizosoleniae*, filamentous heterocyst-forming cyanobacteria, have been found in mutualism with several diatom genera, including *Hemiaulus*, *Rhizosolenia* and *Chaetoceros* (Zehr & Ward 2002). The interactions are highly dynamic, demonstrated by the fact that the length, location and number of *Richelia* and *Calothrix* trichomes per diatom partner, and phylogeny of the symbionts, differ in each symbiosis (Jahson *et al.* 1995; Foster & Zehr 2006).

Unicellular nitrogen-fixing cyanobacteria also are thought to interact with microalgae. Using metagenomic techniques, Zehr *et al.* (2003) identified two novel groups of unicellular diazotrophic picocyanobacteria, UCYN-A and UCYN-B (now formally classified as *Crocosphaera watsonii*).

Whole-genome amplification of UCYN-A (Candidatus *Atelocyanobacterium thalassa*) revealed that it lacks genes for photosystem II and the TCA and Calvin cycles, but retains sufficient electron transfer capacity through alternative electron donors to generate energy and reducing power from light (Tripp *et al.* 2010). Since this metabolic capability is insufficient for an autotrophic lifestyle, it was hypothesised that UCYN-A was mutualistic with photosynthetic picoeukaryotes (PPEs), with which it frequently co-occurred. Microscopy of individual cells coupled to secondary ion mass spectrometry (referred to as nano-SIMS), in combinations with halogen in situ hybridization confirmed the interaction visually by demonstrating the delivery of carbon from the PPEs to UCYN-A cells (Thompson *et al.* 2012). This is a powerful example of how single cell imaging techniques can be used to show directly the transfer of metabolites between interacting species. Further investigations revealed that at least two different UCYN-A phylotypes exist (Bombar *et al.* 2014), the clade UCYN-A1, which is predominantly coastal and found in symbiosis with an uncultured small prymnesiophyte, and the clade UCYN-A2, its open ocean relative which is found in symbiosis with the larger *Braarudosphaera bigelowii.* Studying the geographic distribution of UCYN-A1 and UCYN-A2 and symbionts recorded within the Tara dataset revealed a strikingly consistent co-occurrence (Cabello *et al.* 2015), despite the absence of conclusive evidence for physical association in these relationships.

Mutualism based on vitamin exchange has also been demonstrated between microalgae and heterotrophic bacteria (e.g. Croft *et al.* 2005; Wagner-Dobler *et al*. 2010). Over half of microalgal species surveyed across all lineages require cobalamin (vitamin B_{12}), ~22% require thiamine (B_1), and 5% require biotin (B_7) (Croft *et al.* 2006). In the case of thiamine and biotin, auxotrophic species have lost the ability to synthesise the metabolite *de novo*, although often retain parts of the metabolic pathway. For example thiamine-requiring algae are commonly able to survive on one or more biosynthetic intermediates (McRose *et al.* 2014). Bioinformatics analysis of the genome of sequenced organisms has shed light on this. Representatives belonging to the abundant and ubiquitous SAR11 clade of marine chemoheterotrophic bacteria have genes for all the thiamine biosynthetic enzymes except for *thiC,* encoding an enzyme required for the synthesis of the pyrimidine moiety, 4-amino-5-hydroxymethyl-2-methylpyrimidine (HMP) (Carini *et al.* 2014). Interestingly, the authors found that whilst the SAR11 isolate 'Candidatus *Pelagibacter ubique*' was able to grow on HMP, addition of thiamine itself to laboratory cultures did not rescue growth. The authors attribute this to the absence of thiamine transporters in the Ca. *P. ubique* genome. They propose that vitamin cycling mediated through partial precursors and the presence/absence of transporters requires cooperation and interactions within marine microbial communities. Another study by Paerl *et al*. (2015) investigated thiamine cycling between the marine bacterium *Pseudoalteromonas* sp. TW7 and cosmopolitan marine picoalga *Ostreococcus lucimarinus* CCE9901. The bacterium was able to enhance the bioavailability of the pyrophosphorylated form of the vitamin, thiamine pyrophosphate (TPP) to *O.*

lucimarinus. Bacterial phosphatase activity was inferred to metabolise TPP to thiamine monophosphate (TMP), which is better able to support *O. lucimarinus* CCE9901.

Non-requirers of thiamine and biotin are able to synthesise the compounds for their own metabolism. The underlying genetic reason for vitamin B_{12} auxotrophy in algae is quite different. This compound is made only by prokaryotes (Warren *et al.* 2002), in a process that requires more than 20 enzyme-catalysed steps from the common tetrapyrrole precursor. Despite the widespread occurrence of B_{12} requirement amongst algae, there is no phylogenetic relationship between them, with auxotrophs and non-auxotrophs often found within the same genus, for example *Chlamydomonas nivalis* requires B_{12}, but *C. reinhardtii*, does not. Vitamin B_{12} is used primarily as an essential cofactor for the enzyme methionine synthase (METH), the key enzyme of one-carbon (C1) metabolism. An alternative B_{12}-independent methionine synthase (METE) exists, but is catalytically less efficient than METH (Gonzalez *et al.* 1992). Auxotrophs have METH only, whereas non-requirers either have both enzymes, as is the case for *C. reinhardtii*, or METE only, such as *Coccomyxa* sp. C-169 and *Cyanidioschyzon merolae* (Croft *et al.* 2005; Helliwell *et al.* 2011), suggesting that B_{12}-dependence arose through the loss of *METE* numerous times in algal evolution. Helliwell *et al.* (2015) were able to demonstrate this experimentally in *C. reinhardtii*, which lost a functional copy of *METE* in fewer than 500 generations of vitamin B_{12} supplementation, providing an indication of the ease with which this might have occurred in other lineages.

The most comprehensive analysis of vitamin B_{12} concentrations in the ocean was along a transect off the coast of California (USA), which detected little to no vitamin B_{12} in surface ocean waters (Sanudo-Wilhelmy *et al*. 2012), providing evidence that this organic micronutrient might be limiting for algal requirers. Indeed enrichment experiments showed that addition of B_{12} stimulates phytoplankton growth in samples of water collected from temperate coastal areas (Sañudo-Wilhelmy *et al*. 2006; Gobler *et al*. 2007), the Southern Ocean's Gerlache Straight (Panzeca *et al*. 2006), and the Ross Sea (Bertrand *et al*. 2007). Since only bacteria are capable of B_{12} biosynthesis (Warren *et al.* 2002), they must be the ultimate source of the vitamin, and it has been suggested that acquisition is through mutualism with bacterial producers (Croft *et al.* 2005). Several laboratory co-cultures have been described where there is delivery of vitamin B_{12} from bacteria to algae in exchange for photosynthate, including for members of the Chlorophyta (Croft *et al.* 2005; Kazamia *et al.* 2012), Alveolata (Wagner-Dobler *et al.* 2010), and diatom (Heterokontophyta) lineages (Durham *et al.* 2014). Even though many of the co-cultures were the result of artificial combination, they are often exceptionally stable. For example co-cultures of the green alga *Lobomonas rostrata* with the alpha-proteobacterium *Mesorhizobium loti* persist at the same algal:bacterial ratio (of ~ 1:30) over many rounds of sub-culturing (Kazamia *et al.* 2012), and follow predictable dynamics indicative of regulation (Grant *et al.* 2014). Interestingly, not all combinations of bacterial producers and algal

auxotrophs led to stable co-cultures (Kazamia *et al.,* 2012), which is indicative of species-specific associations and not simple metabolic fitting.

Looking for evidence of symbioses between algae and bacteria for vitamin B_{12} acquisition in the natural environment, Bertrand *et al.* (2015) identified *Oceanospirillaceae* ASP10-02a as a possible vitamin B_{12} producer in sea-ice edge microbial communities from the Southern Ocean, which are characterised by extensive blooms of diatoms during the Antarctic summer. Metatranscriptomics data showed that *Oceanospirillaceae* ASP10-02a contributed more than 70% of reads from cobalamin biosynthesis-associated genes, and peptides of one particular enzyme, CbiA, from this species were abundant in a proteomics dataset. Interestingly, they also observed transcripts for proteins involved in uptake and salvage of B_{12} from other bacteria, such as *Methylophaga,* which does not encode the complete biosynthetic pathway. This implies several different phytoplankton–bacterial interactions for provision of this organic micronutrient that might involve both positive and negative feedback loops.

The above examples describe studies of microbial interactions where the metabolic requirement or auxotrophy is known. However, is it possible to infer metabolic exchanges based on known recurring species correlations? The large datasets cataloguing microbial diversity offer this possibility. Zelezniak *et al.* (2015) analysed a compilation of 16S rRNA sequences to obtain the species composition for 1,297 communities from habitats including soil, water and the human gut. For 261 species whose genomes are sequenced and mapped, the authors constructed whole-genome metabolic models using the ModelSEED pipeline (Henry *et al.* 2010). Two metrics, the Metabolic Resource Overlap (MRO) and the Metabolic Interaction Potential (MIP), were used as indicators for competition or possible mutualism, respectively. On average, communities had higher MRO scores than predicted by chance, and intriguingly the authors found high MIP scores in nutritionally rich habitats where incentive for metabolic cross-feeding would be predicted to be lower. The 'Species METabolic interaction ANAlysis' (SMETANA) algorithm was used to identify metabolic exchanges in a community that was modelled as living on minimal medium. It predicted that the most frequently exchanged metabolites were amino acids and sugars. SMETANA was verified against a well-studied three-species bacterial community (Fig. 3A; Miller *et al.* 2010), and co-cultures of the yeast *Saccharomyces cerevisiae* and *C. reinhardtii* (Hom & Murray 2014), summarised in Fig. 3B. The model correctly reproduced known exchanges and identified further possible links in the system, namely that *S. cerevisiae* might also be able to deliver aspartate, glutamate, glutamine and serine to *C. reinhardtii*, although there is as yet no experimental evidence that *C. reinhardtii* can be supported by these nutrients. This highlights an important limitation of bioinformatics approaches when used in isolation, namely that genetic capacity may not accurately predict physiology. A further example of this comes from studies of the production and use of vitamin B_{12} amongst different phytoplankton classes. Cyanobacteria are known producers of B_{12}, and have been proposed to be a major source of

the vitamin to B_{12}-requiring eukaryotic algae (Bonnet *et al*. 2010). However, biochemical and physiological investigations revealed that cyanobacterial species such as strains of the abundant marine genus *Synechococcus* produce a form of B_{12}, so-called pseudocobalamin, which is considerably less bioavailable to eukaryotic algae (Helliwell *et al.* 2016).

From metabolic fitting to complex regulatory dynamics

Whilst exchange of nutrients is considered to be at the heart of associations between marine microorganisms, there is often more complicated ecology involved. For example, Seyedsayamdost *et al.* (2011) describe a 'Jekyll-and-Hyde' association between the bloom-forming coccolithophore *Emiliania huxleyi* and the roseobacter *Phaeobacter gallaeciensis*. During a bloom, *E. huxleyi* can account for 80–90% of the phytoplankton in an area, and up to 60% of the bacterial community belong to the roseobacter clade (Alpha-proteobacteria) (Gonzalez *et al.* 2000), so associations between the two are thought likely. The reduced system described by Seyedsayamdost *et al.* (2011) was brought into the laboratory for analysis and exhibited two stages. Initially *P. gallaciensis* promoted algal growth by biosynthesizing auxins and antibiotics against algal competitors, and *E. huxleyi* released dimethylsulphopropionate (DMSP) in return, which the bacteria used as a carbon source. However, when the algae began to senesce, *P. gallaeciensis* switched its metabolism to the production and secretion of selective algaecides, killing their symbionts and switching to saprotrophy (Fig. 2).

A further hormone-regulated interaction between algae and bacteria was described by Amin *et al.* (2015), who showed that cultures of the ubiquitous diatom *Pseudo-nitzschia multiseries* were regulated by a *Sulfitobacter*-related species of bacterium, referred to as SA11, through the secretion of the auxin indole-3-acetic acid. The association was chosen based on the previous observation of regular associations between different geographic isolates of *Pseudo-nitzschia multiseries* originating from the Atlantic Ocean and the north Pacific Ocean with clades of bacteria belonging to Alpha-proteobacteria (*Sulfitobacter*), Gamma-proteobacteria (*Marinobacter*), Beta-proteobacteria (*Limnobacter*), and Bacteroidetes (*Croceibacter*) (Amin *et al.* 2012). In the laboratory study the *Sulfitobacter* cells were demonstrated to receive organic carbon required for growth, as well as taurine, a sulfonated metabolite from the diatom, and responded to stimulation by DMSP, which diatoms also produce. In return, the bacteria enriched the co-culture medium with ammonium, the preferred nitrogen source for diatoms, by switching their own metabolic preference to nitrate (Amin *et al.* 2015).

The *Pseudo-nitzschia* genus of diatoms includes bloom-forming toxin-producing species, as well as non-toxic representatives. Working with isolates of *Pseudo-nitzschia* collected from the Santa Cruz Wharf in California, Sison-Mangus *et al.* (2014) compared the bacterial associations of toxic *Pseudo-*

nitzschia with non-toxic species. Different diatom strains had their own unique bacterial communities. To investigate the physiology of interactions, transplant experiments were performed, where bacteria associated with one host were co-cultured with a different *Pseudo-nitzschia* isolate. A change of behaviour was shown for certain bacteria that were mutualistic to their native diatom but were commensal or parasitic to foreign hosts. Moreover, the algae exhibited plasticity with regards to domoic acid toxin production, depending on the bacterial species with which they were co-cultured: less domoic acid was produced when in association with their cognate bacteria.

Further evidence of complex regulatory symbiotic interactions between marine microorganisms comes from the work of Decelle *et al.* (2012). Using molecular techniques in culture-free studies it was shown that heterotrophic amoeboid protists belonging to the Acantharia (Radiolaria), which are some of the most abundant grazers in nutrient poor oceans, often engage in photosymbiosis with *Phaeocystis* species, a lineage of haptophyte eukaryotic microalgae ubiquitous in the marine environment. Photosymbioses between heterotrophic hosts and photosynthetic microalgae are widespread and prevalent in the oceanic plankton (Decelle *et al.* 2015). In the model interaction described by Decelle *et al*. (2012), more than 100 individual acantharian protists were isolated from seven geographical locations, and in each case were found to contain *Phaeocystis* cells in endosymbiosis with the host. There was no consistent relationship between the phylogenies of the interacting organisms, implying a facultative mutualism on the part of the *Phaeocystis*. Taken together with the observation that in each case the *Phaeocystis* species were also known to have an extensive free-living population, the authors inferred that the symbiosis was selected for by local biogeography. However, the morphology of the haptophytes during the endosymbiosis was altered significantly, which suggests intricate metabolic associations, which profoundly affect cellular structure and function and remain to be unravelled.

Understanding how symbiosis affects evolution of microorganisms and *vice-versa*

Within the natural environment, the complexity of microbial communities remains challenging to interpret both functionally (i.e. ecologically) and from an evolutionary perspective, particularly in the absence of theoretical models that link microbial ecology to evolutionary theory. A pivotal question about the evolution of interacting microorganisms is whether metabolic exchanges are the outcomes of stable ecological associations, or coincidental by-products of these exchanges constrained by abiotic pressures.

In the oligotrophic oceans, habitats characterised by severe nutrient limitation, there is a strong selective pressure for streamlined genomes, leading to cells with lower DNA and associated protein

content, and therefore lower cellular requirements for P and N (Dufresne *et al.* 2008). The dominant producers in the tropical oligotrophic oceans are picocyanobacteria belonging to the genus *Prochlorococcus* (Olson *et al.* 1990), which have the smallest genomes of any free-living phototroph, with some isolates encoding only ~1,700 genes (Rocap *et al.* 2003). Similarly, the dominant heterotrophs in these regions belong to the SAR11 clade. Ca. *P. ubique*, the first cultured member of the SAR11 clade, has ~1,400 genes (Giovannoni *et al.* 2005). Although it encodes complete biosynthetic pathways for all 20 amino acids it has no pseudogenes, introns, transposons or extrachromosomal elements, and the shortest intergenic spacers yet observed for any cell.

However, reductive genome evolution often causes loss of function. For example, *Prochlorococcus* has a smaller suite of oxidative-stress genes than its closest relatives from the *Synechococcus* genus (Regelsberger *et al.* 2002). Specifically, no *Prochlorococcus* isolates encode catalase-peroxidase (*katG*), a haem-dependent enzyme that is thought to be the primary defence against external hydrogen peroxide (H_2O_2) in cyanobacteria (Morris *et al.* 2008; Scanlan *et al.* 2009). Laboratory experiments demonstrated that *Prochlorococcus* grows better in the presence of heterotrophic bacteria, which reduce the concentrations of H_2O_2 (Morris *et al.* 2008). It was shown that these 'helpers' of *Prochlorococcus* such as *Alteromonas* EZ55 encode *katG,* and were able to deplete H_2O_2 from seawater to levels that were no longer toxic to *Prochlorococcus* (Morris *et al.* 2011).

The work on *Prochlorococcus* led to the proposal of the Black Queen Hypothesis (BQH; Morris *et al.* 2012), a game-theory model for the evolution of mutualism in microbial communities. To our knowledge, it is currently the only theoretical framework that links community microbial ecology to the evolution of dependency in individual species. The BQH (summarised in Fig. 4) posits that communities evolve to sustain a division of labour amongst the individual players. If an essential function is lost from a subset of species, this provides selection pressure for maintaining the provision of a shared 'public good' in helper species. In the case described, *Prochlorococcus* does not produce catalase but is protected from H_2O_2 by mutualist heterotrophs.

The BQH therefore applies only to 'leaky' functions and scenarios, where 'public goods' are traded against a background of severe nutrient limitation. This is in contrast to the prediction by Zelezniak *et al.* (2015) that interactions abound even under conditions of nutrient sufficiency. Moreover, experimental evolution approaches provide evidence that loss of function is the result of nutrient *sufficiency*. Growth of *C. reinhardtii* in elevated CO_2 over 1000 generations resulted in several lines that exhibited a markedly reduced growth rate at ambient levels, and these were no better than the ancestral line in growth at high CO_2 (Collins & Bell 2004). Similarly, loss of *METE* gave rise to a B_{12}-dependent *C. reinhardtii* mutant after ~500 generations in 1 µg/L B_{12} (Helliwell *et al.* 2015), a level >100 higher than ambient (Sanudo-Wilhelmy *et al*. 2012). It is worth mentioning that the evolved B_{12}-dependent clone of *C. reinhardtii* could be grown in co-culture with B_{12}-synthesising

rhizobial bacteria in medium without either B_{12} or a fixed carbon source, in apparently regulated mutualism.

Taking these observations further it is possible to propose a model for the evolution of mutualism among algal lineages generally, using the delivery of B_{12} as example. If a B_{12}-independent alga that can use the vitamin when available (a 'Forager') is in contact for prolonged periods with a bacterium that synthesizes B_{12}, this might cause the loss of the algal *METE*. At this point the alga will be dependent on the bacterial B_{12}-producer, and it may well then change its behaviour to 'farm' the bacterium, providing photosynthate or other nutrients, to ensure its persistence within the vicinity, and therefore a secure supply of an essential nutrient (Fig. 5). This hypothesis, which we term *Foraging-to-Farming,* is an alternative but complementary scenario to the BQH. The Foraging-to-Farming model proposes that mutualism can evolve as accidental consequence of metabolic exchanges, under fluctuating conditions of resource availability. The starting point assumes the ability for both dependent and independent lifestyles, but dependency evolves as a consequence of recurrent ecological associations, which loosen the pressure to maintain the genetic capacity required for independence, thus reinforcing the mutualism. In the evolved *metE* mutant of *C. reinhardtii* (Helliwell *et al.* 2015) dependence is a consequence of a transposable element inserting into a single gene, disrupting its function, but it is possible to envisage multiple pathways to genetic degradation. Once genes are no longer essential, the lack of selection pressure means that they can accumulate deleterious mutations due to drift. Evidence for this comes from the presence of *METE* pseudogenes in several B_{12}-dependent algal species (Helliwell *et al.* 2011).

Such proposed origins for the evolution of dependency are not confined to B_{12}, but have been found for other vitamins (Helliwell *et al.* 2013). For instance, analysis of the genomes of individual members of the symbiotic microbiome of the tsetse fly gut has revealed inactivation of genes involved in thiamine biosynthesis, suggesting symbionts may 'divide the labour' of producing this compound, each contributing intermediates of different branches of the thiamine pathway, and thus sharing the cost of maintaining a supply of the vitamin (Belda *et al*. 2010). Thiamine biosynthesis genes in bacteria are subject to repression by elevated external levels of thiamine (Winkler *et al.* 2002), again providing a possible explanation for ease of their loss, and indeed identification of pseudogenes of thiamine biosynthetic genes in both the eukaryote host and bacterial endosymbiont *Sodalis glossinidius* suggests this metabolic complementation may have occurred recently. Similarly, we propose our theorem could be applied to production of other metabolites, such as amino acids and nucleotides. In an innovative study, Campbell *et al.* (2016) devised a system using stochastic episomal segregation progressively to introduce metabolic auxotrophies into a population of *S. cerevisiae*. They found that despite a gradual loss of prototrophy for the metabolites histidine, leucine, uracil and methionine, self-establishing communities of *S. cerevisiae* could maintain metabolic efficiency

499 through cooperative metabolite exchange (Campbell *et al.* 2016). The observation that cells stopped
500 making these metabolites in the presence of a ready-available source from neighbouring cells could
501 have contributed to the evolution of dependence, thus reinforcing a stable and obligate association, as
502 would be predicted by the Foraging-to-Farming hypothesis.

503 The similarity between the BQH and the Foraging-to-Farming hypothesis is that they both lead to
504 accidental dependency, as an ecological cul-de-sac. However, in the BQH this is driven by an abiotic
505 pressure for streamlined genomes that leads to loss of function. For the Foraging-to-Farming model,
506 the evolution of mutualism is driven by the presence of symbionts. There is an argument that the
507 evolution of dependency is in the interest of the organisms being 'farmed', i.e. the providers of the
508 public good or service. In the example we use, the farming lifestyle of the alga would guarantee a
509 supply of fixed carbon for the vitamin B_{12} producers, and is therefore in the interest of the bacteria.
510 The function does not have to be 'leaky', therefore, but could be selected for. The metabolic burden of
511 producing B_{12} could be outweighed by the quantities of fixed carbon provided, a premise that can be
512 tested experimentally. Examples of rapid symbiont-driven evolution abound in insect-endosymbiont
513 literature (Himler *et al.* 2011; White 2011) but have never been reported for aquatic microorganisms.
514 In contrast, the BQH assumes that the helpers are left burdened by the public function they deliver,
515 essentially having not evolved fast enough to lose it, as other members of the community have. Being
516 the last carriers of this function, loss of helpers from the community would lead to the collapse of the
517 whole system. It is logical to hypothesise from this that their numbers in the community would
518 therefore be as low as possible, without the total collapse of the system, another testable aspect of the
519 model. In both instances, the net outcome is the partitioning of functions between members of stable
520 communities, where the community metabolome, or 'meta-metabolome', is streamlined to reduce
521 redundancies.

522 **Placing evolutionary theories into an ecological context that is relevant for microorganisms**

523 Valid evolutionary theories for planktonic microbial communities must be relevant for the physical
524 environment that the microbes inhabit, whilst accommodating their unique lifestyles. In the absence of
525 complex behaviour (as contrasted with macroorganisms), and a reliance on metabolic function as the
526 main determinant of lifestyles, the boundary between ecology and evolution is blurred. The classical
527 tenet that stable, repeated ecological associations lead to evolutionary outcomes simply may not hold,
528 as evolutionary dynamics, manifested in genetic differences that result in metabolic dependencies,
529 may precede and drive ecological associations. The fast growth rates of the species in question argue
530 further in favour of blurring of the conventional distinction and directionality between ecology and
531 evolution.

Finally, it remains an open question whether physical associations are a pre-requisite for stable associations and symbiosis, a tenet of classical terrestrial ecology (Boucher 1985). The work of Decelle *et al*. (2012) discussed above presents evidence of an ancient symbiosis, dated to approximately 92.8 Mya that requires contact but is nonetheless facultative, with species that are known to engage in endosymbiosis also abundant as free-living organisms. It has also been argued that metabolic exchange could not be species-specific, since this would require recognition of partners at a distance, which is not feasible in the marine environment (Droop 2007). However, developments in physical studies of the marine environment are redefining the concept of an "operational scale" that is relevant to lifestyles of microbes (Stocker 2012). The argument in favour of directed interactions is that marine microorganisms are capable of chemotaxis, sensing and swimming not only along gradients of dissolved micronutrients but also towards other species (Fenchel & Finlay 2004; Grossart *et al.* 2001; Gardes *et al.* 2011), with mean velocities exceeding 60 to 80 µm/s (Hutz *et al.* 2011); for comparison swimming speeds for *Escherichia coli* have been estimated at 15 to 30 µm/s (Chattopadhyay *et al.* 2006). At these speeds, the experienced distance between microorganisms is reduced, and species-specific interactions may be possible. Moreover, the discovery of new physical microstructures in the ocean may play a role in bringing microbes into closer proximity. The role of "marine snow" in creating microhabitats is long recognised (Alldredge & Silver 1988; Kiørboe & Jackson 2001), whilst other structures, such as marine microgels, have been characterised only recently. In the latter case, the bacterial abundance in local patches exceeds the seawater average by 10^4 fold (Verdugo 2012). Finally, it is important to note that a study of interactions on these microscopic scales requires sampling and preservation of biological material collected from the environment that is not so destructive or invasive as to eliminate motility.

Conclusion

Overall, the increasing recognition of the importance of the microbiome as a community of interacting microorganisms undoubtedly owes a great deal to analysis of metagenomes, metatranscriptomes, and metaproteomes. As well as establishing the identity of the many species present in an environmental sample, these methods offer the means to determine the overall metabolic capability of a community. The increasing depth of coverage is now facilitating integrated systems approaches that can indicate hitherto unknown interactions. However, just as interactions between microbes are intricate, complex, and dynamic, ~~it is likely that~~ an understanding of microbial ecology ~~will~~ requires a combination of several interdisciplinary and complementary approaches. Defined combinations of known and well-characterised partners in co-culture in the laboratory have provided insight into the interactions at the molecular and cellular levels, the results of which often explain field observations, as well as indicate potential associations that are then validated in environmental samples. More importantly they are

567 providing the basis to develop principles in ecology that represent the lifestyle and dynamics of
568 microbial communities.
569
570 **Acknowledgements**

571 EK is grateful for funding from UKERC and EU FP7 DEMA project, grant agreement no. 309086.
572 KEH was supported by the UK Biotechnology and Biological Sciences Research Council (BBSRC),
573 grant BB/I013164/1.

574

References

Alldredge, A.L. & Silver, M.W. (1988). Characteristics, dynamics and significance of marine snow. *Progr. Oceanogr.*, 20, 41-82.

Amin, S.A., Green, D.H., Hart, M.C., Küpper, F.C., Sunda, W.G. & Carrano, C. J. (2009). Photolysis of iron–siderophore chelates promotes bacterial–algal mutualism. *PNAS*, 106, 17071-17076.

Amin, S.A., Hmelo, L.R., van Tol, H.M., Durham, B.P., Carlson, L.T., Heal, K.R. *et al*. (2015). Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature*, 522, 98–101.

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

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A. & Thingstad, F. (1983). The ecological role of water column microbes in the sea. *Mar. Ecol. Prog. Ser.*, 10, 257-263.

Behrenfeld, M.J. & Kolber, Z. S. (1999). Widespread iron limitation of phytoplankton in the South Pacific Ocean. *Science*, 283, 840-843.

Behrenfeld, M.J., Bale, A.J., Kolber, Z.S., Aiken, J. & Falkowski, P. G. (1996). Confirmation of iron limitation of phytoplankton photosynthesis in the equatorial Pacific Ocean. *Nature*, 383, 508-511.

Belda, E., Moya, A., Bentley, S. & Silva, F. J. (2010). Mobile genetic element proliferation and gene inactivation impact over the genome structure and metabolic capabilities of *Sodalis glossinidius*, the secondary endosymbiont of tsetse flies. *BMC genomics*, 11, 449.

Bertrand, E.M., Saito, M.A., Rose, J.M., Riesselman, C.R., Lohan, M.C., Noble, A.E. *et al.* (2007). Vitamin B12 and iron colimitation of phytoplankton growth in the Ross Sea. *Limnol. Oceanogr.*, 52, 1079–1093.

Bertrand, E.M., McCrow, J.P., Moustafa, A., Zheng, H., McQuaid, J.B., Delmont, T.O. *et al*. (2015). Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *PNAS*, 112, 9938-9943.

Bombar, D., Heller, P., Sanchez-Baracaldo, P., Carter, B.J. & Zehr, J.P. (2014). Comparative genomics reveals surprising divergence of two closely related strains of uncultivated UCYN-A cyanobacteria. *ISME J.*, 8, 2530-2542.

605 Bonnet, S., Webb, E., Panzeca, C., Karl, D. M., Capone, D. G. & Sanudo-Wilhelmy, S. A.
606 (2010). Vitamin B_{12} excretion by cultures of the marine cyanobacteria *Crocosphaera* and
607 *Synechococcus*. *Limnol. Oceanogr*. 55, 1959–1964.

608 Bork, P., Bowler, C., de Vargas, C., Gorsky, G., Karsenti, E. & Wincker, P. (2015). Tara Oceans
609 studies plankton at planetary scale. *Science*, 348, 873-873.

610 Boucher, D.H. (1985). The biology of mutualism: Ecology and evolution. Oxford University
611 Press, New York.

612 Cabello, A.M., Cornejo-Castillo, F.M., Raho, N., Blasco, D.,Vidal, M., Audic, S. *et al*. (2015).
613 Global distribution and vertical patterns of a prymnesiophyte–cyanobacteria obligate symbiosis.
614 *ISME J*., 10, 693-706.

615 Campbell, K., Vowinckel, J., Mülleder, M., Malmsheimer, S., Lawrence, N., Calvani, E., *et al.*
616 (2015). Self-establishing communities enable cooperative metabolite exchange in a eukaryote.
617 eLife, 4, e09943.

618 Carini, P., Campbell, E.O., Morré, J., Sañudo-Wilhelmy, S.A., Thrash, J.C., Bennett, S.E. *et al*.
619 (2014). Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its
620 distribution in the Sargasso Sea. *ISME J*, 8, 1727-1738.

621 Chaffron, S., Rehrauer, H., Pernthaler, J. & von Mering, C. (2010). A global network of
622 coexisting microbes from environmental and whole-genome sequence data. *Genome Research*,
623 20, 947-959.

624 Chattopadhyay, S., Moldovan, R, Yeung, C. & Wu, X.L. (2006). Swimming efficiency of
625 bacterium *Escherichia coli*. *PNAS*, 103, 13712–13717.

626 Collins, S. & Bell, G. (2004). Phenotypic consequences of 1,000 generations of selection at
627 elevated CO_2 in a green alga. *Nature*, 431, 566-569.

628 Connon, S.A. & Giovannoni, S.J. (2002). High-throughput methods for culturing microorganisms
629 in very-low-nutrient media yield diverse new marine isolates. *Appl. Environ. Microbiol*., 68,
630 3878-3885.

631 Cooper, E.D., Bentlage, B., Gibbons, T.R., Bachvaroff, T.R. & Delwiche, C.F. (2014).
632 Metatranscriptome profiling of a harmful algal bloom. *Harmful Algae* 37, 75-83.

633 Croft, M.T., Lawrence, A.D., Raux-Deery, E., Warren, M.J. & Smith, A.G. (2005). Algae acquire
634 vitamin B_{12} through a symbiotic relationship with bacteria. *Nature*, 438, 90-93.

Croft, M.T., Warren, M.J. & Smith, A.G. (2006). Algae need their vitamins. *Eukaryotic Cell*, 5, 1175-1183.

Decelle, J., Probert, I., Bittner, L., Desdevises, Y., Colin, S., de Vargas, C. *et al.* (2012). An original mode of symbiosis in open ocean plankton. *PNAS*, 109, 18000-18005.

Decelle, J., Colin, S., & Foster, R. A. (2015). Photosymbiosis in marine planktonic protists. In *Marine Protists* (pp. 465-500). Springer Japan.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R, Lara E, Berney C, Le Bescot N, Probert, I & Carmichael, M. (2015). Eukaryotic plankton diversity in the sunlit ocean. *Science*, 348, 1261605.

Dorrell, R.G. & Smith, A.G. (2011) Do red and green make brown? Perspectives on plastid acquisitions within Chromalveolates. *Eukaryot. Cell*, 10, 856–868.

Droop, M. R. (2007). Vitamins, phytoplankton and bacteria: symbiosis or scavenging? *J. Plankton Res.*, 29, 107-113.

Dufresne, A., Ostrowski, M., Scanlan, D.J., Garczarek, L., Mazard, S., Palenik, B.P. *et al*. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biol*, 9, R90.

Durham, B.P., Sharma, S., Luo, H., Smith, C.B., Amin, S.A., Bender, S.J. *et al*. (2014). Cryptic carbon and sulfur cycling between surface ocean plankton. *PNAS*, 112, 453-457.

Falkowski, P.G., Fenchel, T. & Delong, E.F. (2008). The microbial engines that drive Earth's biogeochemical cycles. *Science*, 320, 1034-1039.

Fenchel, T.O.M. & Finlay, B.J. (2004). The ubiquity of small species: patterns of local and global diversity. *Bioscience*, 54, 777-784.

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

Flynn, K.J., Stoecker, D.K., Mitra, A., Raven, J.A., Glibert, P.M., Hansen, P.J., *et al.* (2012). Misuse of the phytoplankton–zooplankton dichotomy: the need to assign organisms as mixotrophs within plankton functional types. *J. Plankton Res.*, doi: 10.1093/plankt/fbs062

Foster, R. A. & Zehr, J. P. (2006). Characterization of diatom–cyanobacteria symbioses on the basis of nifH, hetR and 16S rRNA sequences. *Environ. Microbiol.*, 8, 1913-1925.

Gärdes, A., Iversen, M.H., Grossart, H.P., Passow, U. & Ullrich, M.S. (2011). Diatom-associated bacteria are required for aggregation of *Thalassiosira weissflogii*. *ISME J*, 5, 436-445.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., *et al*. (2005). Genome streamlining in a cosmopolitan oceanic bacterium. *Science*, 309, 1242-1245.

Gobler, C.J., Norman, C., Panzeca, C., Taylor, G.T., & Sañudo-Wilhelmy, S.A. (2007). Effect of B vitamins (B_1, B_{12}) and inorganic nutrients on algal bloom dynamics in a coastal ecosystem. *Aquat. Microb. Ecol.*, 49, 181–194.

Gonzalez, J.C., Banerjee, R.V., Huang, S., Sumner, J.S. & Matthews, R.G. (1992). Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*: two solutions to the same chemical problem. *Biochemistry*, 31, 6045-6056.

González, J.M., Simó, R., Massana, R., Covert, J.S., Casamayor, E.O., Pedrós-Alió, C. *et al*. (2000). Bacterial community structure associated with a dimethylsulfoniopropionate-producing North Atlantic algal bloom. *Appl. Environ. Microbiol.*, 66, 4237-4246.

Grant, M.A., Kazamia, E., Cicuta, P. & Smith, A. G. (2014). Direct exchange of vitamin B_{12} is demonstrated by modelling the growth dynamics of algal–bacterial co-cultures. *ISME J*, 8, 1418-1427.

Grossart, H. P., Riemann, L. & Azam, F. (2001). Bacterial motility in the sea and its ecological implications. *Aquat. Microb. Ecol.*, 25, 247-258.

Harke, M.J., Davis, T.W., Watson, S.B. & Gobler, C.J. (2016). Nutrient-controlled niche differentiation of western Lake Erie cyanobacterial populations revealed via metatranscriptomic surveys. *Environ. Sci. Technol.*, 50, 604-615.

Hartmann, M., Grob, C., Tarran, G. A., Martin, A. P., Burkill, P. H., Scanlan, D. J., & Zubkov, M. V. (2012). Mixotrophic basis of Atlantic oligotrophic ecosystems. *PNAS*, 109, 5756-5760.

Heidelberg, K.B., Gilbert, J.A. & Joint, I. (2010). Marine genomics: at the interface of marine microbial ecology and biodiscovery. *Microb. Biotechnol.*, 3, 531-543.

Helliwell, K.E., Collins, S., Kazamia, E., Purton, S., Wheeler, G.L. & Smith, A.G. (2015). Fundamental shift in vitamin B_{12} eco-physiology of a model alga demonstrated by experimental evolution. *ISME J*, 9, 1446-1455.

Helliwell, K.E., Lawrence, A.D. Holzer, A., Kudahl, U.J., Sasso, S., Kräutler, B., Scanlan, D.J., Warren, M.J. & Smith, A.G. (2016). Cyanobacteria and eukaryotic algae use different chemical

694 variants of vitamin B_{12}. *Curr. Biol.*, in press (accepted).

695 Helliwell, K.E., Wheeler, G.L., Leptos, K.C., Goldstein, R.E. & Smith, A. G. (2011). Insights
696 into the evolution of vitamin B_{12} auxotrophy from sequenced algal genomes. *Mol. Biol. Evol.*, 28,
697 2921–2933.

698 Helliwell, K.E., Wheeler, G.L. & Smith, A.G. (2013). Widespread decay of vitamin-related
699 pathways: coincidence or consequence? *Trends Genet.*, 29, 469-478.

700 Henry, C.S., DeJongh, M., Best, A.A., Frybarger, P.M., Linsay, B. & Stevens, R.L. (2010). High-
701 throughput generation, optimization and analysis of genome-scale metabolic models. *Nat.*
702 *Biotechnol.*, 28, 977-982.

703 Himler, A.G., Adachi-Hagimori, T., Bergen, J.E., Kozuch, A., Kelly, S.E., Tabashnik, B.E. *et al*.
704 (2011). Rapid spread of a bacterial symbiont in an invasive whitefly is driven by fitness benefits
705 and female bias. *Science*, 332, 254-256.

706 Hom, E.F. & Murray, A.W. (2014). Plant-fungal ecology. Niche engineering demonstrates a
707 latent capacity for fungal-algal mutualism. *Science*, 345, 94–98.

708 Hütz, A., Schubert, K. & Overmann, J. (2011). *Thalassospira* sp. isolated from the oligotrophic
709 eastern Mediterranean Sea exhibits chemotaxis toward inorganic phosphate during starvation.
710 *Appl. Environ. Microbiol*., 77, 4412-4421.

711 Jahson, S., Rai, A.N. & Bergman, B. (1995). Intracellular cyanobiont *Richelia intracellularis*:
712 ultrastructure and immuno-localisation of phycoerythrin, nitrogenase, Rubisco and glutamine
713 synthetase. *Mar. Biol.*, 124, 1-8.

714 Joint, I., Mühling, M. & Querellou, J. (2010). Culturing marine bacteria – an essential
715 prerequisite for biodiscovery. *Microb. Biotechnol*., 3, 564-575.

716 Jolley, E. T. & Jones, A. K. (1977). The interaction between *Navicula muralis Grunow* and an
717 associated species of *Flavobacterium*. *Brit. Phycol. J.*, 12(4), 315-328.

718 Jones, R.I. (2000). Mixotrophy in planktonic protists: an overview. *Freshwat. Biol.* 45, 219-226.

719 Kazamia, E., Czesnick, H., Nguyen, T.T.V., Croft, M.T., Sherwood, E., Sasso, S. *et al.* (2012).
720 Mutualistic interactions between vitamin B_{12}-dependent algae and heterotrophic bacteria exhibit
721 regulation. *Environ. Microbiol*., 14, 1466-1476.

722 Kiørboe, T. & Jackson, G.A. (2001). Marine snow, organic solute plumes, and optimal

723 chemosensory behavior of bacteria. *Limnol. Oceanogr.*, 46, 1309-1318.

724 Leary, D.H., Li, R.W., Hamdan, L.J., Hervey, W.J. 4th, Lebedev, N., Wang, Z., *et al.* (2014).
725 Integrated metagenomic and metaproteomic analyses of marine biofilm communities. *Biofouling*.
726 30, 1211-1223.

727 Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F. *et al*. (2015).
728 Determinants of community structure in the global plankton interactome. *Science*, 348, 1262073.

729 Mende, D.R., Sunagawa, S., Zeller, G. & P. Bork (2013). Accurate and universal delineation of
730 prokaryotic species. *Nat. Methods*. 10, 881–884.

731 McKie-Krisberg, Z. M., & Sanders, R. W. (2014). Phagotrophy by the picoeukaryotic green alga
732 *Micromonas*: implications for Arctic Oceans. *ISME J*., 8, 1953-1961.

733 McRose, D., Guo, J., Monier, A., Sudek, S., Wilken, S., Yan, T. *et al*. (2014). Alternatives to
734 vitamin B_1 uptake revealed with discovery of riboswitches in multiple marine eukaryotic
735 lineages. *ISME J*, 8, 2517-2529.

736 Miller, L.D., Mosher, J.J., Venkateswaran, A., Yang, Z.K., Palumbo, A.V., Phelps, T.J. *et al.*
737 (2010). Establishment and metabolic analysis of a model microbial community for understanding
738 trophic and electron accepting interactions of subsurface anaerobic environments. *BMC*
739 *microbiol.*, 10, 149.

740 Miller, T.R. & Belas, R. (2004). Dimethylsulfoniopropionate metabolism by *Pfiesteria*-associated
741 Roseobacter spp. *Appl. Environm. Microbiol.*, 70(6), 3383-3391.

742 Moran, M. A., Satinsky, B., Gifford, S. M., Luo, H., Rivers, A., Chan, L. K., *et al.* (2013). Sizing
743 up metatranscriptomics. *ISME J.*, 7, 237-243.

744 Morris, J.J., Johnson, Z.I., Szul, M.J., Keller, M. & Zinser, E.R. (2011). Dependence of the
745 cyanobacterium *Prochlorococcus* on hydrogen peroxide scavenging microbes for growth at the
746 ocean's surface. *PloS ONE*, 6, e16805-e16805.

747 Morris, J.J., Kirkegaard, R., Szul, M.J., Johnson, Z.I. & Zinser, E.R. (2008). Facilitation of robust
748 growth of *Prochlorococcus* colonies and dilute liquid cultures by "helper" heterotrophic bacteria.
749 *Appl. Environ. Microbiol.*, 74, 4530-4534.

750 Morris, J.J., Lenski, R.E. & Zinser, E.R. (2012). The Black Queen Hypothesis: evolution of
751 dependencies through adaptive gene loss. *MBio*, 3, e00036-12.

752 Muscatine, L. & Porter, J. (1977). Reef corals: Mutualistic symbioses adapted to nutrient-poor
753 environments. *Bioscience,* 27, 454–460.

754 Olson, R.J., Chisholm, S.W., Zettler, E.R., Altabet, M. & Dusenberry, J.A. (1990). Spatial and
755 temporal distributions of prochlorophyte picoplankton in the North Atlantic Ocean. Deep Sea
756 Res. Part A. *Oceanogr. Res. Papers* 37, 1033–1051.

757 Pace, N.R. (1997). A molecular view of microbial diversity and the biosphere. *Science*, 276, 734-
758 740.

759 Paerl, R.W., Bertrand, E.M., Allen, A.E., Palenik, B. & Azam, F. (2015). Vitamin B_1
760 ecophysiology of marine picoeukaryotic algae: Strain-specific differences and a new role for
761 bacteria in vitamin cycling. *Limnol. Oceanogr.*, 60, 215-228.

762 Panzeca, C., Tovar-Sanchez, A., Agustí, S., Reche, I., Duarte, C.M., Taylor, G.T. *et al*. (2006). B
763 vitamins as regulators of phytoplanktondynamics. *Eos Trans. Am. Geophys. Union,* 87, 593–596.

764 Prosser, J.I., Bohannan, B.J.M., Curtis, T.P., Ellis, R.J., Firestone, M.K., Freckleton, R.P. *et al*.
765 (2007). The role of ecological theory in microbial ecology. *Nat. Rev. Microbiol*., 5, 384-392.

766 Regelsberger, G., Jakopitsch, C., Plasser, L., Schwaiger, H., Furtmüller, P.G., Peschek, G.A. *et*
767 *al*. (2002). Occurrence and biochemistry of hydroperoxidases in oxygenic phototrophic
768 prokaryotes (cyanobacteria). *Plant Physiol. Biochem.*, 40, 479-490.

769 Rocap, G., Larimer, F.W., Lamerdin, J., Malfatti, S., Chain, P., Ahlgren, N.A., *et al*. (2003).
770 Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation.
771 *Nature,* 424, 1042-1047.

772 Sachs, J.L. & Wilcox, T.P. (2006) A shift to parasitism in the jellyfish symbiont *Symbiodinium*
773 microadriaticum. *Proc. R. Soc. B.* 273: 425-429.

774 Sañudo-Wilhelmy, S.A., Gobler, C.J., Okbamichael, M., & Taylor, G.T. (2006). Regulation of
775 phytoplankton dynamics by vitamin B_{12}. *Geophys. Res. Lett*. 33, 10–13.

776 Sañudo-Wilhelmy, S.A., Cutter, L.S., Durazo, R., Smail, E.A., Gómez-Consarnau, L., Webb, E.
777 *et al*. (2012). Multiple B-vitamin depletion in large areas of the coastal ocean. *PNAS*, 109, 14041-
778 14045.

779 Scanlan, D.J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W.R. *et al*. (2009).
780 Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.*, 73, 249-299.

781 Seyedsayamdost, M.R., Case, R.J., Kolter, R. & Clardy, J. (2011). The Jekyll-and-Hyde
782 chemistry of *Phaeobacter gallaeciensis*. *Nat. Chem.*, 3, 331-335.

783 Sison-Mangus, M.P., Jiang, S., Tran, K.N. & Kudela, R.M. (2014). Host-specific adaptation
784 governs the interaction of the marine diatom, *Pseudo-nitzschia* and their microbiota. *ISME J*, 8,
785 63-76.

786 Sogin, M. L., Morrison, H.G., Huber, J.A., Welch, D.M., Huse, S.M., Neal P.R., Arrieta, J.M. *et*
787 *al.* (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *PNAS*,
788 103, 12115-12120.

789 Stat, M., Morris, E. & Gates, R.D. (2008). Functional diversity in coral–dinoflagellate symbiosis.
790 *PNAS*, 105, 9256-9261.

791 Stocker, R. (2012). Marine microbes see a sea of gradients. *Science*, 338, 628-633.

792 Thompson, A.W., Foster, R.A., Krupke, A., Carter, B.J., Musat, N., Vaulot, D. *et al*. (2012).
793 Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science*, 337, 1546-
794 1550.

795 Tripp, H.J., Bench, S.R., Turk, K.A., Foster, R.A., Desany, B.A., Niazi, F., Affourtit, J.P. *et al*.
796 (2010). Metabolic streamlining in an open-ocean nitrogen-fixing cyanobacterium. *Nature*, 464,
797 90-94.

798 Tyrrell, T. (1999). The relative influences of nitrogen and phosphorus on oceanic primary
799 production. *Nature*, 400, 525-531.

800 van der Heijden, M.G.A., Bardgett, R.D. & van Straalen, N.M. (2008). The unseen majority: soil
801 microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters*
802 11: 296–310.

803 Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., Wu, D *et al*.
804 (2004). Environmental genome shotgun sequencing of the Sargasso Sea. *Science*, 304, 66-74.

805 Verdugo, P. (2012). Marine microgels. *Annual Review of Marine Science*, 4, 375-400.

806 Wagner-Döbler, I., Ballhausen, B., Berger, M., Brinkhoff, T., Buchholz, I., Bunk, B. *et al.*
807 (2010). The complete genome sequence of the algal symbiont *Dinoroseobacter shibae*: a
808 hitchhiker's guide to life in the sea. *ISME J.*, 4, 61-77.

809 Warren, M.J., Raux, E., Schubert, H.L. & Escalante-Semerena, J.C. (2002). The biosynthesis of

adenosylcobalamin (vitamin B_{12}). *Nat. Prod. Rep.*, 19, 390-412.

Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H.A., Voget, S., Simon, M., *et al.* (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.*, 87, 378-389.

White, J.A. (2011). Caught in the act: Rapid, symbiont-driven evolution. *Bioessays*, 33, 823-829.

Williamson, S.J., Rusch, D.B., Yooseph, S., Halpern, A.L., Heidelberg, K.B., Glass, J.I., *et al.* (2008). The Sorcerer II Global Ocean Sampling Expedition: metagenomic characterization of viruses within aquatic microbial samples. PloS one, 3, e1456.

Winkler, W., Nahvi, A. & Breaker, R.R. (2002). Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature*, 419, 952-956.

Yager, P.L., Connelly, T.L., Mortazavi, B., Wommack, K.E., Bano, N., Bauer, J.E. *et al.* (2001). Dynamic bacterial and viral response to an algal bloom at subzero temperatures. *Limnol. Oceanogr.*, 46, 790-801.

Zehr, J.P. & Kudela, R.M. (2011). Nitrogen cycle of the open ocean: from genes to ecosystems. *Ann. Rev. Mar. Sci.*, 3, 197-225.

Zehr, J.P. & Ward, B.B. (2002). Nitrogen cycling in the ocean: new perspectives on processes and paradigms. *Appl. Environ. Microbiol.*, 68, 1015-1024.

Zehr, J.P., Crumbliss, L.L., Church, M.J., Omoregie, E.O. & Jenkins, B.D. (2003). Nitrogenase genes in PCR and RT-PCR reagents: implications for studies of diversity of functional genes. *Biotechniques*, 35, 996-1013.

Zelezniak, A., Andrejev, S., Ponomarova, O., Mende, D.R., Bork, P. *et al.* (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. *PNAS*, 112, 6449-6454.

835 **Figure captions**

836

837 **Fig. 1. Unicellular organisms dominate the eukaryotic lineage of the tree of life.** A schematic
838 diagram of the eukaryotic tree of life based on Dorrell & Smith (2011). Multicellularity has evolved
839 only seven times (highlighted in blue); all other lineages are essentially microbial.

840

841 **Fig. 2 Schematic of phytoplankton lifestyles.** The photic zone is dominated by the primary
842 producers prokaryotic cyanobacteria and eukaryotic algae, which fix CO_2 by oxygenic photosynthesis,
843 but there are several layers of complexity, examples of which are shown here. Ecologically significant
844 groups, including diatoms, bloom and are subject to viral lysis and grazing by heterotrophic
845 zooplankton. This releases organic nutrients into solution, which both cyanobacteria and algae can
846 utilise via mixotrophy. In addition, many algae such as dinoflagellates consume bacterial prey via
847 phagotrophy, resulting in net CO_2 release and O_2 consumption. In addition to these trophic processes
848 is a complex series of interactions or symbioses, which have shaped the ecology and evolution of
849 microbes in aquatic communities. Many examples of mutualism are known, where algae supply fixed
850 carbon (photosynthate) in exchange for specific nutrients such as vitamins. Parasitism can also arise,
851 as in the case of senescing haptophytes, where bacterial partners that were initially mutualistic
852 produce algaecides to accelerate the process, indicating that interactions can be dynamic. Intimate
853 physical associations, in the form of endosymbiosis, such as between the amoeboid protistan
854 Radiolarians and haptophytes, are also frequent.

855

856 **Fig. 3 Microbial systems with well-characterised metabolic exchanges**. These have been used to
857 validate the SMETANA algorithm devised by Zelezniak *et al.* (2015). In each case solid arrows
858 represent known exchanges tested in physiological studies and dotted arrows mark potential novel
859 interactions predicted by SMETANA. Panel A) A three-species community based on cellobiose
860 degradation (Miller *et al.* 2010). SMETANA predicts that pyruvate and hydrogen may also be
861 delivered to *G. sulfurreducens* by *C. cellulolyticum.* Panel B) An algal-fungal mutualism between *C.*
862 *reinhardtii* and *Saccharomyces cerevisiae* (Hom & Murray 2014). As well as refining the likely forms
863 of N and S that are exchanged, SMETANA predicted that during co-culture growth aspartate,
864 glutamine and serine could also be delivered from the yeast to the alga.

865

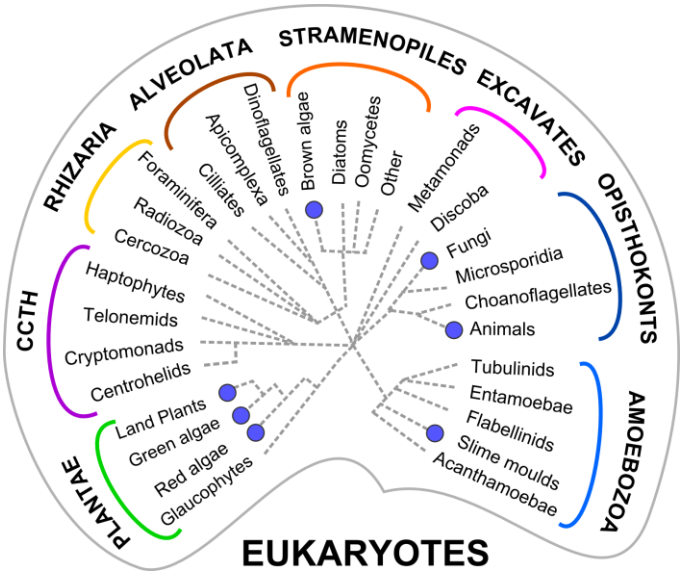
866 **Fig. 4 Black Queen Hypothesis (BQH).** The Black Queen refers to the Queen of Spades in the card
867 game Hearts, where players try to avoid ending up with this card, since it carries the greatest number

of negative points. In the microbial community illustrated, the ability to detoxify H_2O_2 is analogous to the Queen of Spades, because it requires *katG*, an enzyme with a high Fe cost. Helper bacteria, such as *Alteromonas* sp. act as a sink for H_2O_2 and keep concentrations low enough for *Prochlorococcus* and other members of the aquatic community, including the numerically dominant heterotrophic Candidatus *Pelagibacter ubique*, to survive. *Prochlorococcus* is the photosynthetic producer in the system, fixing carbon that is made available to the other species in the community.

Fig. 5 Foraging-to-Farming hypothesis. Algae with both isoforms of methionine synthase are unhindered in the absence of B_{12}. However, they remain facultative opportunists and will use the vitamin if it is available, much like foragers that maximise use of sporadic environmental opportunities. If a reliable source of the vitamin is available for sufficient time, for example from surrounding loosely-associated bacteria, the *METE* gene will be repressed, and may be lost. However, environmental conditions may not be stable so vitamin B_{12} may become scarce. In such circumstances algae that release a carbon source will actively maintain a viable population of bacteria, and will consequently have a selective advantage over those that do not. Here, "farming" the bacteria for the resource becomes an evolutionary stable strategy. Eventually *METE* is lost altogether and the alga becomes absolutely dependent on bacteria for survival turning a previously loose interaction into an obligate one.

887 **Figure files**

888 **Fig 1.**



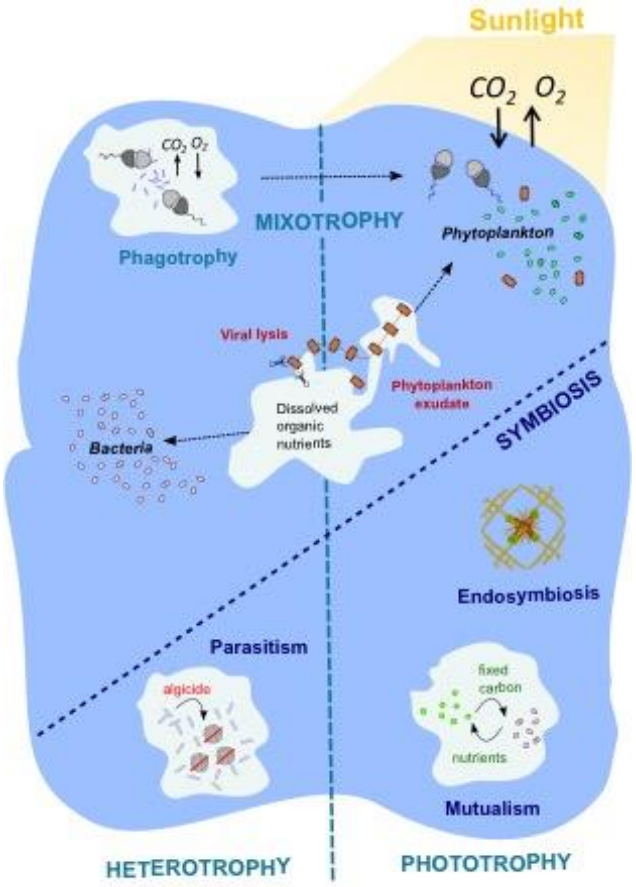
889

890

891

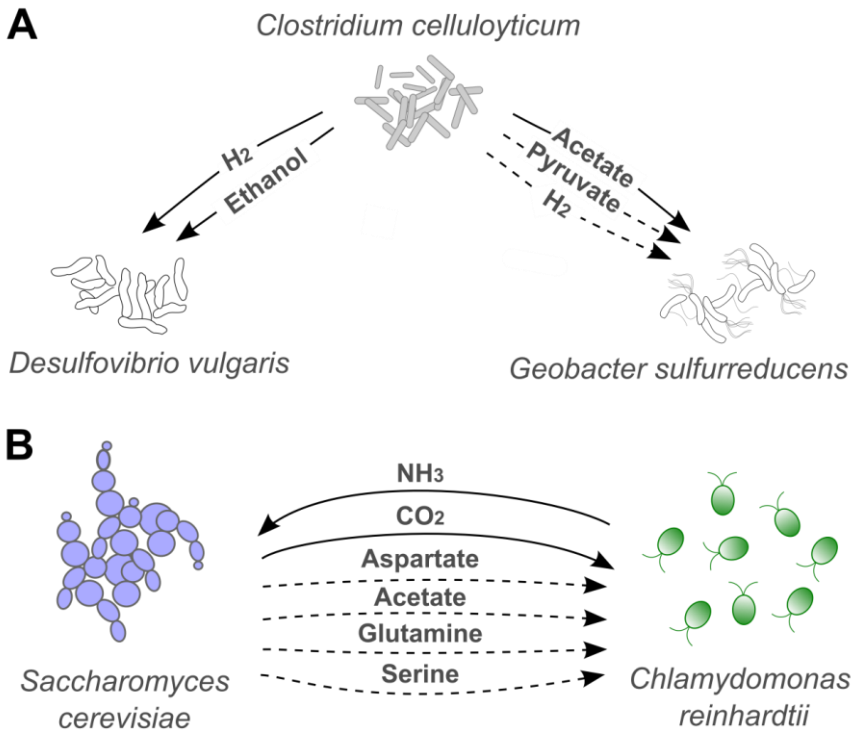
892

893 **Fig. 2**


Sunlight
CO_2 O_2
CO_2 O_2
MIXOTROPHY
Phagotrophy
Phytoplankton
Viral lysis
Phytoplankton exudate
Dissolved organic nutrients
Bacteria
SYMBIOSIS
Endosymbiosis
Parasitism
algicide
fixed carbon
nutrients
Mutualism
HETEROTROPHY
PHOTOTROPHY

894

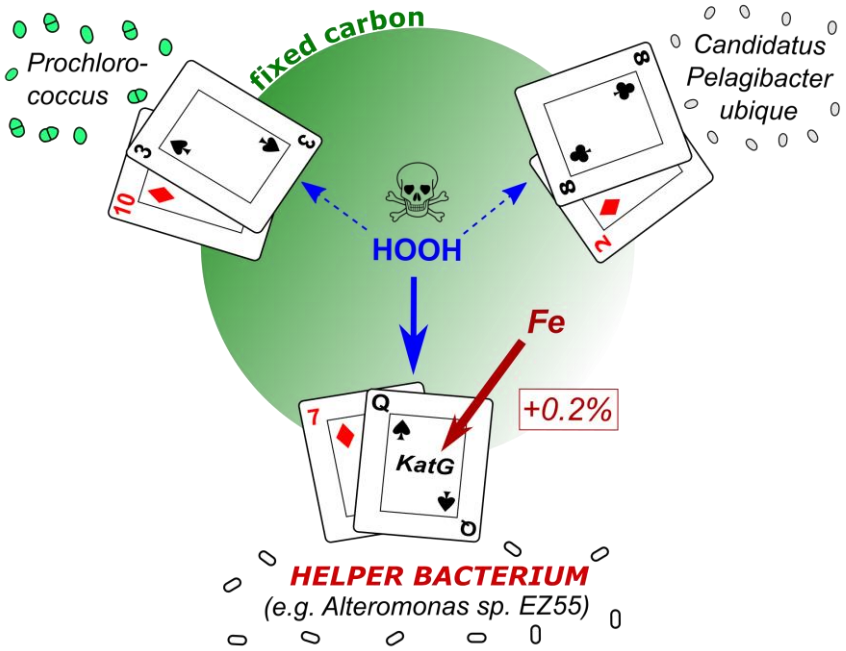
895 **Fig 3.**



896

897

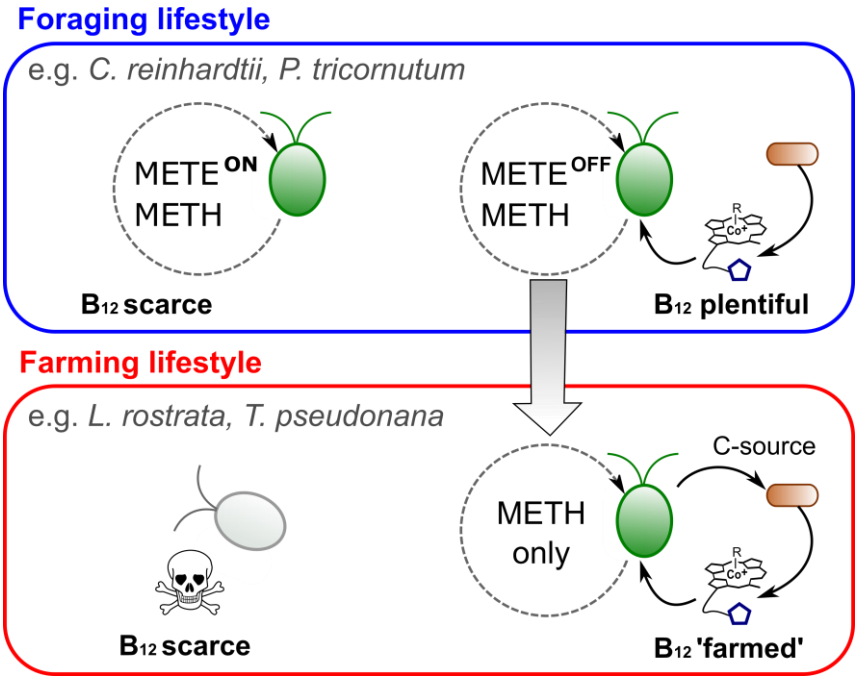
898 **Fig 4.**



899

900

901 **Fig 5.**


Foraging lifestyle
e.g. C. reinhardtii, P. tricornutum
METE ON
METH
METE OFF
METH
B12 scarce
B12 plentiful
Farming lifestyle
e.g. L. rostrata, T. pseudonana
C-source
METH
only
B12 scarce
B12 'farmed'

902