

**Title: Insights into the Evolution of Vitamin B<sub>12</sub> Auxotrophy from Sequenced Algal Genomes**

Research Article

Katherine E. Helliwell,<sup>[a]</sup> Glen L. Wheeler,<sup>[b,c]</sup> Kyriacos C. Leptos,<sup>[d]</sup> Raymond E. Goldstein,<sup>[d]</sup> and Alison G. Smith<sup>[a\*]</sup>

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

*[b] Plymouth Marine Laboratory, Plymouth PL1 3DH, UK*

*[c] Marine Biological Association of UK, The Laboratory, Citadel Hill, Plymouth, PL1 2PB, UK*

*[d] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, UK*

\*To whom correspondence should be addressed, email: [as25@cam.ac.uk](mailto:as25@cam.ac.uk), tel.: 01223 333952

KEYWORDS: ALGAE, COBALAMIN, PSEUDOGENE, METHIONINE SYNTHASE, VITAMIN B<sub>12</sub>

**Running title:** Evolution of Vitamin B<sub>12</sub> Auxotrophy

## Insights into the Evolution of Vitamin B<sub>12</sub> Auxotrophy from Sequenced Algal Genomes

### Abstract

Vitamin B<sub>12</sub> (cobalamin) is a dietary requirement for humans since it is an essential cofactor for two enzymes, methylmalonyl-CoA mutase and methionine synthase (METH). Land plants and fungi neither synthesize or require cobalamin because they do not contain methylmalonyl-CoA mutase, and have an alternative B<sub>12</sub>-independent methionine synthase (METE). Within the algal kingdom, approximately half of all microalgal species need the vitamin as a growth supplement, but there is no phylogenetic relationship between these species, suggesting that the auxotrophy arose multiple times through evolution. We set out to determine the underlying cellular mechanisms for this observation by investigating elements of B<sub>12</sub> metabolism in the sequenced genomes of 15 different algal species, with representatives of the red, green and brown algae, diatoms, and coccolithophores, and including both macro- and microalgae, and from marine and freshwater environments. From this analysis, together with growth assays, we found a strong correlation between the absence of a functional *METE* gene and B<sub>12</sub> auxotrophy. The presence of a *METE* unitary pseudogene in the B<sub>12</sub>-dependent green algae *Volvox carteri* and *Gonium pectorale*, close relatives of the B<sub>12</sub>-independent *Chlamydomonas reinhardtii*, suggest that B<sub>12</sub>-dependence evolved recently in these lineages. In both *C. reinhardtii* and the diatom *Phaeodactylum tricornutum*, growth in the presence of cobalamin leads to repression of *METE* transcription, providing a mechanism for gene loss. Thus varying environmental conditions are likely to have been the reason for the multiple independent origins of B<sub>12</sub> auxotrophy in these organisms. Since the ultimate source of cobalamin is from prokaryotes, the selective loss of *METE* in different

algal lineages will have had important physiological and ecological consequences for these organisms in terms of their dependence on bacteria.

## Introduction

Vitamins are essential for all organisms because they provide the precursors to enzyme cofactors important for metabolism. Animals must obtain these organic micronutrients in their diet, but plants and microorganisms generally synthesize *de novo* the cofactors they need (Smith et al. 2007). It is thus perhaps surprising that, despite their photosynthetic lifestyle, many algae exhibit vitamin auxotrophy, that is the inability to synthesize an organic nutrient essential for growth. Over half of all microalgal species require an exogenous supply of vitamin B<sub>12</sub> (cobalamin), while just over 20% require vitamin B<sub>1</sub> (thiamine) and a smaller proportion (5%) require biotin (vitamin B<sub>7</sub>) (Croft, Warren and Smith 2006). However, vitamin auxotrophy is not phylogenetically related, but instead distributed throughout the algal clades implying that it has arisen multiple times throughout evolution (Croft et al. 2005). Biotin and thiamine auxotrophy are the result of the loss of one or more of the biosynthetic enzymes to make the vitamin (Croft, Warren and Smith 2006), but for cobalamin, this does not appear to be the case. Those algae that are not dependent on B<sub>12</sub> do not contain it unless it is supplied in the medium (Croft et al. 2005). In fact, B<sub>12</sub> biosynthesis appears to be confined to prokaryotes, and there is no evidence that any eukaryote is capable of producing the vitamin (Warren et al. 2002). Interestingly not all bacteria make B<sub>12</sub> – some, such as *E. coli*, will use it if it is available, whilst other species have no enzymes that use B<sub>12</sub>. A notable example is the most abundant prokaryote in the ocean SAR11 clade *Pelagibacter* (Giovannoni et al. 2005).

Cobalamin is a complex Co<sup>2+</sup>-containing modified tetrapyrrole that acts as a cofactor for enzymes involved in C1 metabolism and certain radical reactions. In prokaryotes, there are over twenty enzymes that have a cobalamin cofactor (Marsh 1999), three of which have been

found in eukaryotes. In humans there are just two B<sub>12</sub>-dependent enzymes, methylmalonyl-CoA mutase (MCM), involved in odd-chain-fatty-acid metabolism in the mitochondria, and methionine synthase (METH), which catalyses the C1 transfer from methyltetrahydrofolate to homocysteine to make methionine. An alternative B<sub>12</sub>-independent form of methionine synthase (METE) is found in land plants and fungi, so they have no requirement for the cofactor, and do not synthesize it. *E. coli* possesses both forms of methionine synthase, and can utilize METH if B<sub>12</sub> is available in the environment, otherwise it uses METE. Similarly, there are both B<sub>12</sub>-independent and B<sub>12</sub>-dependent forms of ribonucleotide reductase (type I & III RNR are B<sub>12</sub> independent, type II is B<sub>12</sub> dependent), involved in deoxyribose biosynthesis (Hamilton 1974; Carell, Seeger and Ju 1980). Many prokaryotes encode both types of enzyme, and again appear to switch between them depending on the environment and the availability of cobalamin. All eukaryotes have the type I isoform, whilst *Euglena gracilis*, an excavate protist, is reported to have B<sub>12</sub>-dependent RNR II as well (Hamilton 1974).

An initial search for the presence of genes encoding B<sub>12</sub>-dependent enzymes in four sequenced algal genomes suggested that B<sub>12</sub> auxotrophy may be related to the form of methionine synthase present (Croft et al. 2005). The B<sub>12</sub>-independent red alga *Cyanidioschyzon merolae* has *METE* only. The green alga *Chlamydomonas reinhardtii* and the diatom *Phaeodactylum tricornutum* have both *METE* and *METH* and are B<sub>12</sub> independent, whereas another diatom *Thalassiosira pseudonana*, which contains *METH* only, is B<sub>12</sub> dependent. However, algae are an extremely diverse group of organisms, so a much broader study is necessary to allow a definitive conclusion to be drawn.

The dominant algal lineages have arisen from a complex succession of endosymbiotic events. Over a billion years ago, engulfment of a photosynthetic cyanobacterium by a heterotrophic protist gave rise to the basal algal groups: green and red algae, and the glaucocystophytes (Delwiche 1999; Reyes-Prieto, Weber and Bhattacharya 2007; Gould,

Waller and McFadden 2008). Secondary endosymbiotic events then followed, in which one or more heterotrophic organisms engulfed either a red or a green alga, giving rise to algae with complex plastids (Moreira and Philippe 2001; Gould, Waller and McFadden 2008). These include the chromalveolate supergroup, which encompasses important marine phytoplankton such as diatoms, coccolithophores, and dinoflagellates, as well as brown algae (as defined in Hackett et al. 2007).

Here we perform a comprehensive survey of B<sub>12</sub>-dependent enzymes and associated proteins in 15 algal species with completed genome sequences. Our results provide strong evidence that the major determinant for the B<sub>12</sub> requirements of algae relates to the isoform(s) of methionine synthase that they possess, and that multiple independent losses of a functional *METE* is the principal factor underlying the evolution of B<sub>12</sub> auxotrophy.

## **Materials and Methods**

### **Strains and Growth Conditions**

The algal strains used were either from culture collections or gifts from colleagues (see Supplementary Table 1, for more information and details of growth conditions). *Salmonella enterica* (AR3621) was a gift from Professor Martin Warren (University of Kent, UK) and was grown as described by Raux et al. (1996).

### **Sequence Similarity Searches and Putative Sequence Analysis and Verification**

BLASTp and TBLASTn (Altschul et al. 1990) sequence similarity searches were performed to assess the presence of each protein in the described algal genomes. The organisms and sequence IDs of the proteins that were used to perform these searches are as follows: *C.*

*reinhardtii* METE (XP\_001702934) and METH (76715), *Homo sapiens* MTRR (AAF16876.1), MCM (AAA59569.1), and CBLB (AAH11831.1), *E. coli* CBLA (P27254.2), *E. gracilis* RNR II (Q2PDF6.1) (protein IDs of all putative hits can be found in Supplementary Table 2). The genome version of each species in which the searches were carried out are as follows: *C. reinhardtii* (v4) (Merchant et al. 2007), *Volvox carteri* f. *nagariensis* (v1) (Prochnik et al. 2010), *Coccomyxa* sp. C-169 (v1) ([http://genome.jgi-psf.org/Coc\\_C169\\_1/Coc\\_C169\\_1.home.html](http://genome.jgi-psf.org/Coc_C169_1/Coc_C169_1.home.html)), *Chlorella variabilis* NC64A (v1) (Blanc et al. 2010), *Micromonas pusilla* (CCMP 1545) (v2) (Worden, Panaud and Piegu 2009), *Ostreococcus tauri* (OTH 95) (v2), *O. lucimarinus* (v2) (Palenik et al. 2007) and *Ostreococcus* sp. RCC809 (v2) ([http://genome.jgi-psf.org/OstRCC809\\_2/OstRCC809\\_2.home.html](http://genome.jgi-psf.org/OstRCC809_2/OstRCC809_2.home.html)), *P. tricornutum* (CCAP 1055/1) (v2) (Armbrust et al. 2004; Bowler et al. 2008), *Fragilariopsis cylindrus* (CCMP 1102) (v1) (<http://genome.jgi-psf.org/Fracy1/Fracy1.home.html>), *T. pseudonana* (CCMP 1335) (v3) (Armbrust et al. 2004), *Aureococcus anophagefferens* (CCMP 1984) (v1) (<http://genome.jgi-psf.org/Auran1/Auran1.home.html>), *Emiliana huxleyi* (CCMP 1516) (v1) (<http://genome.jgi-psf.org/Emihu1/Emihu1.home.html>). All of the above can be found at <http://genome.jgi-psf.org/>. The *Ectocarpus siliculosus* (CCAP 1310) (v1) sequence (Cock et al. 2010) can be found at <http://bioinformatics.psb.ugent.be/genomes/view/Ectocarpus-siliculosus>, and that for *C. merolae* (Matsuzaki et al. 2004) at <http://merolae.biol.s.u-tokyo.ac.jp/>.

To verify the identity of all putative orthologous proteins we employed the following techniques: 1) multiple sequence alignment using Jalview (Waterhouse et al. 2009) to check visually each gene model, 2) Genewise 2.0 analysis (Birney, Clamp and Durbin 2004) to identify missing introns in poor gene models and 3) Pfam (24.0) analysis (Finn et al. 2008) to identify conserved functional domains in each protein (METE: PF01717, PF08267, METH:PF02965, PF02965, PF02607, PF00809, PF02574, RNR II: PF09747, PF02867,

MCM: PF01642, PF02310, MTRR: PF00258, PF0667, PF00175, CBLA: PF3308, PF00875, CBLB: PF01923).

### **Treatment of Algae Prior to B<sub>12</sub> Growth Assays**

Bacteria are a potential source of B<sub>12</sub>, so prior to the growth assays, cultures were either obtained axenic (*P. tricornutum* and *F. cylindrus*) or treated to remove contaminating bacteria. *V. carteri* EVE was made axenic by several serial washings of spheroids in 6-well microplates using a sterile pipette. The final washed spheroids were tested for the presence of bacterial contamination on LB plates. *G. pectorale* was made axenic by serial streak outs on SVM agar plates followed by incubation at 28 °C. This was done multiple times until bacterial colonies were growing distinctly from algal colonies. Finally, an isolated algal colony was used to generate an axenic stock. *Ostreococcus* sp. RCC809, *O. tauri* (OTH 95), *E. huxleyi*, and *M. pusilla* were treated with antibiotics. Cultures were grown for one week with penicillin (1 mg/mL), kanamycin (25 mg/mL) and neomycin (20 mg/mL). An additional sub-culture without antibiotics for approximately 6 days was then grown, before inoculating the experimental vitamin B<sub>12</sub> growth assays. For *E. huxleyi*, before the antibiotic treatment, cells were filtered on a 1 micron filter (to attempt to remove external bacteria) and then decalcified to reduce the possibility that bacteria may have been present underneath the coccoliths (calcifying form only). This was achieved by addition of sterile hydrochloric acid to pH 4, followed by restoration to pH 8.5 by the addition of sterile sodium hydroxide). Cultures were stained with the nucleic acid-specific stain 4',6- diamidino-2-phenylindole (DAPI) (1 ng/mL, 5 minutes at 20°C), placed in a microscopy dish and viewed under epifluorescent illumination (excitation 330- 380 nm, emission above 420 nm) using an



Eclipse E1000 microscope (Nikon, Tokyo, Japan). Bacteria were clearly visible in non-treated control cultures, but were not seen in the antibiotic treated cultures used in this study.

### **Assessment of Vitamin B<sub>12</sub> Requirements of Algal Species**

The vitamin B<sub>12</sub> requirements of the algal species (treated as above) were assessed by growing each alga in the appropriate medium, plus or minus B<sub>12</sub>. The concentration of B<sub>12</sub> used in the plus condition was as described in the media appropriate to each alga (see Supplementary Table 1). Aliquots (1 mL) of culture were then transferred to fresh media within 5-21 days depending on the speed of growth, which varies between species. Up to 5 sub-cultures were carried out, or until the algae had died in the experimental condition (minus B<sub>12</sub>) and not in the control (plus B<sub>12</sub>) as described in Croft et al. 2005. Three biological replicates were carried out for each condition. If the cultures were still viable after B<sub>12</sub> starvation they were visually checked with DAPI, and/or plated onto LB plates (for freshwater species) or high salt LB (for marine species).

### **B<sub>12</sub> Bioassay**

B<sub>12</sub> bioassays on cell lysates of *E. huxleyi* were carried out as described by Raux et al. (1996), using *Salmonella enterica* (AR3621). For preparation of the algal cell lysates, cells were washed 3 times and re-suspended in 0.9% NaCl. The cells were then subjected to three freeze-thaw cycles and placed at 100°C for 15 minutes before cooling to room temperature.

### **Molecular Methods**

DNA was extracted from algal cells grown in liquid culture using the phenol-chloroform method described in Newman et al. 1990. Total RNA was extracted as described by Witman et al. (1972). RNA quality was assessed using a Nanodrop Spectrophotometer (ND-1000). A 5 µg sample of RNA was treated with RNase-free DNase (Promega) for 1 hour at 37°C to remove contaminating DNA. Reverse transcription was performed using Superscript TM II transcriptase (Invitrogen) with oligo dT primers according to the manufacturer's instructions. PCR was used to amplify sequences from genomic DNA and cDNA using Taq DNA polymerase (Bioline) (see Supplementary Table 3, for details of primers). For the semi-quantitative RT-PCR analysis in *P. tricornutum* PCR amplification was carried out for 26–38 cycles, to determine the optimal cycle number for each primer pair and allow comparisons between samples (Supplementary Figure 4). The following cycle numbers for each primer set was used throughout: *METE* (35 cycles), *METH* (33 cycles) and *HISTONE H4* (33 cycles).

### **Phylogenetic Analysis of Methionine Synthases**

Methionine synthase sequences were selected from both prokaryotes and other non-algal eukaryotes. In order to ensure that the dataset included appropriate prokaryote sequences, BLASTp searches of the GenBank non-redundant database using methionine synthase genes identified in algal genomes were used to select the most similar sequences in prokaryotes. An alignment of *METE* or *METH* was constructed using MUSCLE (Edgar 2004) and manually corrected using BioEdit (Hall 1999) where appropriate, to ensure only unambiguous residues were compared. Maximum likelihood phylogenetic analysis was performed using PhyML within the Bosque software package (Ramirez-Flandes and Ulloa 2008), based on the WAG substitution matrix (Whelan and Goldman 2001). One hundred bootstrap replicates were run for each analysis.

## Results

### Survey of B<sub>12</sub>-dependent Enzymes in Sequenced Algal Genomes

We examined 15 algal genomes, including one or more representatives from the Rhodophyta (red algae), Chlorophyta (green algae), diatoms, Haptophyta and brown macroalgae, for the presence of enzymes that require vitamin B<sub>12</sub>. We also searched for accessory proteins involved in supplying the cofactor to these B<sub>12</sub>-dependent enzymes. MCM utilises the adenosylated form of cobalamin (AdoCbl) and requires two accessory proteins, CBLA for vitamin B<sub>12</sub> transport into the mitochondria, and CBLB for AdoCbl synthesis (Dobson et al. 2002a; Dobson et al. 2002b) (see Figure 1). Conversely, METH uses methylcobalamin (MeCbl), and requires the molecular chaperone methionine synthase reductase (MTRR) to regenerate the methylated cofactor (Dobson et al. 2002a, Yamada et al. 2006). BLAST searches were performed with the protein sequences as queries (as described in Materials and Methods) using default parameters. All hits were verified by multiple sequence alignment and Pfam analyses, (see Supplementary Table 2, for accession number details).

We found that all species have the B<sub>12</sub>-dependent methionine synthase, *METH*, except *Coccomyxa* sp. C-169 and *C. merolae*, (Table 1). Although partial sequences were found for *A. anophagefferens* and *F. cylindrus*, these appear to be the result of gaps in the genome assembly. All genomes contained the gene encoding MTRR, the accessory protein for METH, even those that did not encode this B<sub>12</sub>-dependent isoform of methionine synthase. In contrast, the gene encoding the vitamin B<sub>12</sub>-independent isoform, *METE*, was present in only 7 of the genomes, and absent altogether from another 7. For *V. carteri* a gene was identified that had sequence similarity to *METE* of *C. reinhardtii*. However, although the genome assembly was complete in this region, when the predicted amino-acid sequence was obtained

using GeneWise 2.0, there were numerous frameshifts and regions of missing sequence, implying that this was a pseudogene. This is considered in more detail below.

MCM was found in the genomes of all species representing the Chromalveolata, but only in one of the species with simple plastids, the green alga *Chlorella variabilis* (NC64A). However the distribution of the accessory proteins was not identical. *A. anophagefferens* does not appear to encode *CBLA*, indicating that MCM would not function in this species, and *CBLB* was found in several of the green algae that do not encode MCM, suggesting that this gene was retained after the loss of the associated B<sub>12</sub>-dependent enzyme. The third B<sub>12</sub>-dependent enzyme RNR II, previously reported in *E. gracilis* (Hamilton 1974), was present only in *E. siliculosus*. This alga also encodes both subunits of B<sub>12</sub>-independent RNR I (Cock et al. 2010).

### **Assessment of Algal Requirements for Vitamin B<sub>12</sub>**

To interpret this genomic dataset, the B<sub>12</sub> requirement of each alga must be established. For eight species this information was already available in existing literature. For the other algal species we performed standard vitamin B<sub>12</sub> growth assays to examine growth with or without this cofactor. In order to avoid false positive results the assays were designed to limit all potential sources of B<sub>12</sub>. Artificial seawater (ASW) was used where possible (see Supplementary Materials Table 1) to exclude a potential source of B<sub>12</sub> from natural seawater, and all cultures used were either obtained axenic, or treated with an antibiotic cocktail prior to the growth assay. Cultures were visually inspected using DAPI-staining and epifluorescence microscopy to look for the presence of bacteria.

After 2-4 sub-cultures it became clear that *V. carteri*, *Ostreococcus* sp. RCC809, *O. tauri* and *M. pusilla* were B<sub>12</sub> dependent (Figure 2A-D). The dramatic decrease in growth within two sub-cultures suggests our precautions to avoid contaminating sources of B<sub>12</sub> were sufficient and that false positives were unlikely. In contrast, for the Arctic diatom *F. cylindrus*, and the haptophyte *E. huxleyi*, no difference was seen in growth with or without B<sub>12</sub>.

### **Methionine Synthase Isoform Determines B<sub>12</sub>-dependence**

An analysis of the B<sub>12</sub>-dependent enzymes in the six algal species that require B<sub>12</sub> for growth (*V. carteri*, *Ostreococcus* sp. RCC809, *O. tauri*, *M. pusilla*, *T. pseudonana* and *A. anophagefferens*) indicates that *METH* is present in all species, whereas only two species possess *MCM*. Moreover, *METE* is absent in each of these B<sub>12</sub>-dependent algal species, indicating that *METH* would be the sole source of methionine synthase activity. Conversely, with one exception, *E. huxleyi*, *METE* is universally present in the B<sub>12</sub>-independent algae, indicating that *METE* enables the organism to live in the absence of the cofactor. Although *E. siliculosus* contains type II RNR, it also encodes B<sub>12</sub>-independent RNR I, explaining the observation that the alga itself does not require B<sub>12</sub> for growth (Boalch 1961). From these observations we conclude that the isoform of methionine synthase is the key factor determining the nutritional requirement for B<sub>12</sub> in algae. However, two of our surveyed species warrant closer scrutiny. Firstly, it must be determined whether the partial *METE* sequence in *V. carteri* is likely to encode an active gene product, and secondly we need to understand how *E. huxleyi* is able to grow in the absence of both *METE* and exogenous B<sub>12</sub>.

### **Characterization of *METE* genes in *V. carteri* and related algae**

As explained above, sequence similarity searches identified a putative *METE* sequence in *V. carteri* but it appeared to be a pseudogene. Closer inspection showed that elements of the N-terminus are missing, there is an in-frame stop codon (white asterisk, Figure 3A) and multiple deletions leading to seven frameshifts (dashed white lines, Figure 3A). The C-terminus of METE contains conserved Zn<sup>2+</sup> binding residues required for catalytic activity (Gonzalez et al. 1992; Pejchal and Ludwig 2005). These residues, equivalent to His660, Cys662, Glu684, and Cys749 in *C. reinhardtii*, are conserved in all of the algal METE sequences we have obtained (black dots, Figure 3B), with the exception of that from *V. carteri*, in which the two cysteines are replaced by arginine and proline respectively. We re-sequenced this region of the *V. carteri* genome and verified that the assembled genomic sequence information was correct. In combination, these features indicate that *METE* in *V. carteri* is no longer functional.

In order to determine whether the *V. carteri* *METE* gene is expressed, we performed an RT-PCR analysis (See Supplementary Table 3 for primer sequences). As *C. reinhardtii* *METE* is repressed in the presence of vitamin B<sub>12</sub>, we depleted *V. carteri* cells of B<sub>12</sub>, taking samples from each subsequent sub-culture, up to the third sub-culture in which the cells failed to grow as a result of B<sub>12</sub> depletion. Two sets of primers were designed, targeted to predicted exons within conserved regions of the *METE* gene. *METH* was found to be expressed in both B<sub>12</sub>-replete and B<sub>12</sub>-depleted treatments (Figure 3C), as is the case in *C. reinhardtii* (Croft et al. 2005). However, despite being able to amplify a *METE* product from genomic DNA, we could not do the same with either primer set using cDNA as the template, indicating that the gene is not expressed.

*V. carteri* and *C. reinhardtii* are members of the chlorophycean order of Volvocales (Herron et al. 2009). Examination of the B<sub>12</sub> requirements within this order (Figure 4A) indicates that B<sub>12</sub> dependence may have arisen on multiple occasions even within this lineage.

We therefore investigated another B<sub>12</sub>-dependent member of the Volvocales, *Gonium pectorale*, for further evidence of recent *METE* gene loss. We confirmed *G. pectorale* is B<sub>12</sub> dependent using the B<sub>12</sub> growth assay (Figure 4B). Using degenerate primers targeted against a conserved region of the *C. reinhardtii* and *V. carteri* *METE* sequences, we were able to amplify a *METE* transcript from *G. pectorale* cDNA of the correct size (237 bp) (Figure 4C), although only in cells grown in the absence of vitamin B<sub>12</sub>; like *C. reinhardtii* the gene appears to be repressed by cobalamin. The larger band is likely to be genomic DNA contamination. The sequence of the 237 bp band exhibited high similarity to that of *C. reinhardtii* (Supplementary Figure 1). However, closer examination of these sequences revealed a significant deletion in *G. pectorale* *METE* that results in a frameshift and a premature stop codon, so that, like *V. carteri*, *G. pectorale* *METE* would not be able to produce a functional enzyme.

Taken together, these results indicate that in both *V. carteri* and *G. pectorale* *METE* has become a pseudogene, resulting in B<sub>12</sub> dependence in these algae. Given that *C. reinhardtii* diverged from *V. carteri* and *G. pectorale* approximately 250 MYA (Herron et al. 2009), this suggests a recent gene loss, and uniquely captures the evolution of B<sub>12</sub> auxotrophy in action. As several other members of the Volvocales are B<sub>12</sub>-independent, we suggest that the two pseudogenes in *V. carteri* and *G. pectorale* represent independent gene loss events.

The *V. carteri* *METE* gene represents a unitary pseudogene as there is no functional duplicate gene, and thus there is an accompanying phenotype. This is presumably also the case in *G. pectorale*, since it too is B<sub>12</sub> dependent. Another notable example of a unitary pseudogene resulting in vitamin auxotrophy is that of L-gulonolactone oxidase involved in the biosynthesis of vitamin C in animals (Nishikimi, Koshizaka and Yagi 1992; Nishikimi et al. 1994). As a consequence of losing this gene, guinea pigs and primates are no longer able to synthesize the vitamin themselves and instead must acquire it in their diet.

## Reassessing the B<sub>12</sub> Requirements of *E. huxleyi*

The haptophyte *E. huxleyi* does not contain *METE* within its genome, but it can survive in the absence of exogenous B<sub>12</sub> (Figure 2F), confirming earlier reports (Guillard 1963; Haines and Guillard 1974; Carlucci and Bowes 1970). Assuming that methionine synthase activity is essential for growth in *E. huxleyi*, either *METE* is present in the *E. huxleyi* genome but absent from the current genome assembly, or the growth assays did not reveal the true B<sub>12</sub> requirements of *E. huxleyi*.

The first suggestion is unlikely. *METE* is not present in either the *E. huxleyi* (CCMP 1516) genome assembly, the unassembled genomic reads, or in the *E. huxleyi* EST collections (Von Dassow et al. 2009). Additionally, using broad specificity degenerate primers (Huang et al. 2005), we were unable to amplify *METE* from genomic DNA of *E. huxleyi* (Supplementary Figure 2). Whilst this is not conclusive evidence for the absence of the gene, our combined data do not support the presence of *METE* in the *E. huxleyi* genome.

We therefore reassessed the B<sub>12</sub> requirements of *E. huxleyi*. We treated cultures of both calcifying and non-calcifying *E. huxleyi* (CCMP 1516) using extended methods to minimize bacterial contamination (see Materials and Methods). Both strains were able to grow in the absence of B<sub>12</sub> for up to five subcultures (Supplementary Figure 3). However, using a bioassay that measures B<sub>12</sub> we were able to detect up to 11.7 ng/mL of B<sub>12</sub> in lysed cell extracts of the *E. huxleyi* cells grown for four subcultures in medium without B<sub>12</sub>. In contrast, Croft et al. (2005) showed that cultures of several other B<sub>12</sub>-independent algae: *Haematococcus pluvialis*, *Cyanidium caldarium*, *Tetraselmis verrucosa* and *C. reinhardtii* have no detectable B<sub>12</sub> after just 2 subcultures. The source of B<sub>12</sub> in *E. huxleyi* is unlikely to be *de novo* synthesis, since only two out of the total of 20 enzymes required for aerobic B<sub>12</sub>



biosynthesis (Warren et al. 2002) were identified in the *E. huxleyi* genome. These genes exhibited weak similarity to *cobB* and *cobN*.

DAPI-staining of the *E. huxleyi* cultures did not reveal any associated bacteria. However, when the cultures were plated on highsalt LB media, growth of bacteria was observed. Analysis of the 16S rRNA gene established that the contaminant for the non-calcifying strain was a member of the *Oceanicaulis* genus, whereas for the calcifying strain it was of the *Aurantimonas* genus, within the order Rhizobiales, whose members are known to synthesize cobalamin. As the vitamin was not depleted in *E. huxleyi*, this result supports the hypothesis that algal species lacking *METE* require B<sub>12</sub> for growth. The apparent very close association of these bacteria with *E. huxleyi*, which could not be disrupted by antibiotic treatment, suggests an intimate relationship that warrants further attention

### **Phylogenetic Analysis of *METE* and *METH***

To examine the evolutionary relationships between the requirement for B<sub>12</sub> and the presence of methionine synthase isoforms, we generated a schematic phylogenetic tree (Figure 5) (Keeling and Palmer 2008), with representatives of the major eukaryote lineages, including land plants, mammals and fungi. It is clear that the B<sub>12</sub> requirement correlates with the absence of a functional *METE*.

A more detailed phylogenetic analysis of *METE* and *METH* sequences from a broad sample of organisms was then performed to enable a distinction to be made between whether multiple independent losses of *METE* resulted in B<sub>12</sub> auxotrophy, or conversely that multiple independent *gains* of *METE* may have led to the evolution of B<sub>12</sub> independence. Aside from the major endosymbiotic gene transfer (EGT) events, abundant horizontal gene transfers (HGT) have been proposed to encode novel metabolic capacities in some algae, for example

in the diatom *P. tricornutum* (Bowler et al. 2008). Our results indicate that the phylogeny of eukaryote *METH* is well resolved, with strong support for the monophyly of *METH* in the chlorophyte, chromalveolate and opisthokont (animal and fungal) lineages (Figure 6A). This argues against multiple HGT events in these lineages and supports the view that the absence of *METH* in the *C. merolae* and *Coccomyxa* sp. C-169 (Table 1) and in land plants reflects independent gene loss.

In contrast, phylogenetic analysis of *METE* sequences indicates a more complex evolutionary history. The majority of algal *METE* sequences forms a well-supported monophyletic clade with cyanobacterial homologues (Figure 6B), suggesting that it was probably acquired during the primary cyanobacterial endosymbiosis, as proposed before for *C. reinhardtii* *METE* (Moustafa and Bhattacharya 2008; Maruyama et al. 2009). This observation argues against multiple HGT events in these lineages, and thus the most likely explanation for the absence of *METE* in representatives of the Chlorophyta and Chromalveolata is as a result of multiple independent losses. This is in accordance with our discovery of *METE* pseudogenes in *V. carteri* and *G. pectorale*.

Surprisingly, *METE* sequences from land plants and the red alga *C. merolae* do not fall into this algal clade, although *METE* from another rhodophyte, *Galdieria sulphuraria*, clusters with the green algal sequences. A more recent HGT event must therefore be considered as a possible origin for *METE* in land plants and *C. merolae*. An alternative hypothesis is that *METE* may have originated from the eukaryote ancestor prior to the cyanobacterial primary endosymbiosis. Therefore, unlike the majority of algae that have retained the cyanobacterial copy of *METE*, the land plants may have retained the ancestral eukaryote gene. This implies that the ancestral members of the Archaeplastida (the eukaryotic supergroup including the Rhodophyta, Chlorophyta, and Glaucocystophyta, as well as the land plants, see Parfrey et al. (2006), for review) retained both forms of *METE* (ancestral

eukaryotic and cyanobacterial), at least until the streptophytes diverged from the chlorophytes. While we have identified no extant Archaeplastida species with both forms of the *METE* gene to support this hypothesis, co-existence followed by multiple gene loss has been proposed to explain the complex distribution of other genes in the green algal lineage, such as *EFL* and *EF-1α* (Noble, Rogers and Keeling 2007). It is worth noting that *Arabidopsis* has three different copies of *METE* (one plastidic and two cytosolic) (Ravanel et al. 2004), all of which belong to the land plant clade. Gene duplication and divergence may therefore have also played a role in the replacement of *METH* or the cyanobacterial *METE* in higher plants.

### **Effect of B<sub>12</sub> on Expression of *METE* in *P. tricornutum***

A comparison of *E. coli* methionine synthase isoforms found that *METH* has a much greater catalytic activity (approximately 100 fold) than *METE* (Gonzalez et al. 1992). This suggests that *METH* would be used preferentially in the presence of B<sub>12</sub>, and indeed in *E. coli*, the *metE* gene is repressed by cobalamin in the medium. In agreement with this, *METE* expression is also strongly repressed by B<sub>12</sub> in *C. reinhardtii* (Croft et al. 2005), and we made a similar observation for *G. pectorale* (Figure 4C). Thus, a prolonged continuous supply of B<sub>12</sub> could provide an opportunity for the accumulation of mutations in *METE*, which in time could lead to the loss of gene function. However, if this is to explain *METE* loss more broadly, it is important to determine whether B<sub>12</sub> represses *METE* expression in other algal species. Therefore, we assessed *METE* regulation in the chromalveolate diatom *P. tricornutum*, which like *C. reinhardtii* has both methionine synthase isoforms.

We used semi-quantitative RT-PCR to analyze the expression of *METE* and *METH* from *P. tricornutum* cells grown in the presence or absence of B<sub>12</sub>. *METH* transcript is detectable in both treatments, whereas *METE* is not expressed in the presence of vitamin B<sub>12</sub> (Figure 7A

and Supplementary Figure 4), and the repression occurs within 14 hours after the addition of the vitamin (Figure 7B). Experiments presented in this paper, alongside those conducted on *C. reinhardtii* (Croft et al. 2005), thus indicate that *METH* is used preferentially in organisms with both methionine synthase isoforms, providing a potential mechanism for *METE* gene loss.

## Discussion

The advent of whole genome sequencing has provided an unparalleled opportunity for comparisons between different organisms, in particular to test evolutionary questions. However, it is essential that sequence analysis is coupled with experimental approaches to provide independent verification of proposed events. We have combined growth assays and expression data with bioinformatics searches to investigate the reason for B<sub>12</sub> auxotrophy within the algal kingdom, and the possible mechanism by which it arose. Our results indicate strongly that the major determinant for the B<sub>12</sub> requirement of algae relates to the isoform(s) of methionine synthase they possess, and the available evidence supports multiple evolutionary losses of *METE* as a key factor in the distribution of B<sub>12</sub> auxotrophy amongst photosynthetic eukaryotes. In contrast, it is unlikely that B<sub>12</sub>-dependent MCM causes the auxotrophy, since it is found in B<sub>12</sub>-independent algal species such as *P. tricornutum*, indicating that the cellular functions of MCM activity may not be vital and thus could be dispensed with in the absence of B<sub>12</sub>. Similarly, the absence of MCM in species that have an absolute requirement for B<sub>12</sub> (eg. *V. catenella*, *M. pusilla*) supports the inference that this enzyme is not a determining factor.

Nonetheless, if we are to gain insight into the factors leading to the loss of *METE*, a fuller understanding of the functional roles of methionine synthase isoforms in different lineages is

required. Methionine synthase plays two major cellular roles, *de novo* synthesis of methionine and the regeneration of methionine following trans-methylation reactions. In land plants, at least, it is proposed that these functions occur in different sub-cellular compartments (the chloroplast and cytosol, respectively) (Ravanel et al. 2004). In certain organisms with both isoforms, *METE* and *METH* could conceivably perform different functions and/or exist in different sub-cellular compartments. For example, *METE* is also an abundant flagella protein in *C. reinhardtii* and has been implicated in protein methylation during flagella resorption (Schneider, Ulland and Sloboda 2008; Sloboda and Howard 2009). Thus, in *C. reinhardtii* there may be a selective pressure to maintain *METE* due to its specialised role in flagella function, although repression of *METE* by B<sub>12</sub> would suggest this is not an essential role, or that *METH* can function in its place. This explanation presumably also applies to flagellated algae such as *M. pusilla* and *V. carteri*, which lack *METE*. Differences in substrate availability may further influence the selective pressure to retain methionine synthase isoforms, since evidence from bacteria, plants and yeast indicate that *METE* is unable to utilize the monoglutamate form of methyltetrahydrofolate as a methyl donor (Burton and Sakami 1969; Whitfield and Weissbach 1970; Ravanel et al. 2004).

Exogenous sources of methionine may be available in the ocean, as one of the most abundant marine prokaryotes, SAR11 clade *Pelagibacter* lacks methionine synthase (Tripp et al. 2008). However, in eukaryotic algae an exogenous source of methionine would be insufficient to compensate for a lack of methionine synthase activity. In all cells, the most important metabolic function for this enzyme is its role as a component of the methylation cycle, and when methionine synthase activity is limiting a phenomenon known as folate trapping is observed (Scott, 1999). This is clearly demonstrated by the fact that methionine is insufficient to support the growth of *Lobomonas rostrata*, a B<sub>12</sub>-dependent green alga, in the absence of B<sub>12</sub>. It is necessary to supplement with folate in addition (Croft et al 2005). Given

these cellular considerations, we consider that the most important environmental factor promoting the selective loss of *METE* in certain algal lineages is a reliable and readily available source of B<sub>12</sub> in the environment, which is produced exclusively by prokaryotes (Warren et al. 2002). Croft et al. (2005) proposed that algae may obtain B<sub>12</sub> via a symbiotic relationship with bacteria, although a direct symbiosis involving B<sub>12</sub> has not yet been demonstrated in natural populations. However, there are many documented examples of close associations between algae and bacteria (Liu et al. 2008; Amin et al. 2009) and, regardless of the initial reason for their formation, these interactions may result in an abundant supply of B<sub>12</sub> for the algal partners. Of particular interest in this regard are the reports of bacterial symbionts associated with *Volvox*, including possible endosymbionts (Hamburger 1958), as well as the tightly associated bacteria associated with *E. huxleyi* found in the current study. Moreover, a recent study showed that vitamin auxotrophy for B<sub>12</sub> (and indeed thiamine and biotin) in harmful algal bloom species (HABs) such as dinoflagellates was much more prevalent than that for non-HAB species, (95% for B<sub>12</sub>, 74% for thiamine and 37% for biotin) (Zhong Tang et al. 2010). Many HABs are phagotrophic (e.g. Jeong, et al. 2005) and therefore probably obtain these micronutrients heterotrophically from their prey.

It is also worth considering the consequences of standard algal stock maintenance practices in the laboratory, in which cultures are grown with a recommended 1-10 nM exogenous vitamin B<sub>12</sub> (amongst other vitamins) regardless of whether they require it or not. Recent measurements indicate the lower range of vitamin B<sub>12</sub> concentrations in the Southern Ocean and North Atlantic Ocean approach potentially limiting values, with the lowest values being 0.4 pM and 0.2 pM respectively (Panzeca et al. 2009), whereas in our hands a minimum of 10 pM is required to support algal growth in batch culture. However, the requirement for a direct symbiosis between bacteria and algae for the delivery of B<sub>12</sub> has been disputed. Chemostat experiments indicate that the growth saturation constants for B<sub>12</sub> in

phytoplankton reside in the sub-picomolar range, suggesting that the concentration of B<sub>12</sub> released through bacterial cell lysis is sufficient to support algal growth, even in open ocean environments (Droop 2007). Nevertheless, there are several reports of B<sub>12</sub> stimulation of natural phytoplankton assemblages suggesting this cofactor may be limiting in certain areas (Sanudo-Wilhelmey et al. 2006; Bertrand et al. 2007). Although it remains to be determined whether B<sub>12</sub> limits phytoplankton productivity in a wider context, one would certainly expect selective pressure to retain *METE* to be higher in environments such as the open ocean where B<sub>12</sub> supply may be low or unreliable. In this regard, a metagenomic study of marine prokaryote communities indicated that methionine and cobalamin-dependent pathways varied significantly with environmental features (Gianoulis et al. 2009). An important factor limiting the ability of bacteria to synthesize cobalamin is the concentration of cobalt (Panzeca et al. 2008), so in turn this may have a significant influence on the distribution of B<sub>12</sub>-auxotrophy in oligotrophic environments. Given that algae alone are responsible for the fixation of 50% of the world's carbon dioxide (Field et al. 1998) and have a vital role in biogeochemical cycling in both marine and freshwater systems, it will be important to understand the extent to which B<sub>12</sub> availability influences both the physiology and the ecology of these organisms.

## Acknowledgements

We would like to thank Herve Moreau (Oceanological Observatory of Banyuls-sur-Mer, France) for donation of the strains *O. tauri* (OTH95), *O. lucimarinus* and *M. pusilla*. We are grateful to Thomas Mock and Jan Strauss (University of East Anglia, UK) for providing us with the culture of *F. cylindrus* and for their helpful advice. Thanks also goes to the Joint Genome Institute and all those contributing to each genome project referred to in this paper. Specifically to the coordinators for specific projects who made data available before publication: Thomas Mock (*F. cylindrus*), Betsy Read (California State University, USA) (*E.*

*huxleyi*), James L. Van Etten (Nebraska Centre for Virology, University of Nebraska-Lincoln, USA) (*Coccomyxa* sp. C-169 and *Chlorella variabilis* NC64A) and Christopher Gobler (School of Marine and Atmospheric Sciences, Stony Brook University, NY, USA) (*A. anophagefferens*). We would also like to thank Karen Weynberg (Plymouth Marine Laboratory, UK) for helping us to grow the *Ostreococcus* strains. Finally, our appreciation goes to Roger Sloboda (Dartmouth College, Hanover, USA) for his helpful insight and discussions. Funding was from the Biotechnology and Biological Sciences Research Council (BBSRC).

**Keywords:** (algae, pseudogene, vitamin auxotrophy, RT-PCR)

---

## Literature Cited

Altschul S, Gish W, Miller W, Myers E, Lipman D. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Amin SA, Green DH, Hart MC, Kupper FC, Sunda WG, Carrano CJ. 2009. Photolysis of iron-siderophore chelates promotes bacterial-algal mutualism. Proc. Natl. Acad. Sci. U. S. A. 106:17071–17076.

Armbrust EV, Berges JA, Bowler C, et al. (46 co-authors). 2004. The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. Science. 306:79–86.

Birney E, Clamp M, Durbin R. 2004. Genewise and genomewise. Genome Res. 14:988–995.



Bertrand EM, Saito MA, Rose JM, Riesselman CR, Lohan MC, Noble AE, Lee PA, DiTullio GR. 2007. Vitamin B<sub>12</sub> and iron co-limitation of phytoplankton growth in the Ross Sea. *Limnol. Oceanogr.* 52:1079–1093.

Blanc G, Duncan G, Agarkova I. et al. (15 co-authors) . 2010. The *Chlorella variabilis* NC64A genome reveals adaptation to photosymbiosis, coevolution with viruses, and cryptic sex. *Plant Cell.* 22: 2943–2955

Boalch GT. 1961. Studies on *Ectocarpus* in culture .2. growth and nutrition of a bacteria-free culture. *Journal of the Marine Biological Association of the United Kingdom.* 41:287.

Bowler C, Allen AE, Badger JH, et al. (77 co-authors). 2008. The *Phaeodactylum* genome reveals the evolutionary history of diatom genomes. *Nature.* 456:239– 244.

Burton EG, Sakami W. 1969. The formation of methionine from the mono-glutamate from of methyltetrahydrofolate by higher plants. *Biochem. Biophys. Res. Commun.* 36:228–34.

Carell EF, Seeger JW, Ju. 1980. Ribonucleotide reductase activity in vitamin B<sub>12</sub> deficient *Euglena gracilis*. *Biochem. J.* 188:573–576.

Carlucci AF, Bowes PM. 1970. Production of vitamin B<sub>12</sub> thiamine and biotin by phytoplankton. *Journal of Phycology.* 6:351–357.

Cock JM, Sterck L, Rouze P, et al. (77 co-authors). 2010. The *Ectocarpus* genome and the independent evolution of multicellularity in brown algae. *Nature*. 465:617–621.

Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. 2005. Algae acquire vitamin B<sub>12</sub> through a symbiotic relationship with bacteria. *Nature*. 438:90–93.

Croft MT, Warren MJ, Smith AG. 2006. Algae need their vitamins. *Eukaryotic Cell*. 5:1175–1183.

Delwiche CF. 1999. Tracing the thread of plastid diversity through the tapestry of life. *American Naturalist*. 154:S164–S177.

Dobson CM, Wai T, Leclerc D, Kadir H, Narang M, Lerner-Ellis JP, Hudson TJ, Rosenblatt DS, Gravel RA. 2002a. Identification of the gene responsible for the cblB complementation group of vitamin B<sub>12</sub>-dependent methylmalonic aciduria. *Hum. Mol. Genet*. 11:3361–3369.

Dobson CM, Wai T, Leclerc D, Wilson A, Wu XC, Dore C, Hudson T, Rosenblatt DS, Gravel RA. 2002b. Identification of the gene responsible for the cblA complementation group of vitamin B<sub>12</sub>-responsive methylmalonic acidemia based on analysis of prokaryotic gene arrangements. *Proc. Natl. Acad. Sci. U. S. A.* 99:15554–15559.

Droop MR. 2007. Vitamins, phytoplankton and bacteria: symbiosis or scavenging? *J. Plankton Res.* 29:107–113.

Dusi H. 1940. Culture of bacteriologiquement pure et nutrition autotrophe d'*Eudorina elegans*. Role du fer pour la formation des colonies. Ann. Inst. Pasteur. 64:340-3

Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 5:1–19.

Field CB, Behrenfeld MJ, Randerson JT, Falkowski P, 1998. Primary production of the biosphere: Integrating terrestrial and oceanic components. Science. 281:237–240.

Finn RD, Tate J, Mistry J, et al. (11 co-authors). 2008. The pfam protein families database. Nucleic Acids Res. 36:D281–D288.

Gianoulis TA, Raes J, Patel PV, et al. (12 co-authors). 2009. Quantifying environmental adaptation of metabolic pathways in metagenomics. Proc. Natl. Acad. Sci. U. S. A. 106:1374–1379.

Giovannoni SJ, Tripp HJ, Givan S, et al. (11 co-authors). 2005. Genome Streamlining in a Cosmopolitan Oceanic Bacterium. Science 309: 1242.

Gonzalez JC, Banerjee RV, Huang S, Sumner JS, Matthews RG. 1992. Comparison of cobalamin-independent and cobalamin-dependent methionine synthases from *Escherichia coli*- 2 solutions to the same chemical problem. Biochemistry. 31:6045–6056.

Gould SB, Waller RR, McFadden GI. 2008. Plastid evolution. *Annu. Rev. Plant Biol.* 59:491–517. Guillard R. 1963. Organic Sources of Nitrogen for Marine Centric Diatoms. 93 edition.

Guillard R. 1963. Organic sources of nitrogen for marine centric diatoms. *Symposium on Marine Microbiology* p. 93.

Guillard R. 1975. *Culture of phytoplankton for feeding marine invertebrates.*, Plenum Press, New York, USA, pp. 26–60.

Guillard R, Hargraves P. 1993. *Stichochrysis immobilis* is a diatom, not a chrysophyte. *Phycologia*. 32:234–236

Guillard R, Ryther J. 1962. Studies of marine planktonic diatoms. i. *Cyclotella nana* Hustedt and *Detonula confervacea* Cleve. *Can. J. Microbiol.* 8: 229-239.

Hackett JD, Su Yoon H, Li S, Reyes-Prieto A, Rummele E, Bhattacharya D. 2007. Phylogenomic analysis supports the monophyly of cryptophytes and haptophytes and the association of rhizaria with chromalveolates. *Mol. Biol. Evol.* 24(8):1702–1713. 2007

Haines and Guillard. 1974. Growth of vitamin B<sub>12</sub>- requiring marine diatoms in mixed laboratory cultures with vitamin B<sub>12</sub> producing marine bacteria. *J. Phycol.* 10: 245-252

Hall. 1999. Bioedit: a user-friendly biological sequence alignment editor and analysis program for windows. *Nucleic Acids Symposium Series No.* 41: 95-98.

Hamburger B. 1958. Bakteriensymbiose bei *Volvox aureus* ehrenberg. Archiv Fur Mikrobiologie. 29:291–310.

Hamilton FD. 1974. Ribonucleotide reductase from *Euglena gracilis*. a 5 deoxyadenosylcobalamin – dependent enzyme. J. Biol. Chem. 249:4428–4434.

Herron MD, Hackett JD, Aylward FO, Michod RE. 2009. Triassic origin and early radiation of multicellular volvocine algae. Proc. Natl. Acad. Sci. U. S. A. 106:3254–3258.

Huang L, Li DY, Wang SX, Zhang SM, Chen JH, Wu XF. 2005. Cloning and identification of methionine synthase gene from *Pichia pastoris*. Acta Biochimica Et Biophysica Sinica. 37:371–378.

Hutner SH. 1948. Essentiality of constituents of sea water for growth of a marine diatom. Transactions of the New York Academy of Sciences. 10:136–141.

Jeong JH, Yoo DY, Park YJ et al. 2005. Feeding by phototrophic red-tide dinoflagellates: five species newly revealed and six species previously known to be mixotrophic. Aquatic Microbial Ecology. 40: 133-150

Keeling PJ, Palmer DP. 2008. Horizontal gene transfer in eukaryotic evolution. Nat. Rev. Genet. 9:605–618.

Keller M, Guillard R. 1985. Factors significant to marine diatom culture. Elsevier, New York.

Keller M, Selvin R, Claus W, Guillard R. 1987. Media for the culture of oceanic ultra-phytoplankton. *Journal of Phycology*. 23:633-638

Kirk DL, Kirk MM. 1983. Protein synthetic patterns during the asexual life-cycle of *Volvox carteri*. *Developmental Biology*. 96:493–506.

Lewin. 1958. Vitamin-bezonoj de algoj. Liu JQ, Lewitus AJ, Brown P, Wilde SB. 2008. Growth promoting effects of a bacterium on raphidophytes and other phytoplankton. *Harmful Algae*. 7:1–10.

Liu JQ, Lewitus AJ, Brown P, Wilde SB. 2008. Growth promoting effects of a bacterium on raphidophytes and other phytoplankton. *Harmful Algae*. 7:1–10.

Mahoney J. 2005. Modification of chemically-defined medium asp12 for picoplankter *Aureococcus anophagefferens*, with limited comparison of physiological requirements of New York and New Jersey isolates. Northeast Fisheries Science Centre Reference Document 05-17.

Marsh. 1999. Coenzyme B<sub>12</sub> (cobalamin)-dependent enzymes. *Essays Biochem*. 34:139154.

Maruyama S, Matsuzaki M, Misawa K, Nozaki H. 2009. Cyanobacterial contribution to the genomes of the plastid-lacking protists. *BMC Evolutionary Biology*. 9.

Matsuzaki M, Misumi O, Shin-I T, et al. (41 co-authors). 2004. Genome sequence of the ultrasmall unicellular red alga *Cyanidioschyzon merolae*. *Nature*. 428:653–657.

McCutcheon JP, McDonald BR, Moran NA. 2009. Convergent evolution of metabolic roles in bacterial cosymbionts of insects. *Proc. Natl. Acad. Sci. U. S. A.* 106:15394–15399.

Merchant SS, Prochnik SE, Vallon O, et al. (101 coauthors). 2007. The *Chlamydomonas* genome reveals the evolution of key animal and plant functions. *Science*. 318:245–251.

Moreira D, Philippe H. 2001. Sure facts and open questions about the origin and evolution of photosynthetic plastids. *Research in Microbiology*. 152:771–780.

Moustafa A, Bhattacharya D. 2008. Phylosort: a user-friendly phylogenetic sorting tool and its application to estimating the cyanobacterial contribution to the nuclear genome of *Chlamydomonas*. *BMC Evolutionary Biology*. 8.

Newman SM, Boynton JE, Gillham NW, Randolph-Anderson BL, Johnson AM, Harris EH. 1990. Transformation of chloroplast ribosomal-RNA genes in *Chlamydomonas*- molecular and genetic-characterization of integration events. *Genetics*. 126:875–888.

Nishikimi M, Fukuyama R, Minoshima S, Shimizu N, Yagi K. 1994. Cloning and chromosomal mapping of the human nonfunctional gene for L-gulono-gamma-lactone oxidase, the enzyme for L-ascorbic-acid biosynthesis missing in man. *J. Biol. Chem.* 269:13685–13688.

Nishikimi M, Koshizaka T, Yagi K. 1992. Cloning of a gene related to the missing key enzyme for biosynthesis of ascorbic-acid in humans. *Biotechnology and Nutrition*. pp. 315–324.

Noble GP, Rogers MB, Keeling PJ. 2007. Complex distribution of EFL and EF-1 alpha proteins in the green algal lineage. *BMC Evolutionary Biology*. 7

Palenik B, Grimwood J, Aerts A, et al. (38 co-authors) 2007. The tiny eukaryote *Ostreococcus* provides genomic insights into the paradox of plankton speciation. *Proc. Natl. Acad. Sci. U. S. A.* 104:7705–7710.

Palmer EG and Starr RC. 1971. Nutrition of *Pandorina morum*. *J. Phycol.* 7:85-9

Panzeca C, Taylor GT, Sanudo-Wilhelmy SA, Leblanc K, Taylor GT, Hutchins DA, Sanudo-Wilhelmy SA. 2008. Potential cobalt limitation of vitamin B<sub>12</sub> synthesis in the North Atlantic Ocean. *Global Biogeochem. Cycles*. 22

Panzeca C, Beck AJ, Tovar-Sanchez A, Segovia-Zavala J, Taylor GT, Gobler CJ, Sanudo-Wilhelmy SA. 2009. Distributions of dissolved vitamin B<sub>12</sub> and Co in coastal and open-ocean environments. *Estuarine Coastal and Shelf Science*. 85:223–230.

Parfrey LW, Barbero E, Lasser E, Dunthorn M, Bhattacharya D, et al. 2006. Evaluating support for the current classification of eukaryotic diversity. *PLoS Genet.* 2(12):e220



Pejchal R, Ludwig ML. 2005. Cobalamin-independent methionine synthase (metE): A face-to-face double barrel that evolved by gene duplication. PLoS Biol. 3:254–265.

Pintner IJ and Provasoli L. 1963. Nutritional Characteristics of Some Chrysomonads. In Symp. Marine Microbiol. ed Oppenheimer CH p. 114-21. Thomas Co., Springfield III.

Pintner IJ and Provasoli L. 1963. The nutrition of *Volvox globator* and *V. tertius*. Proc. IX. Int. Bot. Congr. Montreal, abstr. 300-1

Pringsheim EG. 1970. Identification and cultivation of European *Volvox* sp. Antonie van Leeuwenhoek. 36: 33-43

Prochnik S, Umen J, Nedelcu A, et al. (28 co-authors). 2010. Genomic analysis of organismal complexity in the multicellular green alga *Volvox carteri*. Science. 329: 223–226

Provasoli L. 1958. Nutrition and ecology of protozoa and algae. A. Rev. Microbiol. 12:279–308. Ramirez-Flandes S, Ulloa O. 2008. Bosque: integrated phylogenetic analysis software. Bioinformatics. 24:2539–2541.

Raux E, Lanois A, Levillaye F, Warren M, Brody E, Rambach A, Thermes C. 1996. *Salmonella typhimurium* cobalamin (vitamin B<sub>12</sub>) biosynthetic genes: Functional studies in *S. typhimurium* and *Escherichia coli*. Journal of Bacteriology. 178:753–767.

Raux E, Lanois A, Warren M, Rambach A, Thermes C. 1998. Cobalamin (vitamin B<sub>12</sub>) biosynthesis : identification and characterization of a *Bacillus megaterium cobI* operon. Biochem. J. 335:159-66

Ravanel S, Block MA, Rippert P, Jabrin S, Curien G, Rebeille F, Douce R. 2004. Methionine metabolism in plants - chloroplasts are autonomous for de novo methionine synthesis and can import S-adenosylmethionine from the cytosol. J. Biol. Chem. 279:22548–22557.

Reyes-Prieto A, Weber APM, Bhattacharya D. 2007. The origin and establishment of the plastid in algae and plants. Annu. Rev. Genet. 41:147–168.

Sanudo-Wilhelmy SA, Gobler CJ, Okbami Michael M, Taylor GT. 2006. Regulation of phytoplankton dynamics by vitamin B<sub>12</sub>. Geophys. Res. Lett. 33.

Saito S. 1972. Growth of *Gonium multicoccum* in synthetic media. J. Phycol. 8:169-75

Schneider, MJ; Ulland, M; Sloboda, RD. 2008 A protein methylation pathway in *Chlamydomonas* flagella is active during flagellar resorption. Molecular Biology of the Cell. 19: 4319-4327

Scott J. Folate and vitamin B<sub>12</sub>. Proceedings of the Nutrition Society. 58:441-448 (1999)

Shihira RI & Krauss. 1965. *Chlorella*: Physiology and Taxonomy of 41 isolates, volume 97.

Sloboda RD, Howard L. 2009. Protein methylation in full length *Chlamydomonas* flagella. *Cell Motility and the Cytoskeleton*. 66:650–660.

Smith AG, Croft MT, Moulin M, Webb ME. 2007. Plants need their vitamins too. *Current Opinion in Plant Biology*. 10:266–275.

Stein JR. 1966. Growth and mating of *Gonium pectorale* (Volvocales) in defined media. *J. Phycol.* 2:27-28

Takenaka S, Takubo K, Watanabe F, Tanno T, Tsuyama S, Nanano Y, Tamura Y. 2003. Occurrence of coenzyme forms of vitamin B<sub>12</sub> in a cultured purple laver (*Porphyra yezoensis*). *Bioscience Biotechnology and Biochemistry*. 67:2480–2482.

Taylor AR, Brownlee C. 2003. A novel Cl<sup>-</sup> inward rectifying current in the plasma membrane of the calcifying marine phytoplankton *Coccolithus pelagicus*. *Plant Physiol*. 131:1391-1400.

Tripp HJ, Kitner JB, Schwalbach MS, Dacey JW, Wilhelm LJ, Giovannoni SJ. 2008. SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature*. 452:741-744

Warren MJ, Raux E, Schubert HL, Escalante-Semerena JC. 2002. The biosynthesis of adenosylcobalamin (vitamin B<sub>12</sub>). *Nat. Prod. Rep.* 19:390–412.

Waterhouse A, Procter J, Martin D, Clamp M, Barton GJ. 2009. Jalview version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics*. 25 (9):1189–1191.

Whelan S, Goldman N. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol. Biol. Evol.* 18:691–699.

Whitfield C, Weissbach H. 1970. Binding of folate substrate to 5-methyltetrahydropteroyltriglutamate-homocysteine transmethylase. *J. Biol. Chem.* 245:402.

Witman GB, Rosenbaum J, Berliner J, Carlson K. 1972. *Chlamydomonas* flagella. I. Isolation and electrophoretic analysis of microtubules, matrix, membranes, and mastigonemes. *J. Cell Biol.* 54:507

Worden AZ, Panaud, Piegu. (48 co-authors). 2009. Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes *Micromonas*. *Science*. 325:147–147.

Von Dassow P, Ogata H, Probert I, Wincker P, Da Silva C, Audic S, Claverie JM, de Vargas C. 2009. Transcriptome analysis of functional differentiation between haploid and diploid cells of *Emiliania huxleyi*, a globally significant photosynthetic calcifying cell. *Genome Biol.* 10(10):R114

Yamada K, Gravel RA, Toraya T, Matthews RG. 2006. Human methionine synthase reductase is a molecular chaperone for human methionine synthase. *Proc. Natl. Acad. Sci. U. S. A.* 103:9476–9481.

Zhang Y, Rodionov DA, Gelfand MS, Gladyshev VN. 2009. Comparative genomic analyses of nickel, cobalt and vitamin B<sub>12</sub> utilization. *BMC Genomics*. 10.

Zhong Tang Y, Koch F and Gobler C. 2010. Most harmful algal bloom species are vitamin B<sub>1</sub> and B<sub>12</sub> auxotrophs. Proc. Natl. Acad. Sci. U. S. A.

---

## Figure Legends

FIG.1. Schematic diagram illustrating the intracellular metabolism of B<sub>12</sub> in eukaryotes. The figure shows the B<sub>12</sub>-dependent enzymes methionine synthase (METH) and methylmalonyl-CoA mutase (MCM), and their accessory proteins (in dashed boxes). METE catalyzes the same reaction as METH, but does not require B<sub>12</sub> as a cofactor. Both use methyltetrahydrofolate as the methyl donor, but for METH the methyl group is transferred to the substrate via cobalamin. AdoCbl, adenosylcobalamin; MeCbl, methylcobalamin; Cbl I, Cbl II refers the oxidation state of the cobalt ion contained within vitamin B<sub>12</sub>.

FIG. 2. Assessment of B<sub>12</sub> requirements of selected algae. Species were grown in the appropriate liquid medium (Supplementary Table 1) with and without B<sub>12</sub> in batch culture over several sub-cultures, or until the cells had died. A) *V. carteri*, B) *Ostreococcus* sp. RCC809 C) *O. tauri*, D) *M. pusilla*, E) *F. cylindrus*, and F) *E. huxleyi*. Black bars: + B<sub>12</sub>, White: no B<sub>12</sub>. The following measurements were used to quantify growth: spheroid number (A), optical density (B, C and D), cell count (E and F). Three replicate cultures were used for each treatment. Error bars denote standard errors.

FIG. 3. *V. carteri* contains a unitary *METE* pseudogene. A) Schematic diagram of the gene model for *C. reinhardtii* (*C. r*) *METE* (upper) compared to that of *V. carteri* (*V. c*; lower) indicating that *V. carteri* gene is truncated with multiple frameshifts and an in-frame stop

codon (boxes: exons, arrows + dashed white line: frameshift, \*: in-frame stop codon). B) Multiple sequence alignment (ClustalW) of METE protein sequences from 8 algae for the region spanning the functionally important conserved residues required for binding  $\text{Zn}^{2+}$  (indicated by a black dot). The two cysteines are absolutely conserved except in *V. carteri*. C) PCR amplification of *METE*, *METH* and actin (*ACTA*) in *V. carteri*, using two different sub-cultures (Sub 1 & 2) grown with and without  $\text{B}_{12}$ . RNA was reversed transcribed and then used with two sets of primers (F1/R1 and F2/R2) targeted to different conserved regions of *METE*, and avoiding intron/exon boundaries in *C. reinhardtii*. However, no *METE* transcripts were detected, although a product of the expected size was amplified from genomic DNA (right-hand boxes), and the *METH* and *ACTA* transcripts were readily detected.

FIG 4.  $\text{B}_{12}$  requirements in the Volvocales A) Schematic tree displaying phylogenetic relationships between Volvocalean species (based on Herron et al. 2009) for which  $\text{B}_{12}$  requirements are already known: *C. reinhardtii* (Lewin 1958), *Gonium pectorale* (Stein 1966), *G. multicoccum* (Saito 1972), *Volvox globata*, *Eudorina elegans*, *V. tertius* (Pintner and Provasoli 1959), *Pandorina morum* (Palmer and Starr 1971), and *V. aureus* (Pringsheim 1966), or for *V. carteri* determined in this study. B) Confirmation of  $\text{B}_{12}$  dependence of *G. pectorale* with a  $\text{B}_{12}$  growth assay. C) RT-PCR analysis of *METE* transcripts in *G. pectorale* grown with and without  $\text{B}_{12}$ . A band of the correct size (237 bp) is obtained in the latter case, and sequencing confirmed it to be from *METE*.

FIG. 5. Schematic tree showing phylogenetic relationships between organisms and their  $\text{B}_{12}$  requirements. The phylogeny is based on Keeling and Palmer (2008). In addition to the algae species used in the current study for which complete genome sequences are known, two prokaryotes are shown (the cyanobacterium *Synechocystis* sp. strain PCC6803 and *E. coli*), as

well as members of the land plants, animals and fungi. The organisms fall into the following distinct phylogenetic groups indicated by shading, from top to bottom: Prokaryota, Chromalveolata, Rhodophyta, Chlorophyta, Streptophyta and the Opisthokonta (in dashed box). Ticks indicate that a species can grow without an exogenous source of B<sub>12</sub>, whereas a star (\*) is shown for *Synechocystis* sp. strain PCC6803. This can synthesize cobalamin (Raux et al. 1998) and therefore does not require an external source of the vitamin even though it lacks *METE*. Black squares indicate the presence of a gene, and partial grey boxes indicate a partial hit. In eukaryotes, a strong correlation between the requirement for B<sub>12</sub> and the presence of *METH* in the absence of a complete *METE* can be observed.

FIG. 6. Phylogenetic analysis of methionine synthase isoforms. A) Maximum likelihood consensus tree constructed for METH using PhyML. 100 bootstraps were run, values above 50% are shown. B) Maximum likelihood consensus tree for METE, constructed as described above.

FIG. 7. *METE* expression is repressed by B<sub>12</sub> in the diatom *P. tricornutum*. A) RT-PCR analysis of transcripts for *METE* and *METH* following growth + B<sub>12</sub> or -B<sub>12</sub>. B) 24 hour time-course of *METE* and *METH* expression following the addition of vitamin B<sub>12</sub> at time 0 h. H4, constitutively expressed histone H4 gene. The following cycle number for each primer set was used throughout: *METE* (35 cycles), *METH* (33 cycles) and *HISTONE H4* (33 cycles), see Supplementary Figure 4.

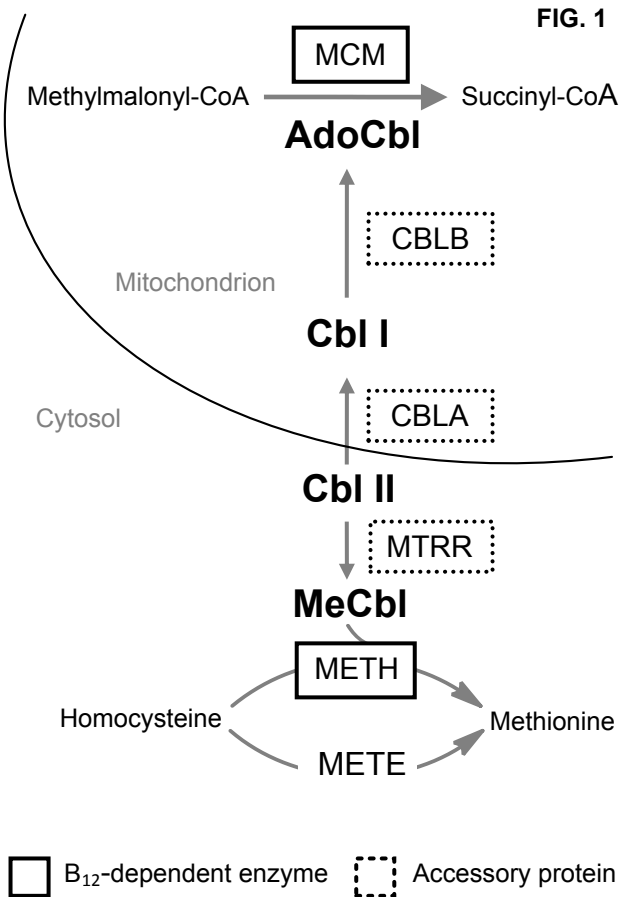
---

Table 1. Analysis of 15 sequenced algal genomes for genes encoding the three B<sub>12</sub>-dependent proteins (and their accessory proteins) known to occur in eukaryotes: <sup>1</sup> A tick indicates that a complete gene was present; an X indicates no hit obtained. <sup>2</sup> an incomplete sequence for *METE* in *V. carteri* was verified by re-sequencing, and was subsequently shown to be a pseudogene (see text for details). <sup>3</sup>The METH results for *F. cylindrus* and *A. anophagefferens* were not complete, but were in regions with gaps in the genome assembly, and so are indicated as “partial”. <sup>4</sup>The B<sub>12</sub> requirements for each of the algal species were either established in this study [1], or taken from the literature: [2] Merchant et al. 2007; [3] Levin 1958; [4] Provasoli 1958; [5] Prochnik et al. 2010; [6] Worden, Panaud and Piegu 2009; [7] Palenik et al. 2007; [8] Shihira 1965; [9] Guillard 1963; [10] Pintner and Provasoli 1963; [11] Bowler et al. 2008; [12] Hutner 1948; [13] Mock et al., unpublished data; [14] Armbrust et al. 2004; [15] Croft et al. 2005; [16] Cock et al. 2010; [17] Boalch 1961; [18] Mahoney 2005; [19] Matsuzaki et al. 2004



		Genes Encoding B <sub>12</sub> - dependent Enzymes				Genes Encoding Accessory Proteins			B <sub>12</sub> Dependent? <sup>4</sup>
		<i>RNR</i>							
Group	Species	<i>METE</i>	<i>METH</i>	<i>II</i>	<i>MCM</i>	<i>MTRR</i>	<i>CBLA</i>	<i>CBLB</i>	
Chlorophyta	<i>Chlamydomonas reinhardtii</i> [2]	✓ <sup>1</sup>	✓	✗	✗	✓	✗	✓	✗[3,4]
	<i>Volvox carteri</i> f. <i>nagariensis</i> EVE [5]	Pseudo <sub>2</sub>	✓	✗	✗	✓	✗	✓	✓[1]
	<i>Micromonas pusilla</i> (CCMP 1545) [6]	✗	✓	✗	✗	✓	✗	✓	✓[1]
	<i>Ostreococcus tauri</i> (OTH95)	✗	✓	✗	✗	✓	✗	✗	✓[1]
	<i>Ostreococcus lucimarinus</i> [7]	✗	✓	✗	✗	✓	✗	✓	?
	<i>Ostreococcus</i> sp. RCC809 [7]	✗	✓	✗	✗	✓	✗	✓	✓[1]
	<i>Coccomyxa</i> sp. C-169	✓	✗	✗	✗	✓	✗	✗	✗ [8]
	<i>Chlorella variabilis</i> NC64A	✓	✓	✗	✓	✓	✓	✓	?
Chromalveolata	<i>Emiliana huxleyi</i> (CCMP1516)	✗	✓	✗	✓	✓	✓	✓	✗[1, 9, 10]
	<i>Phaeodactylum tricornutum</i> (CCAP1055/1) [11]	✓	✓	✗	✓	✓	✓	✓	✗ [12]
	<i>Fragilariopsis cylindrus</i> ( CCMP 1102) [13]	✓	Partial <sub>3</sub>	✗	✓	✓	✓	✓	✗ [1]
	<i>Thalassiosira pseudonana</i> [14]	✗	✓	✗	✓	✓	✓	✓	✓[15]
	<i>Ectocarpus siliculosus</i> [16]	✓	✓	✓	✓	✓	✓	✓	✗ [17]
	<i>Aureococcus anophagefferens</i>	✗	Partial	✗	✓	✓	✗	✓	✓[18]
Rhodophyta	<i>Cyanidioschyzon merolae</i> [19]	✓	✗	✗	✗	✓	✗	✗	✗[15]

FIG. 1



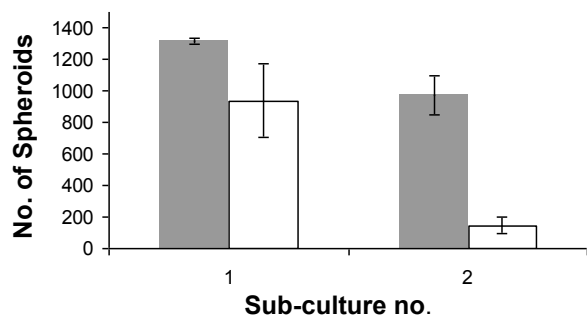
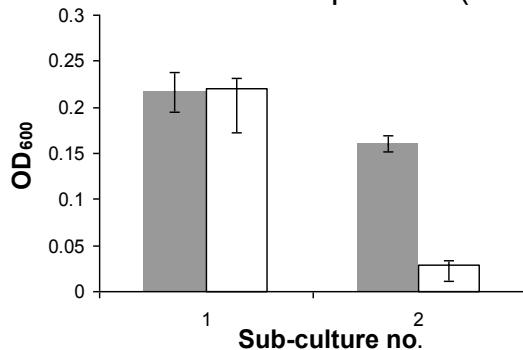
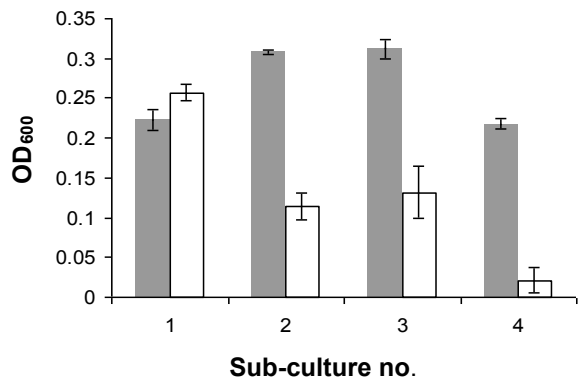
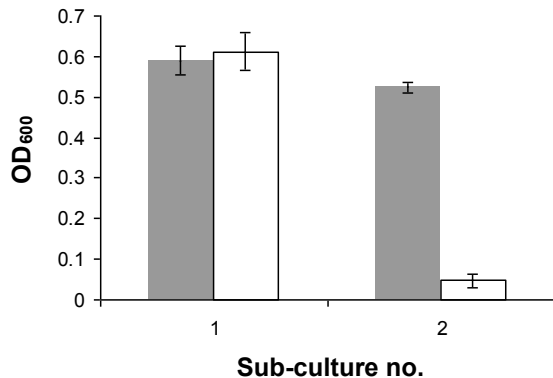
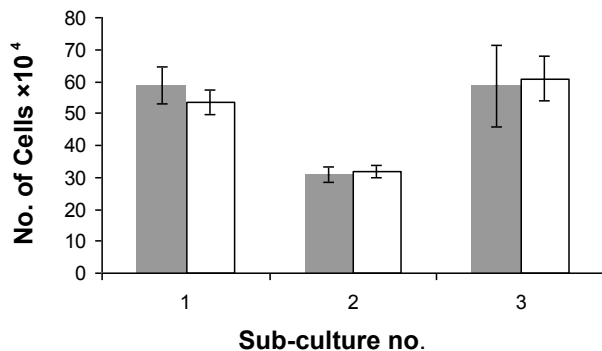
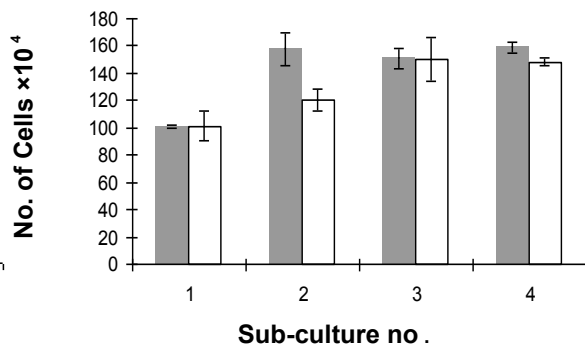
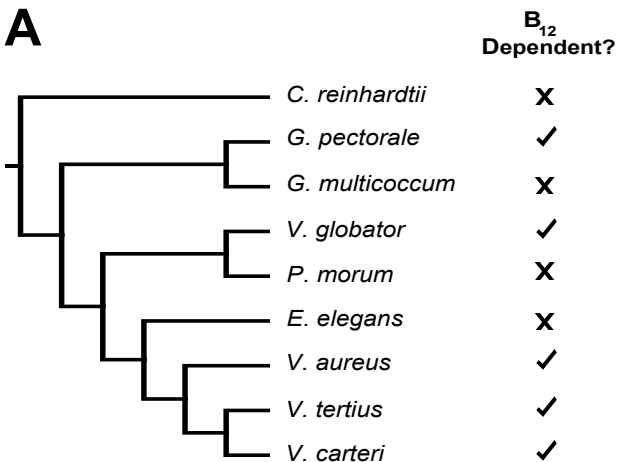
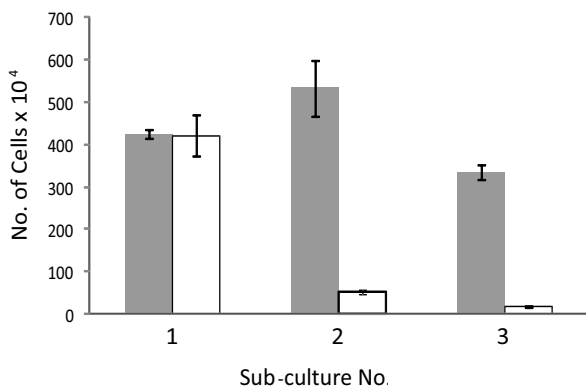
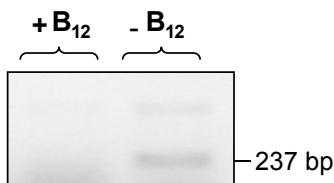
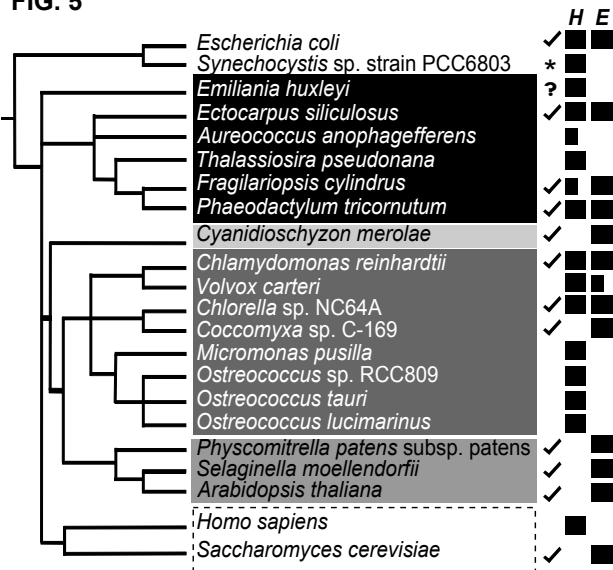
**A** *Volvox carteri* f. *nagariensis* EVE**B** *Ostreococcus* Deep Strain (RCC 809)**C** *Ostreococcus tauri* (OTH 95)**D** *Micromonas pusilla* (CCMP 1545)**E** *Fragilariopsis cylindrus* (CCMP 1102)**F** *Emiliana huxleyi* (CCMP 1516)



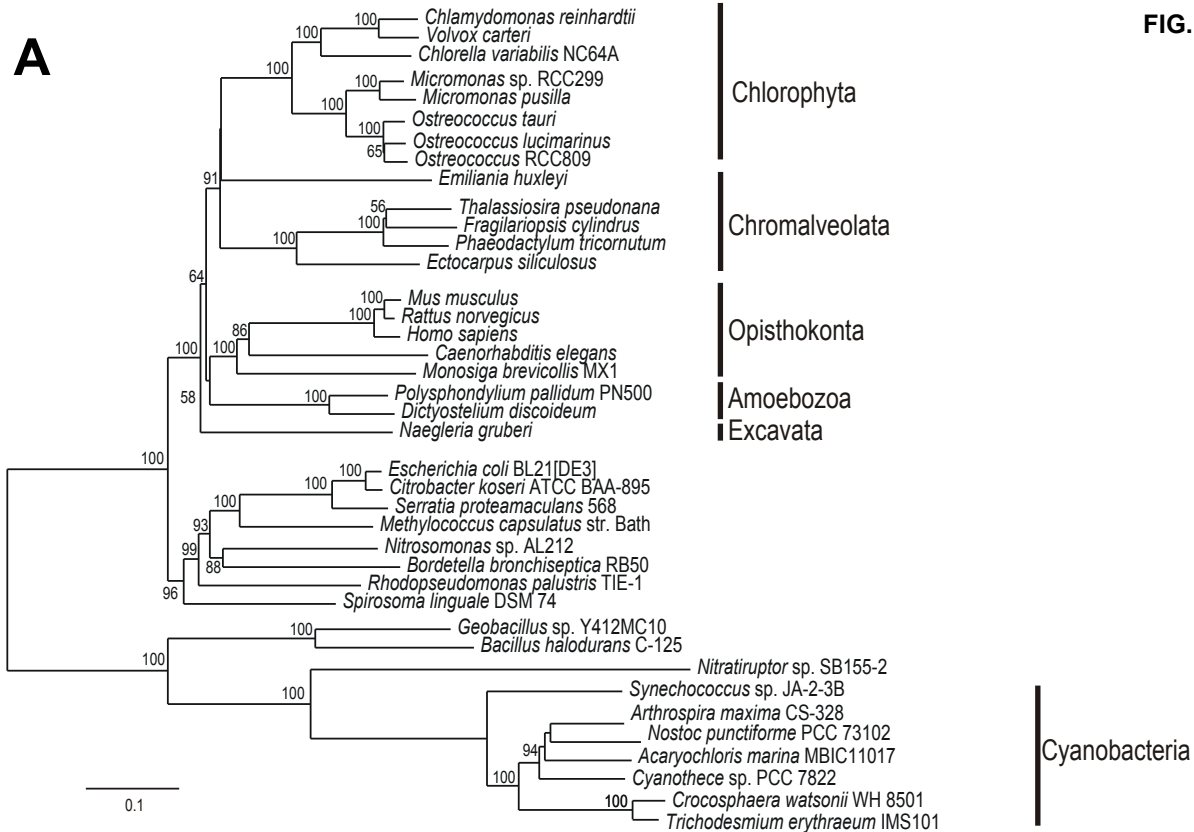
FIG. 4

**A****B****C**

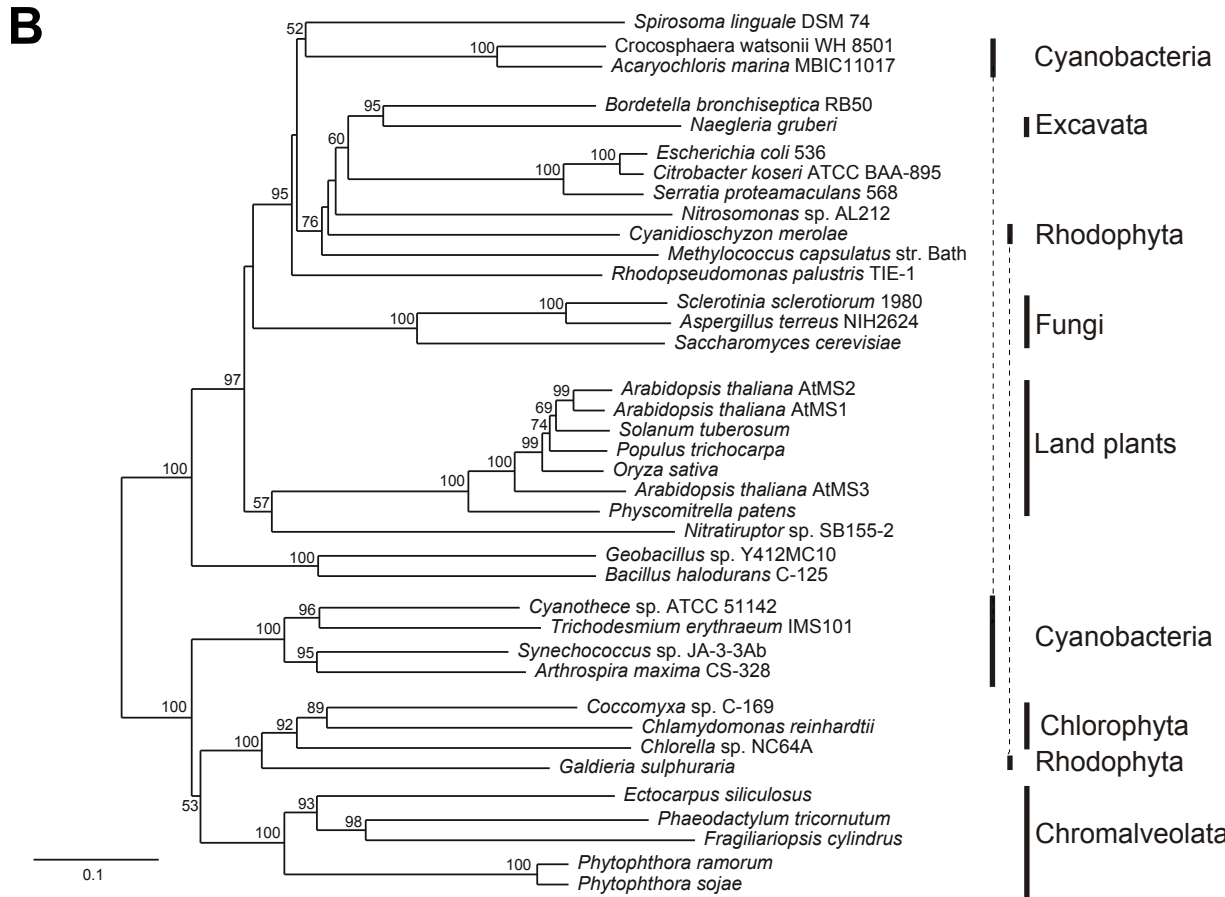
**FIG. 5**

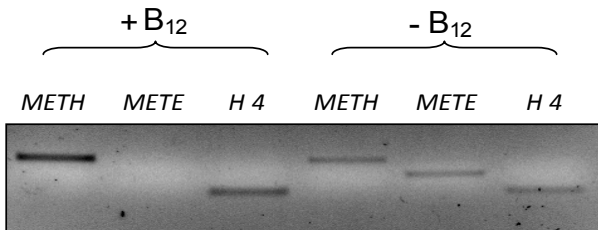


A



B



**A****B**