Insights into the Evolution of Vitamin B$_{12}$ Auxotrophy from Sequenced Algal Genomes

Katherine E. Helliwell,¹ Glen L. Wheeler,²,³ Kyriacos C. Leptos,⁴ Raymond E. Goldstein,⁴ and Alison G. Smith*,¹

¹Department of Plant Sciences, University of Cambridge, Cambridge, United Kingdom
²Plymouth Marine Laboratory, Plymouth, United Kingdom
³Marine Biological Association of United Kingdom, The Laboratory, Plymouth, United Kingdom
⁴Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, United Kingdom

*Corresponding author: E-mail: as25@cam.ac.uk.
Associate editor: Charles Delwiche

Abstract

Vitamin B$_{12}$ (cobalamin) is a dietary requirement for humans because 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$_{12}$-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$_{12}$ metabolism in the sequenced genomes of 15 different algal species, with representatives of the red, green, and brown algae, diatoms, and coccolithophores, 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$_{12}$ auxotrophy. The presence of a METE unitary pseudogene in the B$_{12}$-dependent green algae Volvox carteri and Gonium pectorale, relatives of the B$_{12}$-independent Chlamydomonas reinhardtii, suggest that B$_{12}$ 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$_{12}$ auxotrophy in these organisms. Because 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.

Key words: algae, cobalamin, pseudogene, methionine synthase, vitamin B$_{12}$, vitamin auxotrophy, RT-PCR.

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, which is the inability to synthesize an organic nutrient essential for growth. Over half of all microalgal species require an exogenous supply of vitamin B$_{12}$ (cobalamin), whilst just over 20% require vitamin B$_{1}$ (thiamine) and a smaller proportion (5%) require biotin (vitamin B$_{7}$) (Croft et al. 2006). However, vitamin auxotrophy is not phylogenetically related but instead distributed throughout the algal clades, implying that it has arisen multiple times through 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 et al. 2006), but for cobalamin, this does not appear to be the case. Those algae that are not dependent on B$_{12}$ do not contain it unless it is supplied in the medium (Croft et al. 2005). In fact, B$_{12}$ 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$_{12}$—some, such as Escherichia coli, will use it if it is available, whereas other species have no enzymes that use B$_{12}$. A notable example is Pelagibacter ubique, a member of the highly abundant SAR11 clade of marine bacteria (Giovannoni et al. 2005).

Cobalamin is a complex Co$^{2+}$-containing modified tetrapyrrole that acts as a cofactor for enzymes involved in C1 metabolism and certain radical reactions. In prokaryotes, there are over 20 enzymes that have a cobalamin cofactor (Marsh 1999), three of which have been found in eukaryotes. In humans, there are just two B$_{12}$-dependent enzymes, methylmalonyl-CoA mutase (MCM), involved in odd-chain fatty acid metabolism in the mitochondria, and methionine synthase (METH), which catalyzes the C1 transfer from methyltetrahydrofolate to homocysteine to make methionine. An alternative B$_{12}$-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$_{12}$ is available in the environment, otherwise it uses METE. Similarly, there are both B$_{12}$-independent and B$_{12}$-dependent forms of ribonucleotide reductase (RNR) (type I & III RNR are B$_{12}$ independent, type II is B$_{12}$ dependent), involved in deoxyribose biosynthesis (Hamilton 1974; Carell and Seeger 1980). Many prokaryotes encode both types of enzyme and again appear to switch between them depending on the availability of cobalamin in the environment. All eukaryotes have the type I isoform, whilst *Euglena gracilis*, an excavae protist, is reported to have B$_{12}$-dependent RNR II as well (Hamilton 1974).

An initial search for the presence of genes encoding B$_{12}$-dependent enzymes in four sequenced algal genomes suggested that B$_{12}$ auxotrophy may be related to the form of methionine synthase present (Crotf et al. 2005). The green alga *Chlamydomonas reinhardtii* and the diatom *Phaeodactylum tricornutum* have both METE and METH and are B$_{12}$ independent, whereas another diatom *Thalassiosira pseudonana*, which contains METH only, is B$_{12}$ dependent. The B$_{12}$-independent red alga *Cyanidioschyzon merolae* has METE only. However, algae are an extremely diverse group of organisms, so a much broader study is necessary to allow definitive conclusions 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 et al. 2007; Gould et al. 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 et al. 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$_{12}$-dependent enzymes and associated proteins in 15 algal species with completed genome sequences that span the algal kingdom. Our results provide strong evidence that the major determinant for the B$_{12}$ 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$_{12}$ 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, Supplementary Material online, 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 (AAFA6876.1), MCM (AAAS9569.1), and CBLB (AAAI1831.1), *E. coli* BBLA (P27254.2), *Eu. gracilis* RNR II (Q2PDF6.1) (protein IDs of all putative hits can be found in supplementary table 2, Supplementary Material online). 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), *Chlorella variabilis* NC64A (v1) (Blanc et al. 2010), *Microspora* (CCMP 1545) (v2) (Worden et al. 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), *P. tricornutum* (CCAP 1055/1) (v2) (Armbust et al. 2004; Bowler et al. 2008), *Emiliania huxleyi* (CCMP 1516) (v1) (http://genome.jgi-psf.org/Emi1/Emi1.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 *Cy. 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 (Binney et al. 2004) to identify missing introns in poor gene models, and 3) Pfam (24.0) analysis (Finne et al. 2008) to identify conserved functional domains in each protein (METE: PF01717, PF08267; METH:PF02965, PF02965, PF02607, PF00809, and PF02574; RNR II: PF09747, PF02867; MCM: PF01642, PF02310; MTRR: PF00258, PF0667, and PF00175; CBLB: PF3308, PF00875; and CBLB: PF01923).

#### Treatment of Algae Prior to B$_{12}$ Growth Assays

Bacteria are a potential source of B$_{12}$, 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 six-well microplates using a sterile pipette. The final washed spheroids were tested for the presence of bacterial contamination on lysogeny broth (LB) plates. *Gonium pectorale* was made axenic by serial streak
outs on standard Volvox medium 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), *Em. huxleyi*, and *M. pusilla* were treated with antibiotics. Cultures were grown for 1 week with penicillin (1 mg/ml), kanamycin (25 mg/ml), and neomycin (20 mg/ml). An additional subculture without antibiotics for approximately 6 days was then grown, before inoculating the experimental vitamin B$_{12}$ growth assays. For *Em. 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 min 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 nontreated control cultures but were not seen in the antibiotic treated cultures used in this study.

**Assessment of Vitamin B$_{12}$ Requirements of Algal Species**

The vitamin B$_{12}$ requirements of the algal species (treated as above) were assessed by growing each alga in the appropriate medium, plus or minus B$_{12}$. The concentration of B$_{12}$ used in the plus condition was as described in the media appropriate to each alga (see supplementary table 1, Supplementary Material online). 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 subcultures were carried out, or until the algae had died in the experimental condition (minus B$_{12}$) and not in the control (plus B$_{12}$) as described in Croft et al. 2005. Three biological replicates were carried out for each condition. If the cultures were still viable after B$_{12}$ starvation, they were visually checked with DAPI and/or plated onto LB plates (for freshwater species) or high salt LB (for marine species).

**B$_{12}$ Bioassay**

B$_{12}$ bioassays on cell lysates of *Em. huxleyi* were carried out as described by Raux et al. (1996), using *S. enterica* (AR3621). For preparation of the algal cell lysates, cells were washed 3 times and resuspended in 0.9% NaCl. The cells were then subjected to three freeze–thaw cycles and placed at 100 °C for 15 min 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 h at 37 °C to remove contaminating DNA. Reverse transcription (RT) was performed using Superscript II transcriptase (Invitrogen) with oligo dT primers according to the manufacturer’s instructions. Polymerase chain reaction (PCR) was used to amplify sequences from genomic DNA and cDNA using Taq DNA polymerase (Bioline) (see supplementary table 3, Supplementary Material online, 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 fig. 1, Supplementary Material online). 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 nonalgal eukaryotes. In order to ensure that the data set included appropriate prokaryote sequences, BlastP searches of the GenBank nonredundant 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 Whelan and Goldman substitution matrix (Whelan and Goldman 2001). One hundred bootstrap replicates were run for each analysis.

**Results**

**Survey of B$_{12}$-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$_{12}$. We also searched for accessory proteins involved in supplying the cofactor to these B$_{12}$-dependent enzymes. MCM utilizes the adenosylated form of cobalamin (AdoCbl) and requires two accessory proteins, CBLA for vitamin B$_{12}$ transport into the mitochondria and CBLB for AdoCbl synthesis (Dobson, Wai, Leclerc, Kadir, et al. 2002; Dobson, Wai, Leclerc, Wilson, et al. 2002) (see fig. 1). Conversely, METH uses methylcobalamin (MeCbl) and requires the molecular chaperone methionine synthase reductase (MTRR) to regenerate the methylated cofactor (Dobson, Wai, Leclerc, Kadir, et al. 2002; 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, Supplementary Material online, for accession number details).
However, the distribution of the accessory proteins was simple plastids, the green alga ing the Chromalveolata, but only in one of the species with missing sequence, implying that this was a pseudogene. When the predicted amino acid sequence was obtained using GeneWise 2.0, there were numerous frameshifts and regions when the genome assembly was complete in this region, had sequence similarity to A. anophagefferens, but for METH, the methyl group is transferred to the substrate via 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 B12.

We found that all species have the B12-dependent methionine synthase, METH, except Coccomyxa sp. C-169 and Cy. 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 B12-dependent isofom of methionine synthase. In contrast, the gene encoding the vitamin B12-independent isoform, METE, was present in only seven of the genomes and absent altogether from another seven. 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 Ch. 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 B12-dependent enzyme. The third B12-dependent enzyme RNR II, previously reported in Eu. gracilis (Hamilton 1974), was present only in Ec. siliculosus. This alga also encodes both subunits of B12-independent RNR 1 (Cock et al. 2010).

Assessment of Algal Requirements for Vitamin B12
To interpret this genomic data set, the B12 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 B12 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 B12. Artificial seawater was used where possible (see supplementary table 1, Supplementary Material online) to exclude a potential source of B12 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 subcultures, it became clear that V. carteri, Ostreococcus sp. RCC809, O. tauri, and M. pusilla were B12 dependent (fig. 2A–D). The dramatic decrease in growth within two subcultures suggests our precautions to avoid contaminating sources of B12 were sufficient and that false positives were unlikely. In contrast, for the Arctic diatom F. cylindrus and the haptophyte Em. huxleyi, no difference was seen in growth with or without B12.

Methionine Synthase Isoform Determines B12 Dependence
An analysis of the B12-dependent enzymes in the six algal species that require B12 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 B12-dependent algal species, indicating that METH would be the sole source of methionine synthase activity. Conversely, with the exception of Em. huxleyi, METE is universally present in the B12-independent algae, indicating that METE enables the organism to live in the absence of the cofactor. Although Ec. siliculosus contains type II RNR, it also encodes B12-independent RNR I, explaining the observation that the alga itself does not require B12 for growth (Boalch 1961). From these observations, we conclude that the isoform of methionine synthase is the key factor determining the nutritional requirement for B12 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 Em. huxleyi is able to grow in the absence of both METE and exogenous B12.
Table 1. Analysis of 15 Sequenced Algal Genomes for Genes Encoding the Three B12-Dependent Proteins (and their accessory proteins) Known to Occur in Eukaryotes.

<table>
<thead>
<tr>
<th>Group</th>
<th>Species</th>
<th>Genes Encoding B12-Dependent Enzymes</th>
<th>Genes Encoding Accessory Proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>METE</td>
<td>METH</td>
</tr>
<tr>
<td>Chlorophyta</td>
<td><em>Chlamydomonas reinhardtii</em> [2]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Volvoxoceratia f. nagariensis</em> EVE [5]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Micromonas pusilla</em> (CCMP 1545) [6]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Ostreococcus oithnia</em> (OTH95)</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Ostreococcus lucimarinus</em> [7]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Ostreococcus sp. RCC809</em> [7]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Coccomyxa sp. C-169</em></td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Chlorera NC64A</em></td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td>Chromalveolata</td>
<td><em>Emiliania huxleyi</em> (CCMP1516)</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Phaeodactylum tricornutum</em> (CCAP1055/1) [11]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Fragilariopsis cylindrus</em> (CCMP 1102) [13]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Thalassiosira pseudonana</em> [14]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Ectocarpus siliculosus</em> [16]</td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td></td>
<td><em>Aureococcus anophagefferens</em></td>
<td>☑</td>
<td>☑</td>
</tr>
<tr>
<td>Rhodophyta</td>
<td><em>Cyanidioschyzon merolae</em> [19]</td>
<td>☑</td>
<td>☑</td>
</tr>
</tbody>
</table>

A tick indicates that a complete gene was present; an × indicates no hit obtained.

An incomplete sequence for METE in *V. carteri* was verified by resequencing and was subsequently shown to be a pseudogene (see text for details).

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.”


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, fig. 3A) and multiple deletions leading to seven frameshifts (dashed white lines, fig. 3A). The C-terminus of METE contains conserved Zn\(^{2+}\)-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, fig. 3B), with the exception of that from *V. carteri*, in which the two cysteines are replaced by arginine and proline, respectively. We resequenced 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, Supplementary Material online, for primer sequences). As *C. reinhardtii* METE is repressed in the presence of vitamin B12, we depleted *V. carteri* cells of B12, taking samples from each subsequent subculture, up to the third subculture in which the cells failed to grow as a result of B12 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 B12-replete and B12-depleted treatments (fig. 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 B12 requirements within this order (fig. 4A) indicates that B12 dependence may have arisen on multiple occasions even within this lineage. We therefore investigated another B12-dependent member of the Volvocales, *G. pectorale*, for further evidence of recent METE gene loss. We confirmed *G. pectorale* is B12 dependent using the B12 growth assay (fig. 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) (fig. 4C), although only in cells grown in the absence of vitamin B12; 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 fig. 2, Supplementary Material online). 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 B12 dependence in these algae. Given that *C. reinhardtii* diverged from *V. carteri* and *G. pectorale* approximately 250 Ma (Herron et al. 2009), this suggests a recent gene loss and uniquely captures the evolution of B12 auxotrophy in action. As several other members of the Volvocales are B12 independent, we suggest that the two pseudogenes in *V. carteri* and *G. pectorale* represent independent gene loss events.
The \textit{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 \textit{G. pectorale} because it too is B12 dependent. Another notable example of a unitary pseudogene resulting in vitamin auxotrophy is that of \textit{L-gulono-\gamma-lactone oxidase} involved in the biosynthesis of vitamin C in animals (Nishikimi et al. 1992, 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.

The \textit{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 \textit{G. pectorale} because it too is B12 dependent. Another notable example of a unitary pseudogene resulting in vitamin auxotrophy is that of \textit{L-gulono-\gamma-lactone oxidase} involved in the biosynthesis of vitamin C in animals (Nishikimi et al. 1992, 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 B12 Requirements of \textit{Em. huxleyi}

The haptophyte \textit{Em. huxleyi} does not contain METE within its genome, but it can survive in the absence of exogenous B12 (fig. 2f), confirming earlier reports (Guillard 1963; Carlucci and Bowes 1970; Haines and Guillard 1974). Assuming that methionine synthase activity is essential for growth in \textit{Em. huxleyi}, either METE is present in the \textit{Em. huxleyi} genome but absent from the current genome assembly or the growth assays did not reveal the true B12 requirements of \textit{Em. huxleyi}.

The first suggestion is unlikely. METE is not present in either the \textit{Em. huxleyi} (CCMP 1516) genome assembly, the unassembled genomic reads or in the \textit{Em. huxleyi} expressed sequence tag 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 \textit{Em. huxleyi} (supplementary fig. 3, Supplementary Material online). Although this is not conclusive evidence for the absence of the gene, our combined data do not support the presence of METE in the \textit{Em. huxleyi} genome.

We therefore reassessed the B12 requirements of \textit{Em. huxleyi}. We treated cultures of both calcifying and noncalcifying \textit{Em. huxleyi} (CCMP 1516) using extended methods to minimize bacterial contamination (see Materials and Methods).
Both strains were able to grow in the absence of B12 for up to five subcultures (supplementary fig. 4, Supplementary Material online). However, using a bioassay that measures B12, we were able to detect up to 11.7 ng/ml of B12 in lysed cell extracts of the *Em. huxleyi* cell grown for four subcultures in medium without B12. In contrast, Croft et al. (2005) showed that cultures of several other B12-independent algae: *Haematococcus pluvialis*, *Cyanidium caldarium*, *Tetraselmis vermiculosa*, and *C. reinhardtii* have no detectable B12 after just two subcultures. The source of B12 in *Em. huxleyi* is unlikely to be de novo synthesis because only two of the total of 20 enzymes required for aerobic B12 biosynthesis (Warren et al. 2002) were identified in the *Em. huxleyi* genome. These genes exhibited weak similarity to *cobB* and *cobN*.

DAPI-staining of the *Em. huxleyi* cultures did not reveal any associated bacteria. However, when the cultures were plated on high salt LB media, growth of bacteria was observed. Analysis of the 16S ribosomal RNA gene established that the contaminant for the noncalcifying 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 *Em. huxleyi*, this result supports the hypothesis that algal species lacking METE require B12 for growth. The apparent very close association of these bacteria with *Em. 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 B12 and the presence of methionine synthase isoforms, we generated a schematic phylogenetic tree (fig. 5) (Keeling and Palmer 2008), with representatives of the major eukaryote lineages, including land plants, mammals, and fungi. It is clear that the B12 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 B12 auxotrophy, or conversely that multiple independent gains of METE may have led to the evolution of B12 independence. Aside from the major endosymbiotic gene transfer events, abundant horizontal gene transfers (HGTs) have been proposed to encode novel metabolic capacities in some algae, for example in the...
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 (fig. 6A). This argues against multiple HGT events in these lineages and supports the view that the absence of METH in the Cy. 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 (fig. 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 Cy. 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 Cy. 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. Although we have identified no extant Archaeplastida species with both forms of the METE gene to support this hypothesis, coexistence 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 et al. 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 B₁₂ on Expression of METE in P. tricornutum**

A comparison of *E. coli* methionine synthase isoforms found that METH has a much greater catalytic activity (ca. 100-fold) than METE (Gonzalez et al. 1992). This suggests that METH would be used preferentially in the presence of B₁₂, 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 B₁₂ in *C. reinhardtii* (Croft et al. 2005), and we made a similar observation for *G. pectorale* (fig. 4C). Thus, a prolonged continuous supply of B₁₂ 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 B12 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 semiquantitative RT-PCR to analyze the expression of METE and METH from P. tricornutum cells grown in the presence or absence of B12. METH transcript is detectable in both treatments, whereas METE is not expressed in the presence of vitamin B12 (fig. 7A and supplementary fig. 1, Supplementary Material online), and the repression occurs within 14 h after the addition of the vitamin (fig. 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 B12 auxotrophy within the algal kingdom and the possible mechanism by which it arose. Our results indicate strongly that the major determinant for the B12 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 B12 auxotrophy amongst photosynthetic eukaryotes. In contrast, it is unlikely that B12-dependent MCM causes the auxotrophy because it is found in B12-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 B12. Similarly, the absence of MCM in species that have an absolute requirement for B12 (e.g., V. carteri, 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 is responsible both for de novo synthesis of methionine, and the regeneration of methionine following transmethylation reactions. In land plants, at least, it is proposed that these functions occur in different subcellular 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 subcellular compartments. For example, METE is also an abundant flagella protein in C. reinhardtii and has been implicated in protein methylation during flagella resorption (Schneider et al. 2008; Sloboda and Howard 2009). Thus, in C. reinhardtii, there may be a selective pressure to maintain METE due to its specialized role in flagella function, although repression of METE by B12 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 because 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 B12-dependent green alga, in the absence of B12. 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$_{12}$ in the environment, which is produced exclusively by prokaryotes (Warren et al. 2002). Croft et al. (2005) proposed that algae may obtain B$_{12}$ via a symbiotic relationship with bacteria, although a direct symbiosis involving B$_{12}$ 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$_{12}$ 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 we found associated with Em huxleyi. Moreover, a recent study showed that vitamin auxotrophy for B$_{12}$ (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$_{12}$, 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$_{12}$ (amongst other vitamins) regardless of whether they require it or not. Recent measurements indicate the lower range of vitamin B$_{12}$ concentrations in the Southern Ocean and North Atlantic Ocean approach potentially limiting values, with the lowest values being 0.4 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$_{12}$ has been disputed. Chemostat experiments indicate that the growth saturation constants for B$_{12}$ in phytoplankton reside in the subpicomolar range, suggesting that the concentration of B$_{12}$ 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$_{12}$ stimulation of natural phytoplankton assemblages suggesting this cofactor may be limiting in certain areas (Sanudo-Wilhelmy et al. 2006; Bertrand et al. 2007). Although it remains to be determined whether B$_{12}$ 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$_{12}$ 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$_{12}$-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$_{12}$ availability influences both the physiology and the ecology of these organisms.

### Supplementary Material

Supplementary tables 1–3 and figures 1–4 are available at Molecular Biology and Evolution online (http://www.mbe.oxfordjournals.org/).

### Acknowledgments

We would like to thank Herve Moreau (Oceanological Observatory of Banyuls-ser-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 go 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) (Em. huxleyi), James L. Van Etten (Nebraska Center for Virology, University of Nebraska-Lincoln, USA) (Coccomyxa sp. C-169 and Ch. 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. We acknowledge funding from the Biotechnology and Biological Sciences Research Council (BBSRC) of the United Kingdom with grant reference BB/F021844/1, and a studentship for KEH.

### References


