

Vitamin B₁₂ biosynthesis gene diversity in the Ross Sea: the identification of a new group of putative polar B₁₂ biosynthesizers

Erin M. Bertrand,^{1,2,3*} Mak A. Saito,²
Young Jae Jeon³ and Brett A. Neilan^{3,4}

¹MIT/WHOI Joint Program in Chemical Oceanography,
Woods Hole, MA 02543, USA.

²Marine Chemistry and Geochemistry Department,
Woods Hole Oceanographic Institution, Woods Hole,
MA 02543, USA.

³School of Biotechnology and Biomolecular Sciences,
University of New South Wales, Sydney, New South
Wales 2052, Australia.

⁴Australian Centre for Astrobiology, University of New
South Wales, Sydney, New South Wales 2052,
Australia.

Summary

Vitamin B₁₂, a cobalt-containing micronutrient, has been shown to limit phytoplankton growth in the Ross Sea of the Southern Ocean. However, B₁₂ biosynthesis potential in this environment remains uncharacterized. Select bacteria and archaea synthesize B₁₂ while many phytoplankton require it for growth. Low ratios of bacterial biomass production to primary productivity and high concentrations of labile cobalt in Antarctic surface water suggest that factors controlling bacterial growth rather than cobalt availability may determine vitamin production rates here. In order to assess B₁₂ biosynthesis potential, degenerate polymerase chain reaction primers were designed to target the genetic locus *cbiA/cobB*, encoding cobyrinic acid a,c-diamide synthase, a B₁₂ biosynthesis protein. Sequencing the DNA complement of Ross Sea 16S rRNA (see *Supporting information*) allowed targeting of *cbiA/cobB* probes to dominant bacterial groups. *CbiA/cobB* DNA sequences were successfully identified in clone libraries from the Ross Sea. To our knowledge, this study represents the first targeted molecular characterization of environmental B₁₂ biosynthesis potential. A newly identified group of *cbiA/cobB* sequences dominated the diversity of the sequences retrieved; their expression was confirmed

via mass spectrometry-based peptide detection. These sequences seem to have originated from a previously undescribed group of bacteria that could dominate the B₁₂ biosynthesizing community in polar systems.

Introduction

Vitamin B₁₂, cobalamin, is produced by select archaea and bacteria (Rodionov *et al.*, 2003) and is required for growth of some bacteria and archaea (Roth *et al.*, 1996), as well as approximately half of all eukaryotic phytoplankton (Croft *et al.*, 2005). Either through cycling of the microbial loop (Karl, 2002; Droop, 2007) or through direct symbiotic interaction (Croft *et al.*, 2005), bacteria and archaea must be the ultimate source of vitamin B₁₂ for auxotrophic eukaryotic phytoplankton. Therefore, in areas of the ocean where vitamin B₁₂ availability has been shown to be a factor in determining phytoplankton growth and community composition, elucidating controls on bacterial production of vitamin B₁₂ is crucial for understanding primary production. These regions, as shown through previous investigations, include the Ross Sea (Bertrand *et al.*, 2007) as well as the Antarctic Peninsula sector of the Southern Ocean (Panzeca *et al.*, 2006), Long Island embayments (Sañudo-Wilhelmy *et al.*, 2006; Gobler *et al.*, 2007), the Sargasso Sea (Menzel and Spaeth, 1962) and the Gulf of Maine (Swift, 1981). In the Ross Sea in particular, vitamin B₁₂ has only been shown to limit phytoplankton growth where bacterial populations are small, suggesting that the dynamics of bacterial growth play a large role in determining patterns of vitamin limitation of phytoplankton growth in polar seas (Bertrand *et al.*, 2007).

Vitamin B₁₂ production rates must be influenced, in part, by bacterial growth, which in the marine water column is thought to be impacted by temperature, dissolved organic matter (DOM) availability and in some cases the availability of nitrogen or iron. However, B₁₂ biosynthesis rates must also be a function of the proportion of the bacterial community contributing to vitamin biosynthesis. Additionally, because cobalt is required for B₁₂ biosynthesis, cobalt availability has the potential to limit vitamin B₁₂ production. Cobalt is present in the marine water column in picomolar

Received 30 September, 2010; accepted 6 January, 2011. *For correspondence. E-mail ebertrand@whoi.edu; Tel. (+61) 508 289 2835; Fax (+61) 508 457 2075.

concentrations (Saito and Moffett, 2002) and in fact, there is evidence that vitamin B₁₂ production in surface waters of the North Atlantic may be limited by cobalt availability (Panzeca *et al.*, 2008).

The Ross Sea is home to massive phytoplankton blooms (Smith and Nelson, 1985) and rapid rates of biomass export out of the surface ocean (DiTullio *et al.*, 2000; Buesseler *et al.*, 2001). Primary production in this region has been shown to be controlled by iron availability as well as irradiance (Martin *et al.*, 1990; Sedwick *et al.*, 2000; Coale *et al.*, 2003) and vitamin B₁₂ availability (Bertrand *et al.*, 2007). Ratios of bacterial biomass production to primary production in the Ross Sea are exceedingly low (Ducklow *et al.*, 2000; 2001), suggesting that sources of B₁₂ relative to sinks might be smaller here than in other regions of the ocean. Cyanobacterial populations in the Ross Sea are extremely small (Marchant *et al.*, 1987; Caron *et al.*, 2000); given the dominance of these organisms in the temperate and tropical oceans (Partensky *et al.*, 1999; Scanlan and West, 2002) and the fact that marine cyanobacteria are known to produce vitamin B₁₂ (Bonnet *et al.*, 2010), the Ross Sea and other polar regions are likely missing a large source of B₁₂ that is available in other locations. Bacterial growth in the Southern Ocean has been observed to be limited by the availability of DOM (Ducklow *et al.*, 2001; Hall and Safi, 2001; Oliver *et al.*, 2004), iron (Pakulski *et al.*, 1996), and combined iron and DOM (Church *et al.*, 2000). Because vitamin B₁₂ production rates must be determined, in part, by bacterial growth, iron and DOM availability may thus impact rates of vitamin production.

One approach to investigate factors regulating B₁₂ biosynthesis rates is to develop molecular biological tools to assess and monitor vitamin production potential by characterizing the portion of the bacterial community capable of vitamin biosynthesis. Biosynthesis of vitamin B₁₂ requires a large number of enzymatic steps, involving over 30 biosynthesis genes (Roth *et al.*, 1996; Raux *et al.*, 2000). *In silico* analyses of sequenced microbial genomes reveal that the genetic potential to synthesize the vitamin is widespread throughout most bacterial and archaeal phyla (Rodionov *et al.*, 2003). However, not all members of a given taxonomic group possess the genetic machinery for B₁₂ biosynthesis (Rodionov *et al.*, 2003), making it difficult to draw conclusions about the B₁₂ biosynthesis potential of a microbial community based on traditional taxonomic assessment of its composition. For instance, of the forty-five strains of *Vibrio* with sequenced genomes, only two (*Vibrio splendidus* LGP32 and *Vibrio* MED222) have the genetic machinery for *de novo* B₁₂ biosynthesis (E. Bertrand, unpubl. obs.). Therefore, when a *Vibrio* strain is identified in an environmental sample by 16S rRNA gene sequencing, it is not possible to infer whether that strain is capable of B₁₂ biosynthesis.

As a result, the genetic potential for vitamin B₁₂ production remains almost entirely uncharacterized in the Ross Sea or any other environment.

There are currently five bacterioplankton strains isolated from Antarctic seawater with fully sequenced genomes (as of July 2010). *Pseudoalteromonas haloplanktis* TAC125 (Medigue *et al.*, 2005) and *Shewanella frigidimarina* NCIMB 400 possess genes encoding the B₁₂-requiring enzyme methionine synthase (MetH) as well as a B₁₂ uptake system and the last few genes in the biosynthesis pathway, rendering these strains capable of B₁₂ scavenging and salvaging, but not *de novo* synthesis. Draft genomes of the Antarctic marine strains *Agreia* PHSC20C1 and *Polaribacter irgensii* 23-P contain no B₁₂ biosynthesis related genes, and no genes encoding known B₁₂-dependent enzymes while the draft genome of *Psychroflexus torques* ATCC 70755 seems to contain at least a partial B₁₂ uptake system as well as genes encoding the vitamin B₁₂-requiring enzymes methionine synthase (MetH) and ribonucleotide reductase (RNR II). Although limited in scope, these genomes are currently the only molecular insight into bacterial B₁₂ metabolism in the Antarctic Southern Ocean. Because none of these genomes contain the potential for *de novo* biosynthesis, they provide no information regarding the identity of primary vitamin biosynthesizers in this vast region.

In order to further characterize cycling and sources of B₁₂ in the Ross Sea, a set of DNA probes were developed to profile B₁₂ biosynthesis genetics in the bacterial communities of the Ross Sea. The genes *cbiA* and *cobB*, which encode cobyrinic acid a,c-diamide synthase, were targeted as a marker for vitamin B₁₂ biosynthesis potential. Cobyrinic acid a,c-diamide synthase is comprised of two domains – one ATP binding domain and one glutamine amidotransferase domain (Debussche *et al.*, 1990; Galperin and Grishin, 2000). The gene products catalyse the amidation of two sites on the corrin ring of vitamin B₁₂ (Fresquet *et al.*, 2004).

Functional gene diversity has been profiled using this approach for many genes critical to biogeochemical cycles, e.g. ammonia monooxygenase and dissimilatory sulfate reductase (Wagner *et al.*, 1998; Hansel *et al.*, 2008), but has yet to be applied to vitamin biosynthesis. The development of these probes has elucidated the importance of particular taxonomic groups in specific processes (Francis *et al.*, 2005) has led to a better understanding of the processes controlling biogeochemical fluxes and has also opened further avenues for research (Church *et al.*, 2005; Steward *et al.*, 2005). This study investigates diversity of a gene involved in vitamin B₁₂ biosynthesis, verifies the presence of the protein it encodes, and places these measurements in the context of bacterial abundance as well as cobalt concentrations and lability in the Ross Sea of the Southern Ocean.

Results and discussion

Relevant characteristics of study site

Bacterial abundance and cobalt concentrations and speciation from station NX 19 in the Ross Sea are shown in Table 1. Bacterial abundance was relatively low for marine environments, but was among the higher values observed in the Ross Sea (Ducklow, 2000) and fell within the range of bacterial abundance in which vitamin limitation of primary productivity was previously observed (Bertrand *et al.*, 2007). Total cobalt concentrations were approximately 30 pM throughout the upper 100 m of the water column at this station. These values were within the range of cobalt concentrations previously observed in the ocean [~5–100 pM; (Saito and Moffett, 2002)]. Most importantly, labile cobalt (Co'), the fraction that is thought to be bioavailable, was present at approximately 10 pM throughout the upper water column. These concentrations were at least an order of magnitude larger than other measurements of surface water Co' outside of upwelling regions, suggesting that cobalt bioavailability at this study location is higher than in other areas of the surface ocean (Saito and Moffett, 2001; Saito *et al.*, 2004; 2005; 2010). When considered together with the low bacterial biomass production to primary production ratios observed in this region (Ducklow, 2000), these cobalt speciation data suggested that bacterial metabolism rather than cobalt availability may control the rates of vitamin production in the Ross Sea as bioavailable Co is present in high concentrations. This reinforces the importance of elucidating which bacterial groups are responsible for vitamin production in this region and what factors control their growth and activity.

Table 1. Bacterial abundance and cobalt concentrations from station NX 19.

Depth (m)	Bacteria (10 ⁵ cells ml ⁻¹)	total dissolved cobalt (pM)	labile cobalt (pM)	Per cent labile (%)
20	1.87	30.6	12.1	39.6
40	2.25	34.6	13.3	38.4
60	1.41	31.9	9.8	30.9
100	1.44	34.1	11.4	33.5

Bacterial abundance was among the higher values observed in the Ross Sea, although this was relatively low when compared with other marine environments. Total cobalt and labile cobalt, the fraction that is thought to be most bioavailable, are also shown. Labile cobalt concentrations here were among the highest surface water values observed using this method. These data show that cobalt availability in these waters was high relative to other marine waters and suggests that cobalt availability is less likely to control the rate of vitamin B₁₂ production in the Ross Sea compared with other marine environments.

Design and application of degenerate primers targeting B₁₂ synthesis genes

In this study, several essential genes coding for enzymes involved in vitamin B₁₂ biosynthesis were initially targeted for DNA probe development including *cbiA/cobB* encoding cobyrinic acid a,c-diamide synthase, *cbiC/cobH* encoding precorrin-8x methylmutase, and *cobT* encoding nicotinate mononucleotide:5,6-dimethylbenzimidazole phosphoribosyltransferase. Each of these genes represent a potential biomarker for vitamin biosynthesis because: (i) when they are present in fully sequenced bacterial and archaeal genomes, the complete B₁₂ biosynthesis pathway is also present and (ii) the genes are homologous in both the oxygen-requiring (*cobB*, *cobH*, *cobT*) and non-oxygen-requiring pathways for vitamin synthesis (*cbiA*, *cbiC*, *cobT*), and thus one set of probes could be designed to target both versions of the biosynthesis pathway. However, attempts to design probes to target the full diversity of these genes as represented in the National Center for Biotechnology Information non-redundant (NCBI nr) database failed because of the high degeneracy of the primers designed. This high degree of degeneracy was required because of extensive diversity between homologous sequences. The B₁₂ biosynthesis pathway is one of the most ancient metabolisms (Raux *et al.*, 1999) and the sequences of the associated biosynthesis genes are highly divergent (Rodionov *et al.*, 2003).

CbiA/cobB is among the most highly conserved genes in the B₁₂ biosynthesis pathway, making it more conducive to DNA probe design. Because initial probe design targeting the full diversity of these sequences in the NCBI nr database failed, clone libraries were created from the amplified DNA complement of 16S rRNA genes from the Ross Sea and sequencing was performed in order to identify the major taxonomic groups present (see *Supporting information*). The major groups identified in this study (Alpha- and Gammaproteobacteria and Bacteroidetes; Fig. S1, Table S1) were similar to those found in other 16S rRNA gene library sequencing studies in pristine Antarctic surface waters (Webster *et al.*, 2004; Gentile *et al.*, 2006; Murray and Grzymalski, 2007). Degenerate primers were then successfully designed to target *cbiA/cobB* genes in the Alpha- and Gammaproteobacteria and Bacteroidetes based on *cbiA/cobB* sequences from these groups in the NCBI nr database.

Application of *cbiA/cobB* primers

Figure 1 shows the conserved domains in the CbiA/CobB protein and the genomic loci targeted by the primers designed in this study. These primers were applied to

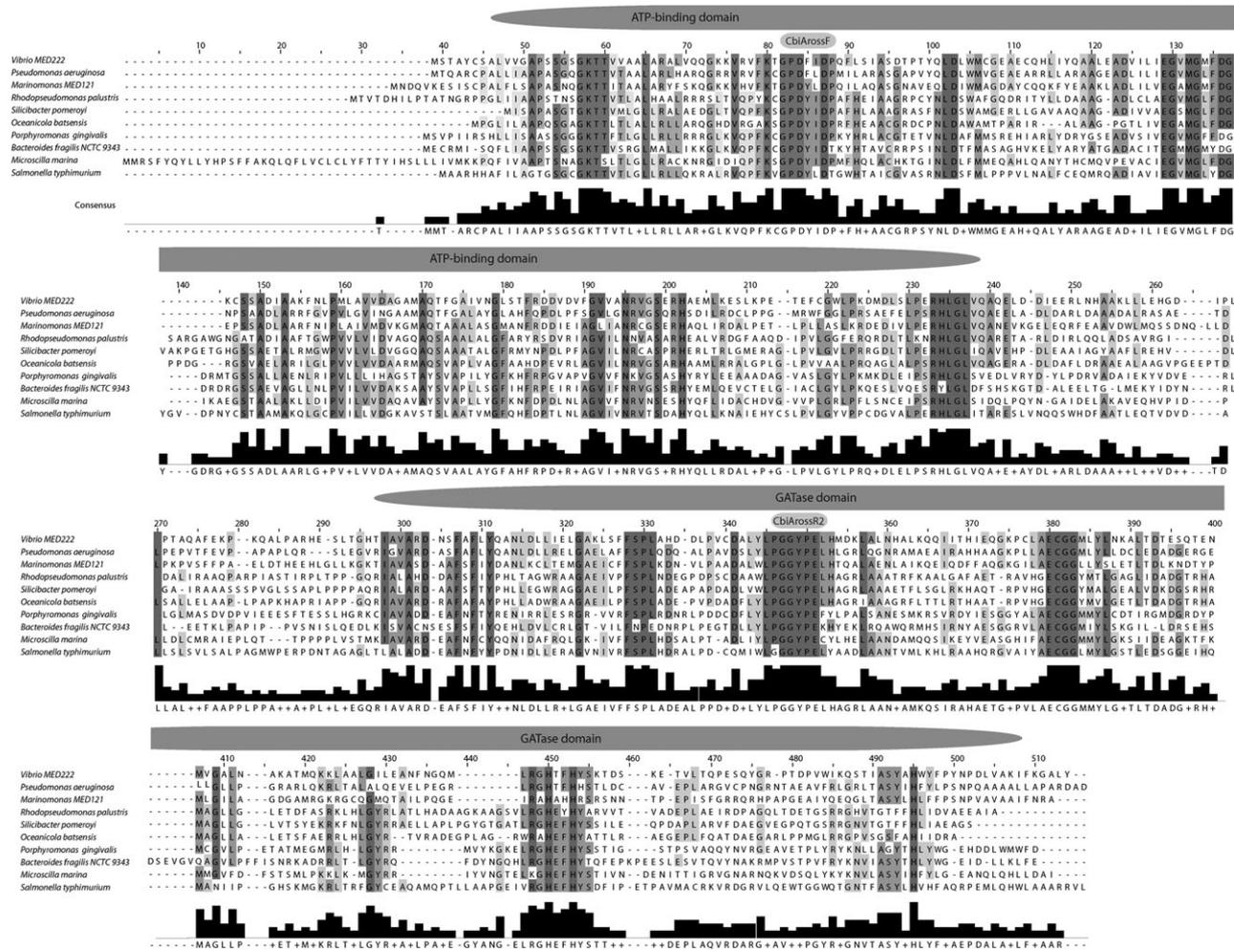


Fig. 1. A Clustal X multiple alignment of CbiA/CobB amino acid sequences from representatives of the Alpha- and Gammaproteobacteria and the Bacteroidetes groups obtained from the NCBI nr database. The alignment was used to design the B₁₂ biosynthesis gene primers used in this study. The most highly conserved regions of the alignment are highlighted in the darkest colours. The location of the primers (CbiArossF and CbiArossR2) are denoted by the light gray ovals above the sequence. The conserved functional domains of the protein, one ATP binding domain and one glutamine amidotransferase domain (Debusche *et al.*, 1990; Galperin and Grishin, 2000), are denoted by darker gray ovals above the sequence.

samples from station NX19, taken at four depths within the upper water column. The libraries created using these primers contained 10–30% *cbiA* sequences along with a variety of sequences from other genes that were amplified as a result of primer degeneracy. Only the sequences homologous to *cbiA/cobB*, as determined by blastx (Altschul *et al.*, 1990) searches against the NCBI nr database, were used for further analyses. From this location, 20 different operational taxonomic units (95% similarity) were recovered from 29 sequences. The results are shown in Fig. 2, reported as a neighbour-joining tree of partial amino acid sequences. The tree is rooted to a sequence of CTP synthetase (PyrG) from *Marinomonas* sp. PyrG catalyses the ATP-dependent amidation of dUTP to form dCTP and is thought to be ancestral to CbiA and

CobB (Leipe *et al.*, 2002). This tree also includes sequences recovered via a blastp search of the sequence CbiA20m167 against NCBI nr and the Community Cyber-infrastructure for Advanced Microbial Ecology Research & Analysis (CAMERA) all metagenomic open reading frame peptides database (Seshadri *et al.*, 2007) on 19 July 2010 using default parameters. Sequences with the lowest *E*-value blastp results are included in Fig. 2. This tree also includes sequences from eukaryotic algal genomes (*Ostreococcus lucimarinus* and *Fragilariopsis cylindrus*, obtained from the Joint Genome Institute's genome portal) that align with bacterial CbiA/CobB. The currently sequenced eukaryotic algal genomes with blastp homologues to CbiA/CobB sequences (those in Fig. 2 as well as *Ostreococcus tauri*, *Micromonas* sp. and *Chlorella* sp.)

do not appear to contain full canonical B₁₂ biosynthesis pathways, although the genome of *Fragilariopsis cylindrus* appears to contain more potentially B₁₂ biosynthesis related proteins than any other eukaryotic algal genome currently available. The significance of the presence of these partial vitamin biosynthesis pathways has yet to be investigated.

Analysis of Group RSB₁₂ sequences

Seventy per cent of the CbiA/CobB sequences retrieved in this study, while homologous to known CbiA/CobB sequences, fell into a distinct group herein referred to as Group RSB₁₂ (Fig. 2). The best blastp matches in NCBI nr to Group RSB₁₂ sequences were included in this tree but did not fall within Group RSB₁₂ (CbiA from *Oceanospirillum* sp. MED92; EAR62324.1); it is evident that this group has no sequenced or otherwise characterized representatives. There are also very few representatives of this group in available metagenomic sequencing databases (Table 2); no sequences from the HOT, BATS or GOS databases in CAMERA fell into this group. There was one sequence from Newcomb Bay of the Antarctic Southern Ocean that exactly matched a sequence from this clone library and fell into Group RSB₁₂. There were also short sequence reads from the Norwegian Fjords that were very similar to a portion of the Group RSB₁₂ sequence, but as these were so short, it was difficult to accurately include them in the Fig. 2 tree and thus determine if they belong to this new group. The lack of Group RSB₁₂ sequence homologues in current databases is perhaps due to the lack of polar metagenome sequence data and the limited genomic sequencing of polar-derived cultured isolates. The complete lack of Group RSB₁₂ sequences in tropical, subtropical and sub-polar metagenomic databases along with its dominance in this clone library suggested that Group RSB₁₂ could be a major contributor to the B₁₂-biosynthesizing community particular to polar regions. While the broad phylogenetic groupings of bacteria found in the Antarctic marine water column are shared with other planktonic ecosystems, there are some species or specific groups that are thought to be predominantly associated with polar systems, including *P. irgensii* as well as some uncultivated gammaproteobacterial species (Murray and Grzymalski, 2007). There is thus some precedent for the increased frequency of specific phylogenetic groups in polar regions, as may be the case for organisms encoding Group RSB₁₂ sequences.

It is notable that *Fragilariopsis cylindrus* appears to possess multiple copies of CbiA/CobB-like genes and that one clusters close to Group RSB₁₂ (Fig. 2). The CbiA/CobB sequences recovered in this study are not likely eukaryotic, as the sample from which the original

DNA was extracted was 2 µm prefiltered. In addition, the 16S sequencing results in Fig. S1 suggest that there was only a small amount of eukaryotic DNA available in these samples because about 10% of the sequences retrieved appear to be eukaryotic in origin. Additionally, within the three longer Antarctic metagenomic reads that contain sequences in Group RSB₁₂ (Table 2), one read contains a partial gene sequence encoding corrinoid adenosyltransferase BtuR/CobO/CobP adjacent to the *cbiA* sequence. This is evidence that this group of RSB₁₂ CbiA – encoding sequences was not isolated from eukaryotes as in many bacterial genomes, B₁₂ biosynthesis genes are adjacent to each other (Rodionov *et al.*, 2003), while in eukaryotic genomes they are not (E. Bertrand, unpubl. obs.).

It was of interest to identify the taxonomic group from which the Group RSB₁₂ sequences originate. Although the primers to target these genes were designed to target only Alpha- and Gammaproteobacteria and the Bacteroidetes, they were not designed to specifically exclude other groups. Therefore, the identities of the bacteria possessing these sequences cannot be narrowed down on this basis. The Group RSB₁₂ sequences clearly clustered with partial CbiA/CobB sequences from Gammaproteobacteria (Fig. 2). However, given the extensive lateral gene transfer that has occurred with B₁₂ biosynthesis genes, it is not possible to definitively confirm that Group RSB₁₂ sequences originate from a Gammaproteobacterial group. Within the three longer metagenomic reads that contain sequences similar to Group RSB₁₂ (Table 2), one also has a partial gene sequence encoding corrinoid adenosyltransferase BtuR/CobO/CobP. The best blastx matches to this sequence in the NCBI nr database were also all from Gammaproteobacteria, further suggesting that Group RSB₁₂ may be Gammaproteobacterial in origin.

Additional insights into the novel origin of this vitamin biosynthesis gene could potentially be achieved by determining the guanine and cytosine (G + C) content and codon usage of the clone library sequences. The mean G + C content of the *cbiA/cobB* sequences from Group RSB₁₂ was 48.9% with a standard deviation of 1.2%. If the per cent G + C content of individual protein coding genes reflects the G + C content of the genomes in which they reside (Muto and Osawa, 1987), then the G + C content of this sequence may be useful in ruling out certain groups of bacteria as sources of these sequences. Unfortunately, no firm conclusions could be made about the origin of these *cbiA/cobB* sequences because their G + C content was in no way unusual, and it is possible that the partial sequences available for these *cbiA/cobB* genes are not long enough to be representative of the entire genomes of organisms from which they originate.

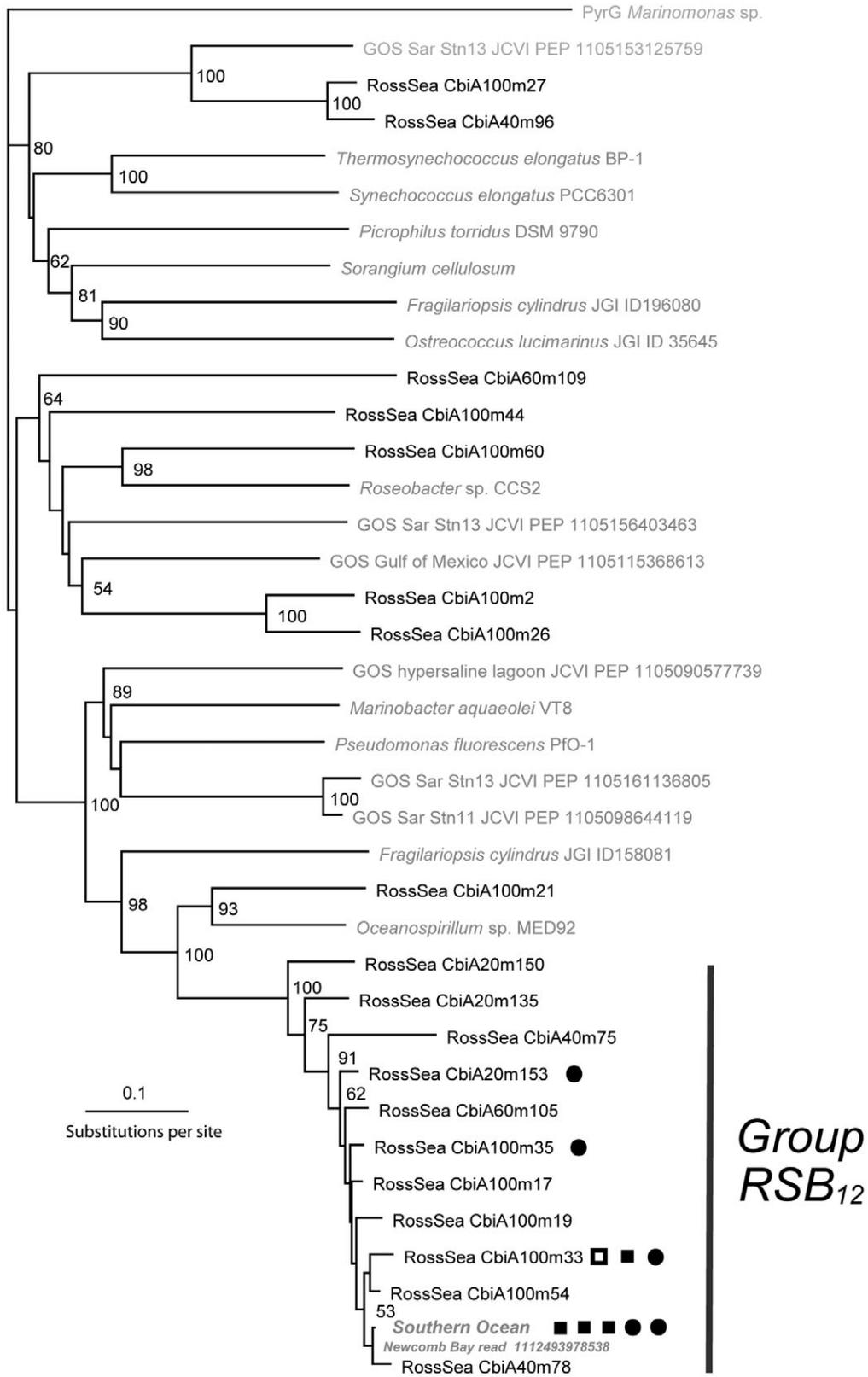


Fig. 2. Neighbour-joining tree of partial CbiA amino acid sequences obtained from the Ross Sea. Sequences obtained in this study from station NX19, 20–100 m depth, are shown in black. One representative of sequence groups with > 95% identity at the nucleic acid level is shown. Each symbol next to a sequence name indicates one additional sequence obtained from the Ross Sea belonging to that group. Closed squares denote that the additional sequence came from 100 m, circles from 20 m, and the open square from 40 m depth. The tree is rooted to a sequence of CTP synthetase (PyrG) from *Marinomonas* sp. PyrG is thought to be ancestral to CbiA and CobB (Leipe *et al.*, 2002). Bootstrap values > 50%, calculated from 4000 replicates, are shown as per cent. Reference sequences from NCBI, the Global Ocean Sequencing dataset, eukaryotic genome sequences from JGI's genome portal, and meta sequence data from the CAMERA database (in gray), including the closest representatives to all environmental samples recovered are shown. Twenty different operational taxonomic units (95% sequence similarity at the nucleic acid level) were recovered from 29 clone sequences. Seventy per cent of the sequences retrieved fell into a distinct group herein referred to as Group RSB₁₂ while the others grouped with sequences represented in available the genomic and metagenomic databases.

Confirmation of expression of Group RSB₁₂ CbiA proteins – peptide biomarkers

In order to evaluate whether Group RSB₁₂ CbiA proteins were expressed in the Ross Sea, we attempted to detect peptides coded for by these sequences using two separate mass spectrometry-based approaches. We conducted global metaproteomic analyses employing reverse phase liquid chromatography coupled to linear ion trap mass spectrometry via electrospray ionization on samples from the Ross Sea. Peptides representing the RSB₁₂ sequences or other CbiA/CobB sequences were not detected, perhaps because of the inability of these global proteomic approaches to detect low abundance components of complex mixtures (data not shown). As an alternative, we employed a targeted proteomic approach using reverse phase liquid chromatography separation coupled to triple quadrupole mass spectrometry with selected reaction monitoring (SRM) using stable isotopically labelled peptide reference standards. SRM is a type of tandem mass spectrometry, meaning that it detects a parent molecular ion and isolates it, then fragments it and detects resulting fragment ions. This method is much more sensitive than the global ion trap mass spectrometry approach in part because only a narrow range of masses are assayed by the mass spectrometer, increasing the likelihood that low abundance components of complex mixtures can be detected.

Three tryptic peptides – peptides produced in the digestion of intact proteins with the enzyme trypsin – were identified as potential biomarkers for Group RSB₁₂ proteins (Table 3). These are coded for by over 70% of the Group RSB₁₂ DNA sequences detected in this study, not encoded by any other known sequences (confirmed via a NCBI nr database search on 15 July, 2010 and a search of JGI eukaryotic algal sequences not yet deposited in GenBank), and are also theoretically observable via the mass spectrometry techniques applied here. As shown in Fig. 3, one of these three peptides was detected via SRM mass spectrometry in a Ross Sea protein sample (peptide sequence: HLGLVQAHEVR). There were three pieces of evidence supporting the claim this peptide from the Group RSB₁₂ CbiA protein was detected. First, the retention time at which the peak for the native

peptide identified here was the same as the synthetic, isotopically labelled standard version. Second, the product distribution patterns (relative peak heights of fragment ions) at this time point were the same for the standard as for the native peptide. Third, this signal was only observed in the protein gel slice that corresponded to the molecular weight of CbiA proteins (43–50 kDa). In addition, none of these peptides were detected in samples of *Crocospaera watsonii* peptides, which served as a negative control as its genome sequence does not encode these peptides. These data suggest that the CbiA protein encoded by Group RSB₁₂ sequences is expressed in the Ross Sea, but that it is among the lower abundance proteins. These data also confirm the potential of targeted proteomic assays, when directed by detailed nucleic acid sequencing information, to confirm the metabolic activity of particular groups of organisms in complex environmental samples.

This study shows that there is a previously unidentified group of *cbiA/cobB* vitamin B₁₂ biosynthesis genes in the Ross Sea of the Southern Ocean that actively produce protein. However, because there are many other proteins required for vitamin biosynthesis, this work does not prove that the organisms expressing Group RSB₁₂ sequences can or do produce vitamin B₁₂. Even though currently sequenced bacterial genomes that possess *cbiA/cobB* sequences also contain the entire vitamin B₁₂ biosynthesis pathway, the fact that only one of many involved genes was documented here is especially important in light of the fact that some eukaryotic algal genomes appear to possess *cbiA/cobB*-like sequences without the rest of the B₁₂ biosynthesis pathway. Hence, this study has identified a putative new group of vitamin B₁₂ biosynthesizers in the Southern Ocean. Further study will be required to link this new group of gene sequences to specific microbes and verify that they in fact produce the vitamin. Once the bacteria encoding Group RSB₁₂ sequences have been identified, these organisms must then be isolated, cultured and manipulated in the laboratory to prove that they do in fact produce vitamin B₁₂.

Once the vitamin biosynthesizing role of these bacteria is confirmed, the distribution, activity and biogeochemical constraints on their growth can be studied to yield insight into controls on B₁₂ production. Further study should also

Table 2. Metagenomic sequence reads similar to group RSB₁₂.

Database	Number of blastn hits (E < 1e-5)	Read IDs for read E < 1e-5	E-value	Read length (bp)	Location	Lat	Long
Antarctic Aquatic	3	NCBI_READ_1112493978538	0	902	Newcomb Bay, Antarctica	66.27	110.533
		NCBI_READ_1112533230362 ^a	4.09E-122	1003	Antarctic Open Water	63.891	112.073
		NCBI_READ_1112493887005	1.24E-26	1013	Newcomb Bay, Antarctica	66.27	110.533
Marine Metagenome from Coastal Waters project at Plymouth Marine Laboratory	4	PML_READ_00281253	3.35E-123	249	Norwegian Fjord	60.269	5.222
		PML_READ_00287170	3.35E-123	249	Norwegian Fjord	60.269	5.222
		PML_READ_00261281	3.35E-123	249	Norwegian Fjord	60.269	5.222
		PML_READ_00343394	1.12E-116	248	Norwegian Fjord	60.269	5.222
BATS: all Metagenomic 454 reads	0	n/a	n/a	n/a	n/a	n/a	n/a
HOT: all metagenomic sequence reads	0	n/a	n/a	n/a	n/a	n/a	n/a
GOS: all metagenomic sequence reads	0	n/a	n/a	n/a	n/a	n/a	n/a

a. Read also includes a partial cobalamin adenosyltransferase (*btuR/cobO/cobP*) encoding sequence, with best blastx identities to gammaproteobacterial adenosyltransferases. Databases in CAMERA were searched via blastn for matches to sequence CbiA20m167, employing default parameters. There were no blastn matches to group RSB₁₂ sequences in available tropical or subtropical sequencing data and only a few matches were found in the small amount of available polar sequencing data. Possible matches, although on very short reads, were found in sequencing data from sub-polar fjords.

identify other vitamin producers and quantify their relative contributions to vitamin availability in the coastal Antarctic Southern Ocean. The relationships between the distributions of these bacteria, their vitamin biosynthesis protein expression and other bacterial groups and processes will lend insight into how vitamin B₁₂ availability and production may also impact bacterial population dynamics. Measuring the abundance and distribution of peptides like those targeted in this study may prove to be an effective method for determining biological activity in a range of geochemically relevant processes in addition to vitamin biosynthesis.

Summary

Factors controlling vitamin B₁₂ availability remain largely unknown. In order to determine B₁₂ supply to phytoplankton, the rate at which the vitamin is transferred from the particulate to the dissolved pool must be determined. This flux will be a function of bacterial community composition; biosynthesis rate within that population; the rate of population turnover by grazing, viral lysis, and other factors, and the efficiency of the microbial loop. Characterizing the microbial community that produces vitamin B₁₂ represents a starting point from which to explore this question.

Primers targeting *cbiA/cobB* were successfully designed and applied to DNA samples from the water column of the Ross Sea. These primers are capable of amplifying *cbiA/cobB* genes in the major groups of bacteria found by 16S rRNA gene sequence analysis at this study site (see *Supporting information*). This study was therefore restricted to detecting functional gene sequences specifically targeted by these primers and as a result was limited to observing only partial diversity of *cbiA/cobB* genes in the Ross Sea. Most of the *cbiA* sequences retrieved fell into a distinct group, Group RSB₁₂, which is not represented in most nucleic acid sequence databases. We hypothesize that this group represents an important subset of B₁₂ biosynthesizers in polar regions and probably originates from members of the Gammaproteobacterial lineage. Detection of peptides characteristic of Group RSB₁₂ CbiA confirmed the expression of these proteins. These sequences comprise the first molecular description of vitamin B₁₂ biosynthesis potential in the Ross Sea, where limitation of phytoplankton growth by vitamin availability has been observed (Bertrand *et al.*, 2007).

Experimental procedures

Sample collection

Whole water samples were collected at 20, 40, 60 and 100 m depths at Station NX19 (30 November, 2006; 76.5°S,

Table 3. Peptides tested as potential Group RSB₁₂ biomarkers, the per cent of the Group RSB₁₂ sequences coding for each peptide, the peptide sequence for the native and synthetic heavy isotopically labelled versions and the selected reaction monitoring (SRM) reactions.

In % of Group RSB ₁₂	Peptide name	Peptide sequence	Parent m/z	Product m/z	Collision energy (V)
75	CbiA_RSB12_1 native	DPEVSLPER	521.264	601.330	19
				401.214	20
				175.119	34
				700.399	21
				608.348	19
75	CbiA_RSB12_1 heavy	DPEVS[L_C13N15]PER	525.044	401.214	20
				175.119	34
				707.413	21
				1008.559	28
				838.453	31
85	CbiA_RSB12_2 native	HLGLVQAHEVR	629.854	739.386	27
				1121.643	28
				1015.576	28
				838.453	31
				739.386	27
85	CbiA_RSB12_2 heavy	HLG[L_C13N15]VQAHEVR	633.363	1128.660	28
				788.463	23
				416.262	25
				916.521	23
				175.119	39
70	CbiA_RSB12_3 native	LLQGTTIGIAR	571.848	788.463	23
				416.262	25
				916.521	23
				175.119	39
				788.463	23
70	CbiA_RSB12_3 heavy	L[L_C13N15]QGTTIGIAR	575.357	416.262	25
				916.521	23
				175.119	39
				788.463	23
				416.262	25

SRMs are described by the mass to charge (m/z) ratios indicative of the particular parent and product ions and collision energies designed to detect the peptides. Collision energy, a measure of the energy applied to the gas that collides with parent ions in the mass spectrometer to produce fragmentation, was optimized to most efficiently generate each product ion.

174.3°W) in the Ross Sea on the CORSACS2 cruise. Water was collected in 10 l trace metal clean Go-Flo bottles (General Oceanics, FL, USA). For DNA samples, approximately 500 ml of water per sample was 2 µm prefiltered and then passed through a 0.2 µm polyethersulfone membrane filter that was retained for DNA extraction and stored at -80°C. Water from the same sampling events was used to count bacterial abundance and quantify total cobalt and labile cobalt concentrations through electrochemical techniques. Protein samples were collected in the Ross Sea on the CORSACS1 cruise (NBP0601) from two stations. For the Station 1 sample (21 December, 2006; 61.4°S, 178.2°W; Southern Ocean, not Ross Sea), 20 l was collected from 10 m depth using a trace metal clean diaphragm pumping system (Bruland *et al.*, 2005), filtered onto a 0.4 µm polycarbonate filter and stored dry at -80°C. For the Station NX 8 sample (12 January 2006; 77.3°S, 175.4°W; Ross Sea) 9 l of water from 10 m and 9 l of water from 30 m was collected as above using Go-Flo bottles and filtered onto a 0.4 µm nitrocellulose filter and stored dry at -80°C. Bacteria and archaea were enumerated using a DAPI (4,6-diamidino-2-phenylindole) staining method (Porter and Feig, 1980).

Cobalt total concentration and speciation measurements

Total cobalt and cobalt speciation analyses were performed by adsorptive cathodic stripping voltammetry using a Metrohm 663 hanging mercury drop electrode and Eco-

Chemie Autolab III as described previously (Saito and Moffett, 2002; Saito *et al.*, 2004). Briefly, cobalt total measurements were made by first UV-irradiating the seawater for 1 h to destroy the strong organic ligands that bind cobalt. Seawater (9.25 ml) was then analysed with 0.2 mM dimethylglyoxime, 0.113 M sodium nitrite and 2.5 mM N-(2-hydroxyethyl)piperazine-N-(3-propanesulfonic acid), as cobalt-ligand, catalyst and buffer respectively. Cathodic stripping voltammetry analysis included a deposition step for 90 s at -0.6 V, and linear sweep stripping from -0.6 V to -1.4 V at 10 V s⁻¹ with successive 25 pM standard CoCl₂ additions. Cobalt speciation was measured similarly but without UV-irradiation and with overnight equilibration with the dimethylglyoxime ligand. Labile cobalt is that which binds to this competitor ligand under the experimental conditions.

DNA extraction

Filter membranes were suspended in a 100 mM NaCl, 100 mM Tris, 1 mM NaCitrate, 5 mM CaCl₂, 25 mM EDTA buffer at pH 8 in the presence of 0.2% w/v pyrophosphate and 0.4 mg ml⁻¹ Poly A and incubated at 37°C for 1 h. Five mg ml⁻¹ lysozyme, 0.3% w/v SDS, 2 mg ml⁻¹ proteinase K, were added and the samples were incubated for 30 min at 50°C after which the SDS concentration was increased to 5%. Samples were then lysed by three freeze-thaw cycles in liquid nitrogen for 5 min and at 60°C until heated through. DNA was extracted twice, each with one volume of 24:24:1 phenol: chloroform: isoamyl-alcohol and then precipitated

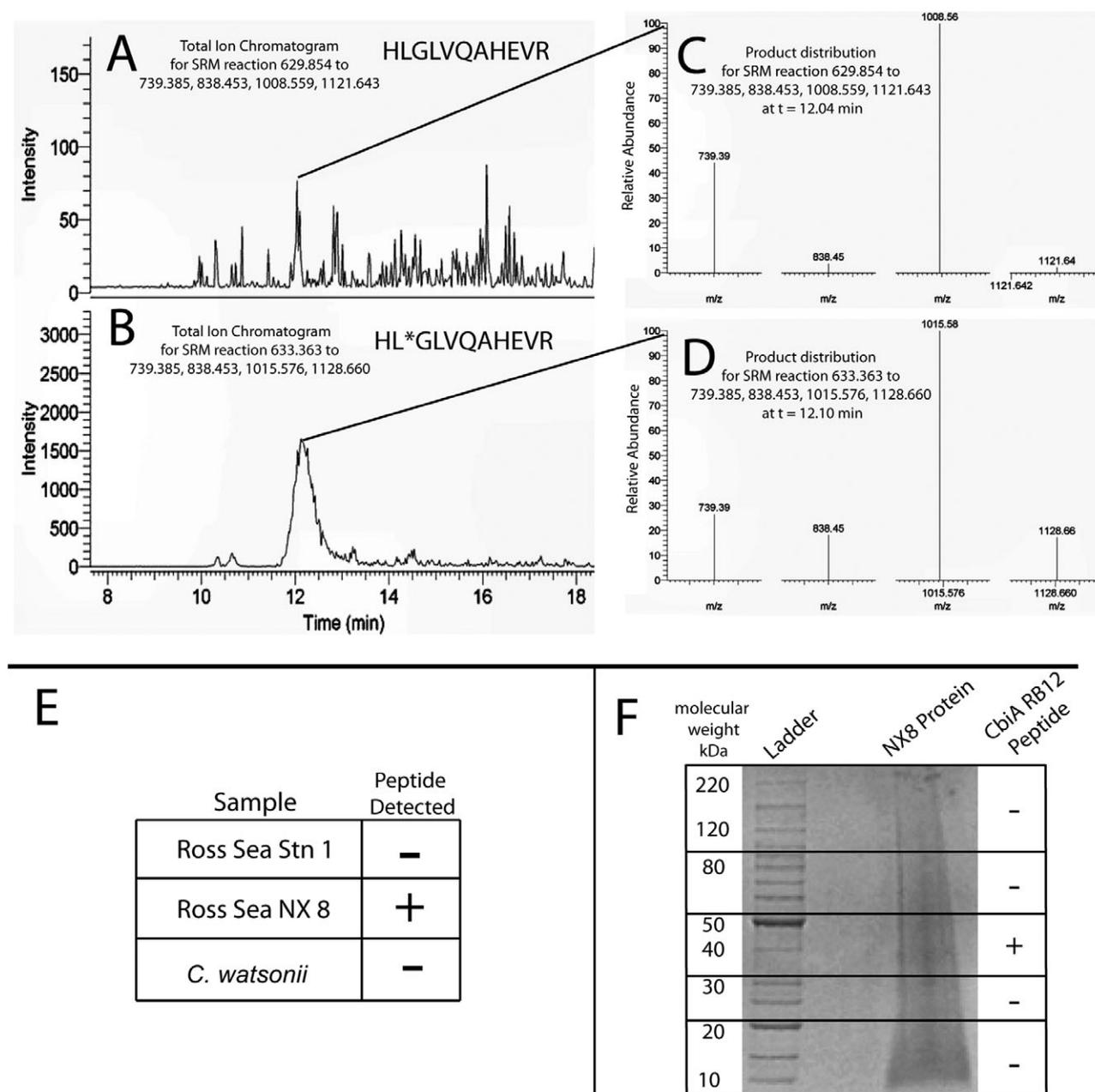


Fig. 3. Evidence supporting the detection of Group RSB₁₂ peptide HLGLVQAHEVR in a Ross Sea protein sample. A shows the total ion chromatogram for the selected reaction monitoring of the native peptide HLGLVQAHEVR (parent ion doubly charged with m/z 629.854 is isolated and fragmented into product ions 1008.559, 838.453, 739.386 and 1121.643) while B shows the selected reaction monitoring of the corresponding heavy labelled peptide (parent ion m/z 633.363 to product ions 1008.559, 838.453, 739.385 and 1128.660). Total ion chromatograms represent the sum of the diagnostic product (fragment ion) intensities created upon fragmenting the isolated parent ions. C shows the product ion distribution at a retention time of 12.1 min for the native peptide while D shows a similar product distribution for the heavy labelled peptide in that same time span. E shows that this peptide was identified in the station NX8 sample, but not the Station 1 sample or the *Crocospaera* control, while F shows that the peptide was only detected in the SDS-PAGE fraction of the station NX8 protein sample containing the mass range that should include the CbiA protein.

three times each with one volume of isopropanol and 0.1 volume of 3 M sodium acetate at pH 5.2 at -20°C for 1 h to overnight and then spun at 12 000 *g* at 4°C for 20 min. Pellets were rinsed once with 70% ethanol and then resuspended in water for further analyses.

Initial B₁₂ biosynthesis gene primer designs

DNA probe design was attempted for *cbiA/cobB* and *cbiC/cobH* genes with the goal of targeting all the diversity of these genes as represented in the NCBI nr database. Probes

tested for *cbiA/cobB* include CbiAF (5'-GGTCCTGAYTWYHTNGA-3'), CbiAF2 (5'-GGTCCNGAYTWYHTNGA-3'), CbiAR2 (5'-TAAYTSNGGRWANCCNCC-3') and CbiAR (5'-TTGTCTSCRCAYTCNCC-3'). Probes tested for *cbiC/cobH* include CbiCF (5'-GGTAACGCNCCNACNGC-3') and CbiCR (5'-TGGACRAANCCNACNGG-3'). DNA probes were also designed to target *cobT* sequences from the NCBI nr database from the major groups found by 16S rRNA gene profiling in pristine Antarctic surface waters (Alpha- and Gammaproteobacteria and the Bacteroidetes groups; see *Supporting information*). The resulting probes were CobTrossF (5'-TTTGCNGSNRAYCAYGG-3') and CobTrossR (5'-TGYTTTSCNAYCATYTCNCC-3'). These probes were tested with DNA samples from NX19, as well as positive control DNA from *Microcystis aeruginosa* PCC 7806 and negative control DNA from *Escherichia coli* under a range of conditions (variations on 94°C for 4 min, 80°C for 2 min (hot start), 10 to 35 cycles of 94°C for 20 s, 45 to 55°C for 30 s and 72°C for 60 s; and 72°C for 7 min).

Final CbiA primer design and application

DNA probes were designed from *cbiA/cobB* sequences from the NCBI nr database based on the major taxonomic groups found by 16S rRNA gene profiling at site NX19 (Alpha- and Gammaproteobacteria and the Bacteroidetes group; see *Supporting information*). The resulting primers were used: CbiArossF (5'-GGTCCNGAYTWYHTNGA-3') and CbiArossR2 (5'-TAATTCNGGRTANCCNCC-3') with the following DNA amplification protocol: 94°C for 4 min, 80°C for 2 min (hot start), 35 cycles of 94°C for 20 s, 53°C for 30 s and 72°C for 60 s; and 72°C for 7 min. The products were gel-purified, subjected to 15 additional amplification cycles and cloned into competent *E. coli* DH5 alpha using the pGEM-T-Easy vector (Promega).

Sequencing and phylogenetic analyses

Sequencing was performed using the BigDye Terminator V3.1 cycle sequencing kit. Sequencing of polymerase chain reaction products was performed using both amplification primers (described above) with electrophoresis and reading carried out using an Applied Biosystems 3730 DNA Analyzer (Applied Biosystems) at either the Ramaciotti Center (University of New South Wales) or by Northwoods DNA. Sequences were edited in FinchTV. Subsequent analyses of *cbiA/cob* partial sequences were conducted on a 650 bp region. Multiple alignments of the translated partial CbiA/CobB amino acid sequences were performed using Clustal X (Version 1.83) (Thompson *et al.*, 1994). Phylogenetic reconstruction of the newly identified partial CbiA/CobB sequences was done using genetic distance methods based on all identified prokaryotic phyla. Bootstrap resampling with 1000 to 4000 iterations was also performed (Neilan *et al.*, 2002). Reference strains were obtained from the NCBI nr database. Fastgroup II (Yu *et al.*, 2006) was used to estimate the number of different sequences detected. These sequence data have been submitted to the GenBank database under accession numbers HQ326507 to HQ326535.

Protein extraction and digestion

Protein was extracted from filters by incubating them on ice for 15 min in 15 ml of extraction buffer (0.1 M Tris HCl, pH 7.5; 5% glycerol, 1% SDS, 10 mM EDTA) with periodic gentle vortexing, then heating to 92–98°C in a water bath for 15 min, gently vortexing for 1 min, then mixing in hybridization oven at 20°C for 2 h, with gentle vortexing for 30 s every 30 min. The extraction buffer was then collected and centrifuged at 7000 r.p.m. at 10°C for 20 min. The supernatant was collected, filtered with a 5 µm filter needle and concentrated via ultrafiltering in a Vivaspin 65 000 Da molecular weight cut-off PES membrane in a centrifugal concentrator spinning at 7000 r.p.m. at 10°C until volume was reduced to ~1 ml (~3 h). Protein was precipitated overnight at –20°C in 5 volumes of ice cold 50% acetone, 50% methanol with 0.5 mM HCl. Samples were spun at 13 000 rcf at 4°C for 20 min; pellets were aspirated and then dried at room temperature in a speedvac and resuspended in 100 µl of extraction buffer.

Protein was separated into five size fractions using SDS-PAGE and then digested into tryptic peptides according to Kinter and Sherman (Kinter and Sherman, 2000). Briefly, 15 µl of protein sample was mixed with 15 µl SDS PAGE loading buffer (0.1 M Tris, pH 7.5, 25% glycerol, 0.14 M SDS, 3 mM bromphenol blue), heated to 95°C for 8 min and proteins were then applied to a 4–20% Tris HCl pre-cast Ready Gel (Bio-Rad) for separation by SDS PAGE. Approximately 15 µg of total protein per sample was applied to the gel, as determined by Bovine Serum Albumin-calibrated DC protein Assay (Bio-Rad). Gels were stained using Coomassie Blue, destained and imaged. Gel lanes containing each protein sample were cut into 5 slices (Fig. 3F, corresponding to >100 kDa, 55–100 kDa, 35–55 kDa, 20–35 kDa and <20 kDa). The slices were cut into roughly 1 mm³ cubes and washed repeatedly with 50% 25 mM NH₄HCO₃, pH 8 50% acetonitrile until completely destained, then dehydrated with 100% acetonitrile. The proteins were reduced using 10 mM dithiothreitol, then alkylated with 55 mM iodoacetamide. Slices were then dehydrated with 100% acetonitrile, dried and rehydrated on ice with trypsin (Promega modified sequencing grade; 1 µg per gel slice) in 25 mM NH₄HCO₃, pH 8. The digest was conducted for 16 h at 37°C. Peptides were then extracted into 5% formic acid, 50% acetonitrile, 45% water and brought to 20 µl final volume and stored at –80°C until analysis.

Peptide biomarker analysis via SRM mass spectrometry

Biomarker peptides for Group RSB₁₂ were identified for detection by SRM. In SRM, parent ions of the correct mass for the peptides of interest are detected and isolated in the mass spectrometer within a narrow mass window. Those ions are then fragmented and the intensity of specific characteristic ions produced in the fragmentation (product ions) are quantified. To be identified positively as the analyte of interest using SRM mass spectrometry, the compound must be detected with the same parent mass and the relative magnitude of the characteristic product ions detected must be the same as an authentic standard. In this case, stable isotopically labelled versions of the peptides of interest serve as internal authentic reference standards and can be included in

the same liquid chromatographic separation as the peptides from field samples. The combination of high parent mass resolution, the identification of multiple fragment ions in correct relative magnitude, and having the correct retention time relative to an authentic standard should greatly reduce the likelihood for false positive identifications.

Peptides chosen as SRM-based biomarkers were present in most Group RSB₁₂ sequences and not in other CbiA sequences detected in this study or in other protein sequences in the NCBI nr and envr_nr databases (as of 15 July, 2010). They were also in the mass range of 800–2400 Da and lacking methionine or cysteine to facilitate mass spectrometry analysis. Heavy isotope-labelled versions of the biomarker peptides were obtained from Sigma-Aldrich (Stemmann *et al.*, 2001) and handled according to the manufacturer's instructions. Heavy and native sequences of peptides, their parent and product ions monitored, and collision energies are shown in Table 3. Each gel slice (two protein samples, five slices each) was analysed in triplicate. Whole-cell tryptic digests of *C. watsonii*, a marine cyanobacteria that does not contain Group RSB₁₂ peptides based on its genome, was used as a negative control. The cyanobacterial sample was processed similarly to the field samples described above except without SDS-PAGE separation. To prepare samples for mass spectrometry analysis, 150 fmol of each heavy labelled peptide was added to approximately 1 µg of protein and the mixture was brought to 50 µl with 2% acetonitrile and 0.1% formic acid for injection.

Selected reaction monitoring analyses were performed on a Thermo Vantage TSQ triple quadrupole mass spectrometer with a Michrom Advance captive spray ion source. The SRM reactions were validated through tuning the mass spectrometer on the heavy labelled peptides described above, introduced by direct infusion, to select the most abundant transitions for monitoring and to optimize the collision energies (Table 3). Q1 operated in 0.2 FWHM (full width at half maximum) resolution while Q3 operated in 0.7 FWHM resolution. All SRMs were monitored over the full chromatographic separation. Reverse phase chromatographic separation consisted of a peptide Cap Trap in-line with a reversed phase Magic C18 AQ column (0.2 × 50 mm, 3 mm particle size, 200 Å pore size) on a Paradigm MS4 HPLC system at a flow rate of 4 µl min⁻¹ with a gradient from 5% buffer B then to 45% buffer B over 25 min and to 95% buffer B in 10 min where A was 0.1% formic acid (Michrom) in water (Fisher LC/MS Optima) and B was 0.1% formic acid in acetonitrile (Fisher LC/MS Optima).

Acknowledgements

We are indebted to Falcia Goh, Tim Salmon and other members of the Neilan Lab at UNSW for technical assistance and helpful discussions. We thank Tyler Goepfert for assistance in collecting field samples, Abigail Noble for assistance with cobalt analyses and to Dan Rogers and Erin Banning for assistance with data analysis and helpful discussions. Thanks to Vladimir V. Bulygin for technical assistance with peptide analyses and to Andrew Allen for helpful discussions regarding eukaryotic and metagenomic sequences. Special thanks to the Captain, Crew, and Raytheon Marine and Science Technical Staff of the R/V N.B. Palmer and

CORSACS science parties including Chief Scientist Jack DiTullio. This research was supported by NSF grants OPP-0440840, OCE-0327225, OCE-0452883, OPP-0732665, OCE-0752291, The Center for Environmental Bioinorganic Chemistry at Princeton, The Center for Microbial Oceanography: Research and Education (CMORE), the Australian Research Council, and EPA STAR and NSF Graduate Research Fellowships to EMB.

References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403–410.
- Bertrand, E.M., Saito, M.A., Rose, J.M., Riesselman, C.R., Lohan, M.C., and Noble, A.E. *et al.* (2007) Vitamin B₁₂ and iron co-limitation of phytoplankton growth in the Ross Sea. *Limnol Oceanogr* **52**: 1079–1093.
- Bonnet, S., Webb, E.A., Panzeca, C., Karl, D.M., Capone, D.G., and Sanudo-Wilhelmy, S.A. (2010) Vitamin B₁₂ excretion by cultures of the marine cyanobacteria *Crocospaera* and *Synechococcus*. *Limnol Oceanogr* **55**: 1959–1964.
- Bruland, K.W., Rue, E.L., Smith, G.J., and DiTullio, G.R. (2005) Iron, macronutrients and diatom blooms in the Peru Upwelling regime: brown and blue waters of Peru. *Mar Chem* **93**: 81–103.
- Buesseler, K., Ball, L., Andrews, J., Cochran, J.K., Hirschberg, D., and Bacon, M.P. *et al.* (2001) Upper ocean export of particulate organic carbon and biogenic silica in the Southern ocean along 107.1W. *Deep Sea Res Part II Top Stud Oceanogr* **48**: 4275–4297.
- Caron, D.A., Dennett, M.A., Lonsdale, D.J., Moran, D.M., and Shalapyonok, L. (2000) Microzooplankton herbivory in the Ross Sea, Antarctica. *Deep Sea Res Part II Top Stud Oceanogr* **47**: 3249–3272.
- Church, M.J., Hutchins, D.A., and Ducklow, H.W. (2000) Limitation of bacterial growth by dissolved organic matter and iron in the Southern Ocean. *Appl Environ Microbiol* **66**: 455–466.
- Church, M.J., Short, C.M., Jenkins, B.D., Karl, D.M., and Zehr, J.P. (2005) Temporal patterns of nitrogenase gene (*nifH*) expression in the oligotrophic North Pacific Ocean. *Appl Environ Microbiol* **71**: 5362–5370.
- Coale, K.H., Wang, X., Tanner, S.J., and Johnson, K.S. (2003) Phytoplankton growth and biological response to iron and zinc addition in the Ross Sea and Antarctic Circumpolar Current along 170W. *Deep Sea Res Part II Top Stud Oceanogr* **50**: 635–653.
- Croft, M.T., Lawrence, A.D., Raux-Deery, E., Warren, M.J., and Smith, A.G. (2005) Algae acquire vitamin B₁₂ through a symbiotic relationship with bacteria. *Nature* **438**: 90–93.
- Debussche, L., Thibaut, D., Cameron, B., Crouzet, J., and Blanche, F. (1990) Purification and Characterization of Cobyirinic Acid *a,c*-Diamide Synthase from *Pseudomonas denitrificans*. *J Bacteriol* **172**: 6239–6244.
- DiTullio, G., Grebmeier, J.M., Arrigo, K.R., Lizotte, M.P., Robinson, D.H., and Leventer, A., *et al.* (2000) Rapid and early export of Phaeocystis antarctica blooms in the Ross Sea, Antarctica. *Nature* **404**: 595–598.
- Droop, M.R. (2007) Vitamins, phytoplankton and bacteria: symbiosis or scavenging? *J Plankton Res* **29**: 107–113.

- Ducklow, H. (2000) Bacterial Production and Biomass in the Oceans. In *Microbial Ecology of the Oceans*. Kirchman, D. (ed.). New York, USA: Wiley-Liss, pp. 85–120.
- Ducklow, H.W., Dickson, M.-L., Kirchman, D.L., Steward, G., Orchardo, J., Marra, J., and Azam, F. (2000) Constraining bacterial production, conversion efficiency and respiration in the Ross Sea, Antarctica, January–February, 1997. *Deep Sea Res Part II Top Stud Oceanogr* **47**: 3227–3247.
- Ducklow, H.W., Carlson, C.A., Church, M., Kirchman, D.L., Smith, D., and Steward, G. (2001) The seasonal development of the bacterioplankton bloom in the Ross Sea, Antarctica, 1994–1997. *Deep Sea Res Part II Top Stud Oceanogr* **48**: 4199–4221.
- Francis, C.A., Roberts, K.J., Beman, M.J., Santoro, A.E., and Oakley, B.B. (2005) Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *Proc Natl Acad Sci USA* **102**: 14683–14688.
- Fresquet, V., Williams, L., and Raushel, F.M. (2004) Mechanism of Cobyric Acid a,c-Diamide Synthetase from *Salmonella typhimurium* LT2. *Biochem* **43**: 10619–10627.
- Galperin, M.Y., and Grishin, N.V. (2000) The synthetase domains of cobalamin biosynthesis amidotransferases CobB and CobQ belong to a new family of ATP-dependent amidoligases, related to dethiobiotin synthetase. *Proteins* **41**: 238–247.
- Gentile, G., Giuliano, L., D'Auria, G., Smedile, F., Azzaro, M., Domenico, M., and Yakimov, M.M. (2006) Study of bacterial communities in Antarctic coastal waters by a combination of 16S rRNA and 16S rDNA sequencing. *Environ Microbiol* **8**: 2150–2161.
- Gobler, C.J., Norman, C., Panzeca, C., Taylor, G.T., and Sanudo-Wilhelmy, S.A. (2007) Effect of B-vitamins and inorganic nutrients on algal bloom dynamics in a coastal ecosystem. *Aquat Microb Ecol* **49**: 181–194.
- Hall, J.A., and Safi, K. (2001) The impact of in situ Fe fertilization on the microbial food web in the Southern Ocean. *Deep Sea Res Part II Top Stud Oceanogr* **48**: 2591–2613.
- Hansel, C.M., Fendorf, S., Jardine, P.M., and Francis, C.A. (2008) Changes in bacterial and archaeal community structure and functional diversity along a geochemically variable soil profile. *Appl Environ Microbiol* **74**: 1620–1633.
- Karl, D.M. (2002) Nutrient dynamics in the deep blue sea. *Trends Microbiol* **10**: 410–418.
- Kinter, M., and Sherman, M. (2000) *Protein Sequencing and Identification Using Tandem Mass Spectrometry*. New York, USA: John Wiley & Sons.
- Leipe, D.D., Wolf, Y.I., Koonin, E.V., and Aravind, L. (2002) Classification and Evolution of P-loop GTPases and Related ATPases. *J Mol Biol* **317**: 41–72.
- Marchant, H.J., Davidson, A.T., and Wright, S. W. (1987) The distribution and abundance of chroococoid cyanobacteria in the Southern Ocean. *Proc NIPR Symp Polar Biol* **1**: 1–9.
- Martin, J.H., Fitzwater, S.E., and Gordon, R.M. (1990) Iron deficiency limits phytoplankton productivity in Antarctic waters. *Global Biogeochem Cycles* **4**: 5–12.
- Medigue, C., Krin, E., Pascal, G., Barbe, V., Bernsel, A., Bertin, P., et al. (2005) Coping with cold: the genome of the versatile marine Antarctica bacterium *Pseudoalteromonas haloplanktis* TAC125. *Genome Res* **15**: 1325–1335.
- Menzel, D.W., and Spaeth, J.P. (1962) Occurrence of vitamin B₁₂ in the Sargasso Sea. *Limnol Oceanogr* **7**: 151–154.
- Murray, A.E., and Grzymski, J.J. (2007) Diversity and genomics of Antarctic marine micro-organisms. *Phil Trans Royal Soc B* **362**: 2259–2271.
- Muto, A., and Osawa, S. (1987) The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc Natl Acad Sci USA* **84**: 166–169.
- Neilan, B.A., Burns, B.P., Relman, D.A., and Lowe, D.R. (2002) Molecular identification of cyanobacteria associated with stromatolites from distinct geographical locations. *Astrobiology* **2**: 271–280.
- Oliver, J.L., Barber, R.T., Smith, W.O., and Ducklow, H.W. (2004) The heterotrophic bacterial response during the Southern Ocean Iron Experiment (SOFEX). *Limnol Oceanogr* **49**: 2129–2140.
- Pakulski, J.D., Coffin, R.B., Kelley, C.A., Holder, S.L., Downer, R., Aas, P., et al. (1996) Iron stimulation of Antarctic bacteria. *Nature* **383**: 133–134.
- Panzeca, C., Tovar-Sanchez, A., Agusti, S., Reche, I., Duarte, M., Taylor, G.T., and Sanudo-Wilhelmy, S.A. (2006) B vitamins as regulators of phytoplankton dynamics. *EOS* **87**: 593–596.
- Panzeca, C., Beck, A., Leblanc, K., Taylor, G.T., Hutchins, D.A., and Sanudo-Wilhelmy, S.A. (2008) Potential cobalt limitation of vitamin B₁₂ synthesis in the North Atlantic Ocean. *Global Biogeochem Cycles* **22**: 223–230.
- Partensky, F., Hess, W.R., and Vaulot, D. (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbial Mol Biol Rev* **63**: 106–127.
- Porter, K.G., and Feig, Y.G. (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25**: 943–948.
- Raux, E., Schubert, H.L., Roper, J.M., Wilson, K.S., and Warren, M.J. (1999) Vitamin B₁₂: insight into biosynthesis's mount improbable. *Bioorganic Chem* **27**: 100–118.
- Raux, E., Schubert, H.L., and Warren, M.J. (2000) Biosynthesis of cobalamin (vitamin B₁₂): a bacterial conundrum. *Cell Mol Life Sci* **57**: 1880–1893.
- Rodionov, D.A., Vitreschak, A.G., Mironov, A.A., and Gelfand, M.S. (2003) Comparative genomics of the vitamin B₁₂ metabolism and regulation in prokaryotes. *J Biol Chem* **278**: 41148–41159.
- Roth, J.R., Lawrence, J.G., and Bobik, T.A. (1996) Cobalamin (coenzyme B₁₂): synthesis and biological significance. *Annu Rev Microbiol* **50**: 137–181.
- Saito, M.A., and Moffett, J.W. (2001) Complexation of cobalt by natural organic ligands in the Sargasso Sea as determined by a new high-sensitivity electrochemical cobalt speciation method suitable for open ocean work. *Mar Chem* **75**: 49–68.
- Saito, M.A., and Moffett, J.W. (2002) Temporal and spatial variability of cobalt in the Atlantic Ocean. *Geochim Cosmochim Acta* **66**: 1943–1953.
- Saito, M.A., Moffett, J.W., and DiTullio, G. (2004) Cobalt and nickel in the Peru upwelling region: a major flux of cobalt utilized as a micronutrient. *Global Biogeochem Cycles* **18**: GB4030.
- Saito, M.A., Rocap, G., and Moffett, J.W. (2005) Production of cobalt binding ligands in a *Synechococcus* feature at the Costa Rica Upwelling Dome. *Limnol Oceanogr* **50**: 279–290.
- Saito, M.A., Goepfert, T.J., Noble, A.E., Bertrand, E.M.,

- Sedwick, P.N., and DiTullio, G.R. (2010) A seasonal study of dissolved cobalt in the Ross Sea, Antarctica: micronutrient behavior, absence of scavenging, and relationships with Zn, Cd, and P. *Biogeosciences* **7**: 4059–4082.
- Sañudo-Wilhelmy, S.A., Okbamichael, M., Gobler, C.J., and Taylor, G.T. (2006) Regulation of phytoplankton dynamics by vitamin B12. *Geophys Res Lett* **33**: 4604–4608.
- Scanlan, D., and West, N.J. (2002) Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microb Ecol* **40**: 1–12.
- Sedwick, P.N., DiTullio, G.R., and Mackey, D.J. (2000) Iron and manganese in the Ross Sea, Antarctica: seasonal iron limitation in Antarctic shelf waters. *J Geophys Res* **105**: 11321–11336.
- Seshadri, R., Kravitz, S.A., Smarr, L., Gilna, P., and Frazier, M. (2007) CAMERA: a community resource for metagenomics. *PLoS Biol* **5**: 0394–0397.
- Smith, W.O.J., and Nelson, D.M. (1985) Phytoplankton bloom produced by a receding ice edge in the Ross Sea: spatial coherence with the density. *Science* **227**: 163–166.
- Stemmann, O., Zou, H., Gerber, S.A., Gygi, S.P., and Kirschner, M.W. (2001) Dual inhibition of sister chromatid separation at metaphase. *Cell* **107**: 715–726.
- Steward, G.F., Jenkins, B.D., Ward, B.W., and Zehr, J.P. (2005) Development and testing of a DNA macroarray to assess nitrogenase (*nifH*) gene diversity. *Appl Environ Microbiol* **70**: 1455–1465.
- Swift, D. (1981) Vitamin levels in the Gulf of Maine and ecological significance of vitamin B12 there. *J Mar Res* **39**: 375–403.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Wagner, M., Roger, A.J., Flax, J.L., Brusseau, G.A., and Stahl, D.A. (1998) Phylogeny of dissimilatory sulfite reductases supports an early origin of sulfate respiration. *J Bacteriol* **180**: 2975–2982.
- Webster, N.S., Negri, A.P., Munro, M.M., and Battershill, C.N. (2004) Diverse microbial communities inhabit Antarctic sponges. *Environ Microbiol* **6**: 288–300.
- Yu, Y., Breitbart, M., McNairnie, P., and Rohwer, F. (2006) FastGroupII: a web-based bioinformatics platform for analyses of large 16S rDNA libraries. *BMC Bioinformatics* **7**: 57.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Neighbour-joining tree of DNA encoding 16S rRNA gene sequences from 20–100 m depths. The tree is rooted with *Methanococcoides* and *Nitrosopumilus* and shown with bootstrap values > 50% (4000 replicates). Reference sequences shown in grey are the closest cultured representatives to each of the environmental sequences displayed. Environmental sequences are labelled in black, as RossSeaNX19100m2 for example, which denotes the 2nd sequence from 100 m depth at station NX19 of the Ross Sea. Major groups are displayed with % of the library classified within that group. Thirty-eight different operational taxonomic units' (97% similarity) were recovered from 41 clone sequences. Alpha- and Gammaproteobacteria and the Bacteroidetes group dominated the 16S rRNA gene sequences recovered from Station NX 19.

Table S1. Identification of representative 16S rRNA gene clone sequences from Station NX19 in the Ross Sea based on percentage identity to sequences in the NCBI nr database.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.