RESEARCH ARTICLE

Investigation of XoxF methanol dehydrogenases reveals new methylo trophic bacteria in pelagic marine and freshwater ecosystems

Arthi Ramachandran1 and David A. Walsh1,2,3,*

1Department of Biology, Concordia University, 7141 Sherbrooke St West, Montreal, Quebec, H4B 1R6, Canada, 2Canadian Institute for Advanced Research, Integrated Microbial Biodiversity Program, Canada and 3Groupe de Recherche Interuniversitaire en Limnologie (GRIL), Montreal, Quebec Canada

*Corresponding author: 7141 Sherbrooke St West, Room GE330.17, Montreal, Quebec H4B 1R6, Canada. Tel: +1-514-848-2424; Ext: 3477; E-mail: david.walsh@concordia.ca

One sentence summary: Through cultivation-independent analysis, new methylo trophic bacteria are discovered and characterized in lakes and oceans.

Editor: Gary King

ABSTRACT

The diversity and distribution of methylo trophic bacteria have been investigated in the oceans and lakes using the methanol dehydrogenase mxaF gene as a functional marker. However, pelagic marine (OM43) and freshwater (LD28 and PRD01a001B) methylo trophs within the Betaproteobacteria lack mxaF, instead possessing a related xoxF4-encoded methanol dehydrogenase. Here, we developed and employed xoxF4 as a complementary functional gene marker to mxaF for studying methylo trophs in aquatic environment. Using xoxF4, we detected OM43-related and LD28-related methylo trophs in the ocean and freshwaters of North America, respectively, and showed the coexistence of these two lineages in a large estuarine system (St Lawrence Estuary). Gene expression patterns of xoxF4 supported a positive relationship between xoxF4-containing methylo troph activity and spring time productivity, suggesting phytoplankton blooms are a source of methylo trophic substrates. Further investigation of methanol dehydrogenase diversity in pelagic ecosystems using comparative metagenomics provided strong support for a widespread distribution of xoxF4 (as well as several distinct xoxF5) containing methylo trophs in marine and freshwater surface waters. In total, these results demonstrate a geographical distribution of OM43/LD28-related methylo trophs that includes marine and freshwaters and suggest that methylo trop he occurring in the water column is an important component of lake and estuary carbon cycling and biogeochemistry.

Keywords: C1 metabolism; LD28; marine OM43 clade; metagenomics; methanol oxidation; methylo trop he

INTRODUCTION

Methanol is an abundant organic compound in the ocean where it can serve as a carbon and energy source for methylo tropic microorganisms. Marine methanol can originate from terrestrial inputs as well as production during phytoplankton growth and decay (Milne et al. 1995; Heikes 2002). In addition to high oceanic concentrations (50–400 nM), the turnover time of methanol in seawater is short, on the order of one to a few days, suggesting that methanol biogeochemistry is a significant component of the marine carbon cycle (Dixon, Beale and Nightingale 2011; Dixon and Nightingale 2012). To elucidate the role of microorganisms in marine methanol cycling, it is important to identify and quantify the diversity and distribution of methylo tropic bacteria in the ocean. To this end, Mcdonald and Murrell (1997) first proposed the use of the mxaF gene, which
encodes the alpha subunit of methanol dehydrogenase (MDH), as a functional gene marker for methylotrophs. MDH is an enzyme containing a pyrroloquinoline quinone (PQQ) cofactor that oxidizes methanol to formaldehyde, and over the years the mxaF gene has been used in numerous cultivation-independent assessments of marine methylotroph diversity (McDonald and Murrell 1997; Neufeld et al. 2008a; Lau et al. 2013).

Although mxaF as a marker has proved useful, an important limitation is that marine methylotrophs that lack the mxaF gene have recently been reported. Cultivation of marine methylotrophs of the OM43 clade of Betaproteobacteria and subsequent genome analysis showed that these methylotrophs lack mxaF, yet possess a homologous xoxF4 gene that is implicated in methanol oxidation (Giovannoni et al. 2008; Chistoserdova 2011b). In fact, previous phylogenetic analyses of MxaF and XoxF proteins revealed that this MDH protein family comprises at least six different clades (MxaF and XoxF1–5) and that methylotrophs may possess multiple MDH homologs (Chistoserdova 2011b; Keltjens et al. 2014). However, representative genomes of the OM43 clade possess only a single xoxF4 gene.

Based on 16S rRNA surveys, the OM43 clade is more abundant and common than other previously described proteobacterial methylotrophs (Sekar et al. 2004; Morris, Longnecker and Giovannoni 2006; Song, Oh and Cho 2009; Taubert et al. 2015, in press). Environmental protocentric studies have shown that the XoxF4 protein is commonly expressed in coastal surface waters, suggesting the OM43 clade may be responsible for a considerable amount of methanol turnover in the coastal ocean (Sowell et al. 2011; Georges et al. 2014). Based on these recent observations, it is therefore likely that the current approaches that rely on the mxaF gene marker do not provide an accurate account of the abundance and diversity of marine methylotrophic bacteria.

Compared to oceans, less is known about methanol oxidation in the surface waters of lakes and rivers, although a large body of work exists on benthic methylotrophy in sediments (Kalyuzhnaya et al. 2012; Chistoserdova and Lidstrom 2013). Given that terrestrial plants and phytoplankton are sources of methanol and methylated compounds in the ocean (Milne et al. 1995; Heikes 2002; Millet et al. 2008), methylotrophy is likely important in the surface waters of lakes as well. Interestingly, 16S rRNA gene sequences affiliated with OM43 are common in lakes of Europe and North America where they are referred to as the LD28 tribe (Zwart et al. 1998, 2003; Salcher et al. 2008; Jezebera et al. 2012). LD28 and OM43 are both members of the broader betIV clade (Newton et al. 2011), which also contains Methylophilus and Methylostera isolates from freshwater sediments and soils (Chistoserdova 2011a). These isolates can use a range of one carbon compounds (Chistoserdova 2011a) and some may also be able to subsist on phenol and humic matter (Hutalte-Schmelzer et al. 2010). Very recently, isolates from two planktonic freshwater lineages (LD28 and closely affiliated PRD01a001B) were cultured from Lake Zurich, Switzerland (Salcher et al. 2015, in press). Subsequent genome analysis confirmed a methylotrophic lifestyle and a reduced set of genes for one carbon (C1) metabolism. Similar to OM43, these freshwater lineages lack mxaF, but possess a single xoxF4 homolog. As such, we hypothesized that xoxF4 genes present in lakes and rivers can also serve as a functional gene marker for freshwater LD28/PRD01a001B methylotrophs.

The objective of this study was to first develop the xoxF4 gene as a complementary functional gene marker to the canonical mxaF gene for studying methylotrophy diversity in aquatic environments. We then investigated the distribution, diversity, and xoxF4 expression levels of methylotrophic bacteria across marine, estuarine, and freshwater environments using a combination of PCr-based analyses of xoxF4 genes and comparative metagenomics of PQQ-dependent methanol dehydrogenases present in publically available marine and freshwater metagenomes.

MATERIALS AND METHODS

Sampling and environmental DNA and RNA extraction

Microbial biomass for DNA and RNA extraction was collected from the surface water (5 m depth) at the compass buoy station in Bedford Basin, Nova Scotia, and along a transect of the St Lawrence Estuary, Quebec (Fig. 1). Sampling of the St Lawrence Estuary occurred during 16–20 May 2011. It should be noted that no biological replicates of DNA or RNA were collected. Seawater was prefiltered (2 L from Bedford Basin and 2.5 L from the St Lawrence Estuary) through a Whatman GF/D filter (2.7 μm cut-off) and cells were collected on a 0.22 μm Sterivex filter. After filtration, 1.8 mL of sucrose-based lysis buffer was added and filters were stored at −80°C. DNA was extracted following the protocol described in Zaitkova et al. (2010). RNA was extracted using the mirVana miRNA Isolation kit (Life Technologies, Burlington, Ontario, Canada) following the modified methodology in Stewart et al. (2012).

Temperature and salinity were measured by CTD. Bacterial production was measured using the 14C-helium incorporation method detailed in Smith and Azam (1992). Bacterial abundance was determined through flow cytometry in which the cells were stained with SYBR Gold as detailed in Li and Dickie (2001). Chlorophyll a measurements were determined in vitro by fluorescence of 90% acetone extracts of plankton as detailed in Li and Dickie (2001).

xoxF4 PCR primer design and amplification

Multiple sequence alignments of PQQ-dependent dehydrogenase genes (xoxF1–5 and mxaF) from whole genomes were generated using MUSCLE (in MEGAS). Primer sets were designed in order to amplify the xoxF4 fragment specifically by maximizing primer matching to marine xoxF4 genes and minimizing primer matching to genes from the other five xoxF/mxaF clades. The multiple sequence alignments were created using the nucleotide sequences of genes encoding PQQ-dependent dehydrogenases acquired from the NCBI database. The primers used for the ~500 bp fragment were 974F (5′-AACCCWTCHTNTGGAAYCC-3′) and 1441R (5′-GRCRMCCACAWGGYTGACC-3′), which were subsequently used for cloning and sequencing. The primers for the ~230 bp fragments were 1004F (5′-CCWGGYGAAYAARTGCT-3′) and 1232R (5′-GGCACATTNYRAAKGGRTG-3′), which were subsequently used for qPCR and RT-qPCR assays (Supplementary Fig. S1). Genomic DNA from Methylotenera mobilis strain #13 (Chistoserdova 2011a) was used as a PCR positive control as this strain possesses both the mxaF and the xoxF4 homologs.

Template genomic DNA from environmental samples was used in a 25 μL reactions containing 16 μL MilliQ, 5 μL 1X Reaction Buffer, 1.25 μL F primer (0.5 μM), 1.25 μL R primer (0.5 μM), 0.5 μL dNTPs (0.2 μM), 0.5 μL Phire Polymerase, 1 μL DNA template (~10 ng). PCR conditions were as follows: 98°C for 30 s for the initial denaturing step; 30 cycles of 98°C for 5 s, 56.5°C (60°C for mxaF) for 5 s, 72°C for 10 s; and 72°C for 1 min for the final
elongation step. PCR products were visualized using a 1% (w/v) agarose gel via electrophoresis.

**Cloning and sequencing of xoxF4 fragment**

The larger fragment (~500 bp) of xoxF4 was PCR-amplified and gel-purified using the QIAGEN QIAquick Gel Extraction Kit. The samples used for the construction of the clone libraries were from January 2011 and May 2011 from Bedford Basin, and Station B and Station 22 from the St Lawrence Estuary. The purified PCR product was cloned into a pJET vector using the CloneJET PCR Cloning Kit. The transformants were grown on LB + Ampicillin plates and the colonies, 24 for each sample, were screened for the insert by colony PCR using the pJET primers. Restriction fragment length polymorphism (RFLP) analysis was performed using the following reaction mix: 7.75 μL H2O, 2 μL NE buffer 4, 0.25 μL HhaI, 10 μL Colony PCR product. The 16 clones chosen for sequencing were those that showed variation in fragment size after the RFLP. Plasmid DNA was isolated using the GeneJet Plasmid Miniprep Kit (Thermo Scientific, Ottawa, Ontario, Canada) and sent for paired end Sanger sequencing (OPERON, Louisville, KY, USA). Phylogenetic analysis was performed using maximum likelihood and a GTR + Gamma distribution (four categories) model of evolution.

**Quantification of xoxF4 gene and mRNA abundances**

Plasmids harboring the xoxF4 fragment were linearized overnight at 37°C using the Eco31I restriction enzyme. The linearized plasmid was gel purified and used as the qPCR and RT-qPCR standards. The absolute quantification was measured through serial dilutions of 10^7 to 10^3 gene copies per microlitre. Template genomic DNA from Bedford Basin and St Lawrence Estuary samples were used in 20 μL reactions containing 14.6 μL MilliQ, 4 μL 5X EvaGreen qPCR mix, 0.4 μL F primer (0.1 μM), 0.4 μL B primer (0.1 μM), 1 μL DNA template. The thermal profile used was 15 min polymerase activation at 95°C followed by the PCR cycling stage with 40 cycles (95°C for 45 s, 55°C for 30 s, 72°C for 1 min) and ending with a melting curve (95°C for 15 s, 55°C for 15 s, 95°C for 15 s). The dilution factor used for the unknowns was between 10^0 and 10^-3 and the reactions were performed in triplicate. In order to calculate the copy number of the xoxF4 gene in 1 mL of seawater, the copy number given by the qPCR was divided by the amount of seawater (mL) filtered for each sample and corrected for the dilution factor.

The RNA from the Bedford Basin and the St Lawrence Estuary was used as the template to synthesize cDNA using the M-MLV Reverse Transcriptase kit (Invitrogen) using the reverse primer 1441R. The results obtained from the samples (Bedford Basin and St Lawrence Estuary) were compared to the standard curve and a quantity of copy number was given. Calculation of mRNA transcript levels in the RT-qPCR analysis was similar to that of gene copy number. The efficiency of the qPCR was ≥ 81% and the R^2 value was ≥ 93% for all runs.

**Metagenomic analysis of PQQ-dependent methanol dehydrogenases**

XoxF4 from the Betaproteobacterial strain HTCC2181 was used as the query protein sequence to search for homologs in metagenomic datasets from 35 marine and two freshwater ecosystems from IMG/M using BLASTp and a 10^-5 E-value cutoff (Supplementary Table S1). Retrieved protein sequences were
clumped at 99% identity using the program cd-hit (Fu et al. 2012) to reduce redundancy, resulting in 808 clusters for which a single sequence was selected as a representative. Sequences greater than 300 amino acids in length were added to a multiple sequence alignment, containing PQQ-dependent dehydrogenase reference sequences including sequences from Lau et al. (2013) and Keltjens et al. (2014). The preliminary tree included metagenomic sequences that were closely related to other PQQ-dependent dehydrogenases (e.g. alcohol and glucose dehydrogenases) and these were trimmed. After clustering and trimming, 44 protein sequences remained for the final phylogenetic analysis using the MEGA6 software. The following parameters were used: maximum-likelihood method, JTT substitution model, gamma distribution model for the rate variation parameters were used: maximum-likelihood method, JTT substitution method, gamma distribution model for the rate variation

### RESULTS

**xoxF4 PCR primer design and specificity**

Previous phylogenetic analysis of MxaF and XoxF proteins revealed that these MDH subunits are members of six different clades (MxaF and XoxF1–5) (Chistoserdova 2011b; Keltjens et al. 2014). While the genomes of certain bacteria may encode multiple xoxF homologs (Keltjens et al. 2014), the publicly available marine OM43 genomes (strains HTCC2181 and HIMB624/KB13) and freshwater LD28 (’Candidatus Methylophilus planktonicus’) and PRD01a001B (’Candidatus Methylophilus turicensis’) genomes only contain a single homolog encoding an XoxF protein. Therefore, we designed PCR primer sets (Supplementary Fig. S1) that would broadly and specifically target the xoxF4 gene in order to maximize the detectable diversity of xoxF4-containing bacteria in the environment. One primer combination (974F/1441R) targeted a ~500 bp fragment of the xoxF4 gene, while a second primer combination (1004F/1232R), targeting an internal ~230 bp fragment, was designed for qPCR.

**xoxF4 diversity in aquatic environments**

We surveyed the methylophilic community in seasonal surface waters of a coastal inlet (Bedford Basin, Nova Scotia), as well as along a 200 km estuarine salinity gradient (St Lawrence Estuary, Quebec) (Fig. 1, Table 1). Using the 974F/1441R primers, we detected xoxF4 genes in all samples, while no mxaF amplification was detected using the updated primers reported in Lau et al. (2013). Cloning and sequencing of the xoxF4 fragments followed by phylogenetic analysis verified the specificity of the 974F/1441R primer combination and identified three distinct and well-supported xoxF4 subclades (Fig. 2). We detected xoxF4 sequences nearly identical to strain HTCC2181 in both Bedford Basin (BB) and the marine end of the St Lawrence Estuary (SLE). The second subclade (xoxF4–2) was composed solely of sequences recovered from BB. The third clade was composed of sequences recovered from the SLE, including sequences recovered from a freshwater location (xoxF4–3), as well as the xoxF4 gene from the LD28 isolate ‘Candidatus Methylophilus planktonicus’. These results expand the known diversity and distribution of xoxF4-harborine lineages in both marine and freshwater habitats of North America.

**Abundance and expression of xoxF4 genes**

Quantification of xoxF4 gene abundance and mRNA transcripts revealed spatiotemporal variation in methylophilic abundance and expression levels in BB and the SLE (Fig. 3). In BB, we compared the xoxF4 community during the winter (December and January) and during the spring phytoplankton bloom (April and May). Over this time interval, xoxF4 gene abundance increased twofold (from an average of 1747 mL⁻¹ to 3499 mL⁻¹), while mRNA expression levels increased 12-fold (from 419 mL⁻¹ to 4964 mL⁻¹). These results demonstrated that xoxF4-harborine bacteria exhibit higher abundance and xoxF4 expression levels in association with productive spring conditions compared to the winter.

In the SLE, the xoxF4 genes were present and expressed across a salinity gradient ranging from freshwater (Station B) to brackish-marine waters (Station 21). However, unlike in BB, we did not observe a positive correlation between abundance and expression. While xoxF4 abundance was maximal at the freshwater end of the SLE (Station B, 9931 mL⁻¹), xoxF4 expression was at its lowest (198 mL⁻¹), suggesting the presence of an abundant, yet inactive freshwater assemblage of methylophilics. In contrast, xoxF4 gene expression was elevated within the brackish estuary and exhibited a maximum at the most productive location near the marine end of the estuary (Station 22, 7345 copies mL⁻¹).

### Table 1. Environmental data from sample locations in Bedford Basin and Saint Lawrence Estuary.

<table>
<thead>
<tr>
<th>Location</th>
<th>Temperature (°C)</th>
<th>Salinity (PSU)</th>
<th>Chlorophyll (mg m⁻³)</th>
<th>Bacterial abundance (×10⁶ mL⁻¹)</th>
<th>Bacterial production (μg C L⁻¹ day⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bedford Basin</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15 December 2010</td>
<td>7.9</td>
<td>29.6</td>
<td>0.97</td>
<td>8.3</td>
<td>n.d.</td>
</tr>
<tr>
<td>11 January 2011</td>
<td>5.1</td>
<td>29.9</td>
<td>0.84</td>
<td>2.8</td>
<td>n.d.</td>
</tr>
<tr>
<td>20 April 2011</td>
<td>4.5</td>
<td>29.9</td>
<td>7.64</td>
<td>5.7</td>
<td>n.d.</td>
</tr>
<tr>
<td>25 May 2011</td>
<td>7.9</td>
<td>28.4</td>
<td>12.69</td>
<td>8.4</td>
<td>n.d.</td>
</tr>
<tr>
<td><strong>Saint Lawrence Estuary</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Station B</td>
<td>10.6</td>
<td>0.08</td>
<td>n.d.</td>
<td>6.5</td>
<td>21.64</td>
</tr>
<tr>
<td>Station SI</td>
<td>4.3</td>
<td>18.5</td>
<td>n.d.</td>
<td>2.8</td>
<td>2.94</td>
</tr>
<tr>
<td>Station 25</td>
<td>5.5</td>
<td>21.2</td>
<td>n.d.</td>
<td>4.2</td>
<td>15.90</td>
</tr>
<tr>
<td>Station 23</td>
<td>4.7</td>
<td>24.3</td>
<td>n.d.</td>
<td>5.7</td>
<td>42.41</td>
</tr>
<tr>
<td>Station 22</td>
<td>5.4</td>
<td>22.2</td>
<td>n.d.</td>
<td>6.9</td>
<td>59.12</td>
</tr>
<tr>
<td>Station 21</td>
<td>6.0</td>
<td>27.1</td>
<td>n.d.</td>
<td>3.7</td>
<td>26.03</td>
</tr>
</tbody>
</table>

Downloaded from: https://academic.oup.com/femsec/article-abstract/91/10/fiv105/2467292 by guest on 01 March 2019
Comparative metagenomics of \textit{mxaF} and \textit{xoxF} diversity and biogeography

Based on our observation of \textit{xoxF4} genes in the freshwater end of the SLE and the recent report of freshwater methylotrophs by Salcher et al. (2015, in press), we were motivated to investigate the distribution of PQQ-dependent methanol dehydrogenase across a diversity of marine and freshwater habitats using metagenomics. To do so, we searched for XoxF/MxaF in publicly available metagenomic datasets, including those from surface waters of the oceans and freshwater lakes in North America (Supplementary Table S1). After clustering and trimming, we identified 44 representative sequences. Phylogenetic analysis placed the metagenome-derived protein sequences into five separate clades, which were previously reported in Keltjens et al. (2014) (XoxF1, 3, 4, 5 and MxaF) (Fig. 4). To visualize the geographical distribution of methylotrophs, we then mapped the distribution of the different clades to the metagenome sample locations (Fig. 5).

Based on our metagenomic analysis, the detection of MxaF was restricted to lakes, while XoxF4 was more broadly distributed across lake and ocean locations. XoxF4 proteins affiliated with OM43 were common in coastal habitats of the North Pacific and North Atlantic Oceans, yet not detectable offshore (Stations 12, 16 and 26 of Line P in the north-eastern Pacific and a single station in the Indian Ocean) (Supplementary Table S1). In support of our observation of freshwater \textit{xoxF4}-containing methylotrophs in the St Lawrence River, we also identified closely related homologs in the epilimnion of two freshwater lakes (Trout Bog and Lake Mendota). Phylogenetic analyses placed these freshwater XoxF4 homologs into two distinct groups. One clade contained cultured isolates (e.g. \textit{Methylphilus} and \textit{Methylotenera}) from freshwater sediments (Hernandez et al. 2015), while the second clade (XoxF4–3) was represented by environmental sequences recovered from the water column and the xoxf4 homolog from ‘\textit{Candidatus Methylphilus planktonicus}’ (Salcher et al. 2015, in press). These observations support the presence of phylogenetically distinct benthic and pelagic lineages of methylotrophic \textit{Betaproteobacteria} in lakes.

Metagenomics also revealed the presence and broad distribution of the XoxF5 clade in freshwater and marine environments. In the ocean, the XoxF5 showed a contrasting latitudinal distribution compared to XoxF4: XoxF4 dominated in the northern hemisphere, while XoxF5 dominated environments near the equator. Moreover, there were 3 distinct XoxF5 subclades, including two that were only identified in the ocean (XoxF5–1, XoxF5–2), and a third that was restricted to lakes (XoxF5–3). While the freshwater clade was closely affiliated with \textit{Methylomicrobium alcaliphilum}, a halophilic methanotrophic bacterium containing the \textit{mxaF} gene (Vuilleumier et al. 2012) the marine subclades represented uncultivated and poorly characterized taxa that may represent novel one carbon compound oxidizers in the sea.

DISCUSSION

OM43 methylotroph abundance and activity in coastal habitats

Although methylotrophy has been studied in the oceans for decades, a combination of cultivation, genomic, and proteomic studies recently demonstrated that the \textit{Betaproteobacteria} OM43 clade represent a previously overlooked component of the
methylotrophic marine community (Morris, Longnecker and Giovannoni 2006; Giovannoni et al. 2008; Sowell et al. 2011; Halsey, Carter and Giovannoni 2012; Georges et al. 2014; Taubert et al. 2015, in press). Using PCR-based assays specific to xoxF4, we detected the presence and expression of xoxF4 genes across a set of samples from the coastal north-west Atlantic Ocean and along the St Lawrence Estuary salinity gradient. The universal detection of the xoxF4 gene and an absence of detectable mxaF demonstrate that OM43 are numerically dominant over mxaF-harboring methylotrophs in the ocean. A recent study of xoxF diversity in the western English Channel also identified xoxF4 genes closely related to the OM43 clade (Taubert et al. 2015, in press). Moreover, our comparative metagenomics analysis showed a common presence of xoxF4 in coastal habitats, and a lack of detectable mxaF, supporting a previous report by Chistoserdova (2011b). In the Atlantic, xoxF4 gene abundance ranged from 1747 to 3499 copies mL$^{-1}$ suggesting that about 1% of the bacterioplankton contain a xoxF4 homolog. Although it is possible that some OM43 bacteria may possess multiple copies of xoxF4, if we assume xoxF4 is a single copy gene in OM43, then this value agrees well with previous studies showing that OM43 constitute ~2% of bacterial cells in coastal microbial communities (Rappé, Vergin and Giovannoni 2000; Sekar et al. 2004; Gallow et al. 2008; Song, Oh and Cho 2009). We observed highest xoxF4 gene expression during the spring phytoplankton bloom. This observation is in agreement with previous work showing a high OM43 abundance associated with diatoms (Morris, Longnecker and Giovannoni 2006), suggesting that diatom-derived methanol and methylated substances such as dimethylsulfoniopropionate (DMSP) and trimethylamine-N-oxide (TMAO) may serve as methylotrophic growth substrates.

The widespread distribution and expression of xoxF4 suggests that OM43 could be playing a role in marine methylotrophy. However, these findings are at odds with earlier studies that implicate Gammaproteobacteria with marine methanol metabolism (Neufeld et al. 2007; Neufeld et al. 2008b). In a series of elegant studies, Neufeld et al. (2007, 2008b, 2008a) used DNA-SIP with [13C]methanol to link methylotrophy with Methylophaga and other Gammaproteobacteria, but not OM43, irrespective of season. One possible explanation is that OM43 are adapted to very low methanol concentrations, and it has been suggested that XoxF-type MDHs oxidize methanol with higher rates and affinities than MxaF-MDHs (Keltjens et al. 2014). However, even in incubations at near in situ methanol concentration (1 μM), Methylophaga are detected among the predominant methylotrophs (Neufeld et al. 2008a). Perhaps OM43 methylotrophs were overlooked in these studies since low resolution 16S rRNA DGGE analysis and/or the mxaF gene marker were used to assess the identity of methylotrophs. On the other hand, it may be that OM43 does not readily assimilate methanol under the incubation conditions as it may be adapted to take advantage of multiple growth substrates simultaneously when available in the environment. In fact, although OM43 (strain HTCC2181) can grow on methanol as a sole carbon and energy source, this strain exhibits synergistic growth when methanol is amended with additional methylated compounds including DMSP and TMAO, which are compatible osmolytes produced by phytoplankton (Halsey, Carter and Giovannoni 2012). In any case, the application of xoxF4 as a functional gene marker for OM43 methylotrophy should be valuable in further deciphering the ecological niches and metabolic preferences of methylotrophs in the ocean.
Figure 4. Phylogenetic analysis of PQQ-dependent methanol dehydrogenase MsxF?XoxF proteins. Protein sequences generated from our study are in red, while those from lake and ocean metagenomes are in blue. Some methylotrophs have multiple homologs of XoxF that can branch into different clades. The following parameters were used: maximum-likelihood method, 100 bootstrap iterations, JTT substitution model, gamma distribution model for the rate variation with four discrete gamma parameters.
Expanding the diversity of putative marine methylotrophs

In addition to the XoxF4 family, our metagenomic survey revealed that the XoxF5 family is common in the ocean. Although XoxF5 function remains undescribed, it is a large family of proteins implicated in one carbon metabolism and found in methylotrophs of Alpha-, Beta- and Gammaproteobacteria (Keltjens et al. 2014). For the most part, detection of XoxF5 was restricted to locations near the equator, while XoxF4 was more prevalent at northern latitudes. Indeed, Salcher et al. (2015, in press) reported that xoxF4-containing freshwater strains from Lake Zurich, Switzerland were psychrophilic, growing more rapidly at 6 °C compared to 22 °C. Hence, the biogeographical divide we observed in the oceans may be driven by environmental conditions. However, further work is required to link these xoxF5 genes to particular phylogenetic lineages and assess their role in marine methylotrophy.

Rare-earth elements and their relationship to methylotrophs

Rare-earth elements (REEs) are used as the cofactor for the XoxF methanol dehydrogenase. These elements are not as rare as their name implies as they are found in every type of soil, sand and sediment (Keltjens et al. 2014). Previous studies indicate that xoxF gene expression is induced by REEs found in the lanthanide group such as cerium (III) and lanthanum (III) (Keltjens et al. 2014). A study conducted by Hara et al. (2009) showed that during phytoplankton growth, the concentration of dissolved REEs decreased in the surface waters suggesting that REEs can be concentrated by phytoplankton. Phytoplankton decay may therefore not only serve as a source of methylotrophic growth substrates, but also as a source of REE cofactors; however, further experimentation is required to test this hypothesis.

Freshwater and estuarine pelagic methanol-oxidizing bacteria

Methylotrophs have been extensively studied in freshwater sediments (Chistoserdova 2011a), yet much less is known about methylotroph distribution and activities in the surface waters of lakes and rivers. Potential sources of methanol and methylated compounds in freshwater include production in anoxic bottom waters and sediments, degradation of methylated sugars, direct production by phytoplankton and transport from terrestrial sources. Using our xoxF4 primer set we identified a group of xoxF4 sequences (xoxF4–3) in the St Lawrence River, and in metagenomes from the epilimnion of Lake Mendota and Trout Bog in North America. These results expand on a recent study by Salcher et al. (2015, in press), who isolated previously uncultivated members of two freshwater lineages (LD28 and PRD01a001) affiliated with the marine OM43 clade. The xoxF4 sequences from our study were members of the LD28 group, and we were unable to identify xoxF4 sequences from the PRD01a001 group. Although primer bias cannot be completely ruled out, in silico analysis would suggest that this is unlikely, as our xoxF4 primers will amplify xoxF4 sequences from both groups. This observation is in line with the findings of Salcher et al. (2015, in press) who demonstrated that LD28 were generally more abundant...
than PRD01a001 in Lake Zurich and this may be a general pattern that extends to freshwaters of North America.

In addition to the identification of LD28-derived xoxF4 in freshwaters, we also identified a mixture of LD28- and OM43-derived xoxF4 genes in brackish waters of the SLE. In a recent study by Taubert et al. (2015, in press), a diversity of xoxF4 genes were identified in a salt marsh demonstrating their common presence in brackish environments. In the SLE, a peak in xoxF4 gene expression was also concomitant with the most productive location (S22) in the estuary. Hence, as shown for marine (Milne et al. 1995; Heikes 2002; Sinha et al. 2007; Halsey, Carter and Giovannoni 2012) and freshwater systems (Millet et al. 2008; Salcher et al. 2015, in press), it appears that methylotrophs are associated with phytoplankton abundance in estuaries as well.

Now that it is possible to differentiate between distinct populations of xoxF4 containing methylotrophs, it will be interesting to assess the activity of OM43 and LD28 methylotrophs along salinity gradients. Moreover, the phylogenetic relationships between the marine and freshwater lineages show that this group of methylotrophs likely has a freshwater ancestry and that the OM43 arose due to colonization of the marine environment. This is supported by the observation that marine OM43 genomes have acquired ‘marine genes’ such as Na+–pumping respiratory proteins by lateral gene transfer from marine Alphaproteobacteria (Walsh, Lafontaine and Grossart 2013) during their evolution. Hence the OM43/LD28 methylotrophs may serve as a useful evolutionary model to investigate habitat transitions, ecological specialization and speciation in the microbial world.

SUPPLEMENTARY DATA
Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS
We wish to thank the captains and crews of the R/V Coriolis II for providing essential assistance during field operations in the SLE and Bill Li, Kevin Pauley and Jeff Anning for logistical support in collecting samples from Bedford Basin. We thank Paul del Giorgio for providing bacterial abundance data and Roxane Manargent for providing bacterial production data. We are also grateful to Ludmila Chistoserdova for providing the M. mobilis culture. We thank Ludmila Chistoserdova and Hans-Peter Grossart for comments on an earlier version of the manuscript and three anonymous reviewers for their helpful input on the submitted manuscript.

FUNDING
This work was supported by National Science and Engineering Research Council of Canada (402214–2011) and Canada Research Chair Program (950–221184) research grants. A.R. was supported by the Concordia Institute for Water, Energy, and Sustainable Systems.

Conflict of interest. None declared.

REFERENCES


Morris, RM, Longnecker, K, Giovannoni, SJ. Pirellula and OM43 are among the dominant lineages identified in an Oregon coast diatom bloom. Environ Microbiol 2006;8:1361–70.


