Molecular evidence for novel mercury methylating microorganisms in sulfate-impacted lakes

Daniel S. Jones^{1,2,6} · Gabriel M. Walker² · Nathan W. Johnson³ · Carl P. J. Mitchell⁴ · Jill K. Coleman Wasik⁵ · Jake V. Bailey²

Received: 31 August 2018 / Revised: 2 February 2019 / Accepted: 8 February 2019 / Published online: 26 February 2019
© International Society for Microbial Ecology 2019

Abstract

Methylmercury (MeHg) is a bioaccumulative neurotoxin that is produced by certain anaerobic microorganisms, but the abundance and importance of different methylating populations in the environment is not well understood. We combined mercury geochemistry, *hgcA* gene cloning, rRNA methods, and metagenomics to compare microbial communities associated with MeHg production in two sulfate-impacted lakes on Minnesota's Mesabi Iron Range. The two lakes represent regional endmembers among sulfate-impacted sites in terms of their dissolved sulfide concentrations and MeHg production potential. rRNA amplicon sequencing indicates that sediments and anoxic bottom waters from both lakes contained diverse communities with multiple clades of sulfate reducing *Deltaproteobacteria* and *Clostridia*. In *hgcA* gene clone libraries, however, *hgcA* sequences were from taxa associated with methanogenesis and iron reduction in addition to sulfate reduction, and the most abundant clones were from unknown groups. We therefore applied metagenomics to identify the unknown populations in the lakes with the capability to methylate mercury, and reconstructed 27 genomic bins with *hgcA*. Some of the most abundant potential methylating populations were from phyla that are not typically associated with MeHg production, including a relative of the *Aminicenantes* (formerly candidate phylum OP8) and members of the *Kiritimatiellaeota* (PVC superphylum) and *Spirochaetes* that, together, were more than 50% of the potential methylators in some samples. These populations do not have genes for sulfate reduction, and likely degrade organic compounds by fermentation or other anaerobic processes. Our results indicate that previously unrecognized populations with *hgcAB* are abundant and may be important for MeHg production in some freshwater ecosystems.

Supplementary information The online version of this article (<https://doi.org/10.1038/s41396-019-0376-1>) contains supplementary material, which is available to authorized users.

✉ Daniel S. Jones
daniel.s.jones@nmt.edu

¹ BioTechnology Institute, University of Minnesota, St. Paul, MN, USA

² Department of Earth Sciences, University of Minnesota, Minneapolis, MN, USA

³ Department of Civil Engineering, University of Minnesota Duluth, Duluth, MN, USA

⁴ Department of Physical and Environmental Sciences, University of Toronto – Scarborough, Toronto, ON, Canada

⁵ Department of Plant and Earth Science, University of Wisconsin River Falls, River Falls, WI, USA

⁶ Present address: Dept. of Earth and Environmental Science, New Mexico Institute of Mining and Technology, Socorro, NM, USA

Introduction

Mercury (Hg) is pervasive in aquatic ecosystems because of past and present anthropogenic emissions and the potential for global atmospheric transport [1–3]. Severe health effects are associated with its methylated form, methylmercury (CH_3Hg^+ , hereafter MeHg), which is a neurotoxin that bioaccumulates in aquatic food webs. MeHg is produced from inorganic mercury by certain anaerobic microorganisms. Sulfate-reducing microorganisms are thought to be the primary mercury methylators in most environments [4, 5], and sulfate introduction to otherwise low-sulfate ecosystems can stimulate MeHg production [6–10]. But more recently, MeHg production has also been associated with iron-reducing microorganisms [11–13], methanogens [14], and other anaerobes [15].

The recent discovery of two genes associated with mercury methylation, *hgcAB*, has dramatically expanded the known diversity of potential methylating organisms [16].

hgcAB are present in the genomes of diverse sulfate- and iron-reducing members of the *Deltaproteobacteria* and *Clostridia* (*Firmicutes*), methanogens in the *Methanomicrobia*, and acetogenic, fermentative, and syntrophic members of the *Firmicutes* [15, 16]. *hgcAB* have also been discovered in members of other microbial phyla, including the *Chloroflexi*, *Chrysiogenetes*, *Nitrospina*, *Spirochaetes* and candidate phyla ACD79, and *Atribacteria* (formerly OP9) [17, 18], demonstrating that the ability to methylate mercury is more widespread across the tree of life than previously believed. *hgcAB* is a valuable marker for the detection of methylating populations in environmental samples (e.g., [19]). *hgcA* or *hgcAB* cloning has identified potential methylating organisms in paddy soils [20, 21], freshwater marshes [22, 23], and other freshwater sediments [24, 25]. Quantitative PCR of *hgcAB* has also been applied to measure potential methylators in coal-ash amended sediments [26] and in waters contaminated by chlor-alkali plant effluent [27, 28]. These discoveries of partial *hgcA* or *hgcAB* sequences have identified new clades that are not closely related to known *hgcAB*-containing organisms and indicate that additional methylator diversity remains to be discovered.

The net rate of MeHg production in anoxic environments is controlled by complex geochemical and ecological dynamics that affect the activity of methylating populations, the bioavailability of inorganic mercury to those populations, and the activity of MeHg demethylators. The availability of electron acceptors and donors [7, 8, 11, 13, 29] exerts a primary control on net MeHg production, while dissolved sulfide and organic matter impact mercury speciation and rates of uptake to methylating organisms [30, 31]. Microbial community ecology also matters. While MeHg production is primarily associated with the *Deltaproteobacteria*, *Clostridia*, and *Methanomicrobia*, the ability to methylate mercury is seemingly randomly distributed within those clades. For example, only specific sulfate reducers in the *Deltaproteobacteria* and *Clostridia* are capable of methylating mercury, and the same genus can include both methylating and non-methylating strains [16, 18, 32]. Likewise, only certain methanogens and iron-reducers can methylate mercury [11, 12, 14, 15], and different organisms vary widely in the rate and extent to which they can methylate [15, 33]. The recent identification of potential methylators in other microbial phyla (e.g., refs. [18, 20, 22–24]) further complicates the picture by showing that taxonomically and metabolically diverse organisms are likely capable of MeHg production. Understanding the distribution of methylating populations in the context of the geochemical constraints known to affect MeHg production will improve our ability to predict MeHg generation in anoxic environments.

In this study, we characterized microorganisms with *hgcAB* and their geochemical context in anoxic sediment and water from two sulfate-impacted lakes in the St. Louis River watershed, northern Minnesota. Several tributaries to the St. Louis River receive discharges from historic and active iron mining operations that raise surface water sulfate concentrations to over 200 mg/L. The origins of MeHg are of interest in the watershed due to restrictive fish consumption advisories, and previous studies have addressed the complex relationship between MeHg accumulation and sulfate loading (e.g., refs. [34–37] and references therein). These studies show that, across the watershed, methylation rates and MeHg accumulation are not always clearly related to sulfate concentration. For example, wetlands and anoxic lake sediments in the watershed that have been chronically impacted by sulfate have similar or lower MeHg production rates than unimpacted sites [34, 36], and more MeHg export occurs from the wetland-rich, low-sulfate tributaries than those impacted by high sulfate loads [35]. Currently, little is known about the microbial communities responsible for MeHg production in the region, and knowledge of the relationship between community composition and MeHg production may help elucidate the origins of variability in methylation rates across the watershed. We therefore performed a detailed microbiological analysis of methylating communities from two lakes that, among sulfate-impacted sites, represent a stark contrast in terms of sulfate loading, sulfide concentration, and MeHg production potential. We combined *hgcA* gene cloning, rRNA methods, and metagenomics with measurements of mercury methylation rates and other geochemical parameters in order to (i) identify populations with the genetic potential to methylate mercury, (ii) determine how and why those populations differ between the two lakes and in the water column versus the sediment, and (iii) evaluate the geochemical implications of differences in methylating communities between the lakes.

Materials and methods

Field sites and sample collection

We collected sediment and water from Lake Manganika (N 47.49°, W 92.57°) and Lake McQuade (N 47.42°, W 92.77°) in the St. Louis River watershed. Both lakes have neutral to slightly alkaline pH, seasonally anoxic bottom waters, and are impacted by sulfate loading from mines on Minnesota's Iron Range [34]. Lake Manganika is a hypereutrophic lake with seasonally anoxic and sulfidic bottom waters (up to 3 mM dissolved sulfide) that receives elevated nutrients from a municipal wastewater treatment plant and sulfate-rich water from a dewatering taconite mine pit [34]. Manganika sediments and bottom waters are perennially sulfidic. In

contrast, Lake McQuade is a mesotrophic lake that has seasonally anoxic bottom waters and low sulfide concentrations in the sediment pore water and bottom water (<0.03 mM) [34].

Samples for this study were collected from Manganika on 6th June 2013, and from both lakes on 31st July 2014. Sediments were collected with an HTH Teknik gravity corer. Samples for molecular analyses were collected by immediately extruding the top 0–2 cm of the core in the field and aseptically collecting material from the center. Samples for DNA extraction were either immediately frozen on dry ice, or first preserved with 5 parts RNeasy (Thermo Fisher Scientific, Waltham, MA, USA) to 1 part sample and then immediately frozen on dry ice. Samples were transferred to -80°C until analysis. Water samples for microbiological analyses were collected at 5.6 m (Manganika 2013), 5.75 m (Manganika 2014), and 4.0 m (McQuade) depth. A water sample was also collected from the Manganika oxycline in 2013 at 3.6 m depth. Biomass was collected from bottom waters by filtering 50–150 mL of water through 0.2 μm polycarbonate or cellulose acetate filters (Whatman). Filters were either immediately frozen on dry ice after collection, or first bathed in RNeasy and then immediately frozen.

Water column profiles of dissolved oxygen, pH, oxidation–reduction potential, and electrical conductivity were measured with a multiparameter Hydrolab S5 sonde (Hydromet). Water samples for geochemical analyses were collected with a peristaltic pump and Teflon tubing. Samples for sulfide determination were immediately filtered (0.2 μm) and preserved with zinc acetate. Samples for dissolved sulfate, nitrate, and iron analyses were filtered (0.2 μm) and acidified to 0.5% (vol/vol) trace metal grade HCl and bubbled gently with a stream of N_2 to minimize the conversion of sulfide to sulfate. Pore waters were collected from the top 2 cm of material from composited cores using Rhizon® samplers connected to PTFE tubing and stainless steel hypodermic needles in an anoxic atmosphere. Filtered (0.2 μm) and unfiltered water samples were collected into mercury-free PETG bottles and acidified with trace metal grade HCl to 0.5% (vol/vol). Sediment cores and water samples for methylation rate assays were stored in coolers during transport, and the methylation assays were initiated less than 8 h after collection upon return to the University of Minnesota Duluth.

Chemical analytical methods

Methylation and demethylation potentials were determined with an enriched stable isotope incubation technique as described previously [13, 38], following the handling procedures of Bailey et al. [34] and Johnson et al. [36]. Briefly, sediment cores were injected at 1 cm intervals with a

mixture of stable isotope-enriched $^{200}\text{Hg}^{2+}$ and $\text{Me}^{201}\text{Hg}^{+}$ (94.3% $^{200}\text{Hg}^{2+}$ and 84.7% $\text{Me}^{201}\text{Hg}^{+}$) at quantities similar to in situ Hg and MeHg concentrations (0.5–2 ng/g MeHg and 50–300 ng/g inorganic Hg), using stock solutions that had first been equilibrated for 1 h with anoxic filtered water or pore water from the water column or sediments. Cores were incubated in a water bath at in situ temperatures for 7 h, and the top 2 cm extruded and frozen until analysis. Water samples were incubated at in situ temperatures for 24 h and then frozen until analysis. Concentrations of total mercury (THg), MeHg, and excess ^{200}Hg and ^{201}Hg as THg and MeHg, were quantified as described in Bailey et al. [34] using an Agilent 7700x ICP-MS hyphenated either to a Tekran 2600 automated THg analyzer or to a thermal desorption and gas chromatography front-end for MeHg. Quantification detection limits are given in Table S1. %MeHg values refer to the percentage of total Hg present as MeHg. As in Mitchell et al. [13], rate constants (k_{meth} and k_{demeth}) determined with this technique are referred to as potential rate constants because the additional mercury that was introduced may be more bioavailable than ambient mercury. k_{demeth} values are often considerably greater than k_{meth} , but since demethylation affects a much smaller pool of Hg, k_{demeth} cannot be considered in isolation as equally affecting the net concentration of MeHg (see Hintelmann et al. [38] for complete treatment of this phenomenon).

Sulfate and nitrate were determined by ion chromatography with a Dionex ICS 1100 system (Method 300.1, US EPA, 1997). Sulfide was determined using the methylene blue method (4500-S2–E, [39]). Major cations and phosphorus were quantified with inductively coupled plasma optical emission spectroscopy (ICP-OES) at the University of Minnesota Analytical Geochemistry Laboratory using a Thermo Scientific iCAP 6500 dual view ICP-OES. For the 2013 Manganika samples, ferrous iron (Fe^{2+}) was determined with the phenanthroline method (3500-FeB, [39]). Additional data from the 2013 Manganika samples were reported in Bailey et al. [40].

Nucleic acid extraction and amplicon library preparation

Total DNA was extracted from sediments using the PowerSoil DNA isolation kit (Mo Bio Laboratories, Inc., Carlsbad, CA, USA), except that aliquots of each sample were bead beaten for 5, 10, and 15 min and subsequently recombined. DNA was isolated from filters using the PowerWater DNA isolation kit (Mo Bio). RNeasy was removed prior to extraction by centrifugation. Small subunit 16S rRNA gene “amplicon” libraries were produced by sequencing the V4 hypervariable region of the 16S rRNA gene with the full-service amplicon sequencing service at the University of Minnesota Genomics Center (UMGC)

[41, 42] with primers 515f (GTG CCA GCM GCC GCG GTA A) and 806r (GGA CTA CHV GGG TWT CTA AT). Amplicon libraries were deposited in the Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) under BioProject ID PRJNA488162.

***hgcA* gene cloning**

Partial *hgcA* genes were amplified with the primer set *hgcA*261F (5'-CGGCATCAAYGTCTGGTGYGC-3') and *hgcA*912R (5'-GGTGTAGGGGTGCAGCCSGTRWAR KT-3') from Schaefer et al. [23]. PCR was performed using HotStarTaq (Qiagen, Venlo, The Netherlands) and 0.5 µM of each primer, with 5 min initial denaturation at 95 °C, 35 cycles of 30 s denaturation at 95 °C, 30 s annealing at 60 °C, and 60 s elongation at 72 °C, with 7 min final elongation at 72 °C. PCR products were purified using the DNA Clean & Concentrator-5 Kit (Zymo Research, Irvine, CA, USA) according to the manufacturer's instructions. Purified products were ligated into the pCR4-TOPO TA vector, which was used to transform One Shot Mach1-T1 chemically competent *Escherichia coli* (Invitrogen, Carlsbad, CA). Transformed *E. coli* were grown on lysogeny broth (LB) agar plates with 50 µg/mL kanamycin, and approximately 50 successful transformants picked and screened by colony PCR with M13 primers. Clones of the correct sized inserts were cleaned with the DNA Clean & Concentrator-5 Kit (Zymo) and sequenced at the UMGC with an ABI 3730XL sequencer and ABI BigDye Terminator chemistry v. 3.1 (Applied Biosystems). Sequences are archived in GenBank under accession numbers MH809186–MH809356.

Metagenome library preparation

We created metagenomes from the bottom water and sediments from each lake (Table 1). For Manganika sediments, DNA extracts from replicate core samples were pooled in proportion to their DNA concentration, as determined with a Nanodrop 2000 instrument (Thermo Scientific) (Table S2). Separate datasets were created from two replicate McQuade sediment samples. Environmental DNA was sequenced on a HiSeq 2500 (Illumina, Inc., San Diego, CA, USA) in high-output mode with 125 bp paired reads, following library preparation by the TruSeq Nano protocol with a target insert size of 650 bp.

Amplicon and clone library sequence analysis

Raw 16S rRNA gene amplicon libraries were processed using the pipeline described in Jones et al. [42], except that the Silva database v.128 [43] was used for taxonomic classification. Hierarchical agglomerative cluster analyses were computed in R v.3.4.0 [44] with Vegan package

v.2.4–3 [45], using Bray–Curtis dissimilarity and unweighted pair-group method using arithmetic averages (UPGMA) clustering. R-mode cluster analysis included only those operational taxonomic units (OTUs) that were ≥2% in one or more samples.

hgcA clone sequences were evaluated for quality using A Plasmid Editor (ApE) (<http://biologylabs.utah.edu/jorgensen/wayned/ape/>), trimmed so that only the *hgcA*261F and *hgcA*912R flanked the sequences, and, for some sequences, forward and reverse reads assembled using Sequencher (Gene Codes Corp., Ann Arbor, MI, USA). Sequences were translated and aligned in a custom ARB database described below [46]. Putative chimeras were identified with Bellerophon 3 [47] and removed.

Assembly, binning, and annotation of metagenomic datasets

Raw metagenomic reads were quality trimmed and filtered with Sickle v.1.33 (mean quality ≥28, minimum length 50) [48], residual adaptors removed with cutadapt v.1.11 [49], and dereplicated (exact forward and reverse sequences) and low complexity sequences removed (DUST score ≥7) with prinseq v.0.20.4 [50]. PhiX reads were removed with BLASTN [51]. The seven datasets were then co-assembled with metaSPAdes [52] using default parameters, and reads from each dataset mapped back to the co-assembly using bowtie2 v.2.2.6 [53]. Automated binning was performed with MaxBin 2.0 [54]. These automated bins were screened for candidate *hgcA* sequences by BLASTX analysis against *hgcA* from the database described below. Each of the scaffolds with a candidate *hgcA* homolog were crudely annotated using prokka v.1.12 [55], and true *hgcA* sequences were then identified by phylogenetic comparison of the open reading frames from prokka to bona fide *hgcA* sequences in ARB [46]. Criteria for the determination of true *hgcA* also included the presence of conserved cysteine motif C93 [56] and a transmembrane region [16]. *hgcB* sequences were identified as ferredoxin homologs (presence of multiple CXXCXXC motifs) immediately adjacent to the *hgcA* in the prokka annotation. Transmembrane helices were predicted with TMHMM v.2.0 [57].

Bins with *hgcA* were manually refined in anvi'o v.2.3.2 [58], and the inclusion of *hgcA*-containing scaffolds in the bins was confirmed by visually inspecting coverage profiles in anvi'o. Bin completeness and contamination were evaluated with CheckM v.1.0.7 [59] and anvi'o. All of the refined bins with *hgcA* were then annotated with prokka, and bins that were >75% complete based on CheckM were submitted for annotation at the Integrated Microbial Genomes (IMG) online workbench [60]. The identities of the bins were initially determined with CheckM and more precisely for some by phylogenetic analysis of translated

Table 1 Mercury concentration and potential methylation rate constants determined using enriched stable isotopes

Lake	Date	k_{meth} (per day) ^a	k_{demeth} (per hour) ^a	MeHg (ng/L)	THg (ng/L)	%MeHg
Sediment						
Manganika	3 June 2013	0.028	0.059	1.36	226.9	0.60
Manganika	31 July 2014	0.014	0.061	2.21	282.3	0.78
McQuade	31 July 2014	0.022	0.036	2.40	163.0	1.47
Bottom water						
Manganika	3 June 2013	0.009	0.014	0.94	1.1	87.98
Manganika	31 July 2014	0.038	BD	2.28	3.7	61.43
McQuade	31 July 2014	0.024	BD	3.60	5.9	60.80
Chemocline						
Manganika	3 June 2013	BD	BD	0.46	2.6	8.26

MeHg and THg concentrations for the bottom waters and chemocline are aqueous concentrations, sediment values are solid-phase concentrations (pore water values are given in Table S3)

BD below detection

^a k_{meth} and k_{demeth} were assessed by additions of isotopically enriched $^{200}\text{Hg}^{2+}$ and $\text{Me}^{201}\text{Hg}^+$

gene sequences for RNA polymerase beta subunit (RpoB) and the DNA recombination and repair protein RecA in custom ARB databases. Some bins were incomplete and did not contain *recA* or *rpoB*. The relative abundance of each bin was determined by calculating the proportion of reads recruited to each bin in each dataset by the total number of reads in that dataset.

Raw sequence data for the metagenomes was deposited in the Sequence Read Archive (SRA; <http://www.ncbi.nlm.nih.gov/sra>) under BioProject ID PRJNA488162, and the assembly and annotated bins are accessible at IMG (GOLD study ID Gs0130353).

Phylogenetic analyses

Contiguous *hgcA* and *hgcB* sequences recovered from the bins were concatenated and used to build an ARB database [46] that also included *hgcAB* from a core set of predicted mercury methylating organisms [16]. Nucleotide sequences were translated in ARB and the amino acid sequences were exported, aligned using M-Coffee [61], and reimported. Prior to phylogenetic analysis, concatenated, full-length *hgcAB* sequences were filtered to mask terminal gaps, internal positions with >50% gaps, and a poorly aligned internal region corresponding to residues 261–283 in *Desulfovibrio desulfuricans* ND132. The final alignment had 386 positions. Maximum likelihood analysis was computed with RAXML v.8.0.24 [62] using the LG model of amino acid substitution [63] with observed amino acid frequencies and the fraction of invariant sites and the shape parameter (α) value estimated from the data (model PROTGAMMAILGF). These parameters were selected using the corrected Akaike information criterion (AICc) in ProtTest v.3.4.2 [64, 65]. Confidence was assessed with 1000 rapid bootstrap replicates, and the HgcAB tree was

rooted with the clade of fused HgcAB sequences identified by Podar et al. [18]. Partial *hgcAB* sequences from meta-genomic bins were added using the evolutionary placement algorithm (EPA) [66]. The alignment and ARB database of concatenated *hgcAB* used for these analyses are provided as Supplementary datasets 1 and 2.

Phylogenetic analysis of partial *hgcA* clones was performed in ARB, using a database of *hgcA* sequences (no *hgcB*) that also included carbon monoxide dehydrogenase/ acetyl-CoA synthases delta subunit sequences (pfam03599). Phylogenetic analysis of *hgcA* clones only included those positions covered by the clone sequences, masking the same positions as for the *hgcAB* phylogeny above. Neighbor joining analysis was performed using the ARB implementation of Phylip using the Jones–Taylor–Thornton correction [67, 68].

RpoB and RecA sequences were identified in the prokka and IMG annotations and aligned in ARB with their top BLASTP matches. Variable regions were omitted by removing positions with more than 50% gaps, resulting in alignments of 1052 and 318 positions for RpoB and RecA, respectively. Maximum likelihood trees were constructed with RAXML using the same evolutionary model as for *hgcAB*, above.

Results

Field observations and geochemistry

Lake Manganika was hypereutrophic and had soupy green surface waters (Figure S1). Near the surface, dissolved oxygen approached 16 mg/L (nearly 200% saturation) and pH exceeded 8.7 (Fig. 1, Table S3). Dissolved oxygen and pH decreased with depth, and anoxic conditions were

Fig. 1 Geochemical profiles from the lakes sampled for this study. % MeHg values from sediment pore waters refer to the dissolved phase. Thick black arrows indicate the location of bottom water and sediment

samples collected for microbiological analyses. The depth on the Y-axis differs between the two lakes

encountered below 4 m (Fig. 1). During the early summer sampling expedition (6 June 2013), the bottom waters had 0.18 mM sulfide and 0.02 mM nitrate, but in late summer (31 July 2014), the bottom waters had 1–2 mM sulfide and

no detectable nitrate. Though sulfate remained above 3 mM throughout the water column in all seasons, higher sulfide and undetectable nitrate in the late summer is consistent with the development of strongly reducing conditions

Table 2 Summary of microbiological analyses

Lake	Date	# rRNA gene libraries	# <i>hgcA</i> clones	Metagenome library size (Gbp)
Sediment				
Manganika	3 June 2013	2	23	5.24
Manganika	31 July 2014	7	33	6.61
McQuade	31 July 2014	2	46 ^a	10.90 ^a
Bottom water				
Manganika	3 June 2013	1	21	6.46
Manganika	31 July 2014	2	29	5.28
McQuade	31 July 2014	3	22	5.69
Chemocline				
Manganika	3 June 2013	1	0	0

^aPrepared from two samples from separate sediment cores

during summer stratification [34]. Sediment pore waters had >2 mM sulfide and <0.5 mM sulfate during both sampling periods, consistent with previous observations that pore water sulfide remains high in Lake Manganika despite seasonal sulfide fluctuations in the overlying water column [34]. The sediment surface was covered with a fluffy gray biomat with small finger-like protrusions that rose 1–2 cm in relief and were a few millimeter in diameter (Figure S1). Similar layers of what appeared to be buried biomat were visible deeper in the cores.

Lake McQuade was mesotrophic [34] and had much higher surface water clarity compared to Manganika (Figure S1) but an order of magnitude lower sulfate (~0.5 mM). Anoxic conditions occurred 3 m below the water surface. A green microbial layer was encountered just below the oxycline at 4 m (Figure S1). This layer was not observed at 4.5 m depth. Sulfide concentrations were ≤0.03 mM in the bottom water and sediment pore water (Fig. 1, Table S3).

In both lakes, between 60% and 90% of the dissolved mercury was methylated in the bottom waters (Table 1). Dissolved mercury in sediment pore waters was 5–25% MeHg (Table S3). Solid phase MeHg was 0.6–1.5% of the total sediment mercury pool (Table 1). Potential mercury methylation rate constants (k_{meth}), as measured using enriched stable isotope labeling, ranged from 0.014 to 0.028 per day in the sediments and 0.009–0.038 per day in the bottom waters. k_{meth} in waters at the oxycline were only measured at Manganika in 2013, and were below detection (Table 1). k_{demeth} was consistently 0.035–0.06 per hour in sediments, but near or below detection limits in bottom waters and in the oxycline.

rRNA gene amplicon libraries

We generated 18 total 16S rRNA gene amplicon libraries (Table 2) with between 59,732 and 505,242 sequences each (Table S2). Replicate libraries created from separate

sediment cores or water filtered from the same site clustered together and had similar taxonomic profiles, independent of sample preservation (Figures S2, S3). Libraries from the early summer Manganika oxycline and bottom waters had relatively more *Betaproteobacteria* than other samples (Figure S2), and the most abundant OTUs included a member of the genus *Albidiferax* sp. (12–19% of sequences) and an OTU from the sulfur- and hydrogen-oxidizing family *Hydrogenophilaceae* (5–13%) [69] (Figure S3). In contrast, the most abundant OTU in the late summer Manganika bottom water libraries was classified as the sulfate-reducing genus *Desulfobacula* (7–10%), likely because strongly reducing conditions had developed in the hypolimnion by that time (Fig. 1). The McQuade bottom water libraries were dominated by a single OTU related to the anoxic phototrophic genus *Chlorobium* (Figure S3), which was likely responsible for the green layer we encountered when sampling just below the oxycline (Figure S1). Additionally, diverse *Bacteroidetes*, *Betaproteobacteria*, and *Deltaproteobacteria* were abundant in the bottom waters of both lakes. Sediment libraries contained many of the same OTUs that were present in the water column, but were also more diverse than the planktonic communities and included taxa that were not abundant in the bottom waters such as *Methanosaeta*, *Aminicenantes*, and *Desulfatiglans* (Figure S3). Manganika sediments were littered with algal detritus, evident from fluorescence microscopy (not shown) and the abundant chloroplast and cyanobacterial sequences in the amplicon libraries (Figure S3).

Sequences from classes typically associated with MeHg production—the *Deltaproteobacteria*, *Clostridia*, and *Methanomicrobia*—were present in most rRNA gene libraries (Figure S4). *Deltaproteobacteria* accounted for 4–18% of sequences in anoxic bottom waters and sediments. The sulfate-reducing *deltaproteobacterial* orders *Desulfobacterales* and *Desulfovibrionales* were present in all bottom water and sediment samples, as was the iron-reducing

Fig. 2 Summary of *hgcA* gene clone libraries. Numbers at the right indicate the number of clones from each clone that were present in each clone library ('-' indicates no clones recovered for that clone). The clades that represent methylators from the *Deltaproteobacteria*, *Clostridia*, and *Methanomicrobia* are shown in bold font. A phylogenetic tree showing uncollapsed nodes is provided as Figure S5

family *Geobacteraceae*. *Peptococcaceae*, a family in the *Clostridia* that includes inorganic sulfur reducers capable of MeHg production, were rare, making up <0.25% of sequences in the sediment libraries and <0.02% in the water libraries. *Methanomicrobia* represented 1.5–3% of the sediment libraries and <0.2% of the water libraries from both lakes.

hgcA gene cloning

We created *hgcA* gene clone libraries from the sediments and bottom waters from each lake, each composed of 21–46 clones (Table 2). All libraries included *hgcA* clones that grouped with sequences from sulfate-reducing *Deltaproteobacteria* of the genera *Desulfobulbus*, *Desulfovibrio*, and *Desulfomicrobium* (group “Deltaproteobacteria clade 1” in Fig. 2). *hgcA* clones from the *Geobacteraceae* were abundant in the less sulfidic samples, including McQuade sediment and surface water libraries and in the Manganika 2013 water column. *hgcA* clones from the *Methanomicrobia* were most abundant in the sediment libraries (Fig. 2). The majority of Manganika sediment clones were from a group that could not be reliably placed in *hgcA* phylogenies (“MNG clone group 1”; Figs. 2, S5) and likely represents unknown *hgcA*-containing organisms. *hgcA* clones also occurred in four other groups that, like MNG clone group 1, could not be identified by *hgcA* phylogeny alone. One of these clades of unknown *hgcA* sequences is represented by clone MCQS4, and the other clades are designated clone/metagenome groups 2 and 3 and Clostridia-like clone/metagenome group (Fig. 2).

Metagenomic reconstruction of *hgcAB*-containing organisms

In order to verify *hgcA* cloning results and identify the unknown *hgcA*-containing populations, we generated metagenomes from the bottom waters and sediments of both lakes (Table 2). Co-assembly of reads from the seven metagenomes resulted in 885,923 contigs >500 bp. Automated binning with MaxBin 2.0 produced 140 bins that were >50% complete. These preliminary bins were screened for *hgcA*, and manual refinement of the *hgcA*-containing bins in anvi'o resulted in 27 bins with an *hgcA* gene, all with <7.5% estimated contamination and 18 of which were >75% complete (Table 3). The manual refinement step resulted in some *hgcA*-containing bins with low percent completion estimates, in some cases <10% (Table 3). These low percent completion bins occurred if the *hgcA*-containing scaffold did not group with the majority of scaffolds in the bin (e.g., Figure S6). Twenty-three of the 27 bins had adjacent *hgcA* and *hgcB* homologs on the same contig. Of the remaining four bins, the *hgcA* gene in bins 61, 67, and 75 occurred at the end of a contig, and bin 158 did not have an *hgcB* homolog on the same contig. Nevertheless, the *hgcA* sequences in these four bins cluster with true *hgcA* homologs based on phylogenetic analysis (Fig. 3). Bin 89 contains a contiguous *hgcAB* with an incomplete *hgcA* at the 5' end. Tables S4 and S5 contain additional detailed information on the metagenomic bins.

In the *HgcAB* peptide alignment, all of the translated metagenomic *hgcA* have the conserved cysteine residue (C93) that has been shown to be essential for methylation in

Table 3 Summary of metagenomic bins with *hgcA*

Bin name ^a	% Complete ^b	% Contam. ^b	Bin size (Mbp)	# Scaffolds	Relative abundance (%) ^c						HgcAB group	Inferred taxonomy	
					McQuade late summer (2014)		Manganika early summer (2013)		Manganika late summer (2014)				
					Sed ^d	Sed ^d	Water	Sed	Water	Sed			Water
bin89	15.79	0	0.9	188	0.001	0.009	0.027	0.002	0.056	0.002	0	Geobacter-like	Order Desulfuromonadales
bin111	94.46	0.88	4.1	295	0.001	0.015	0	0.21	0.18	0.71	3.7	Delta 1	Desulfobacula sp.
bin47	79.72	1.49	3.8	475	0.040	0.064	0.42	0.020	0.17	0.020	0.13	Delta 1	Order Desulfobacterales
bin55_r2	47.37	3.51	3.2	270	0.044	0.037	0.41	0	0	0.001	0	Delta 1	Order Desulfobacterales
bin158	77.46	0	3.5	373	0	0	0.001	0.005	0.004	0.028	0.34	Delta 1	Order Desulfovibrionales
bin61	21.47	0	1.4	328	0.006	0.005	0.095	0	0	0	0	Delta 1	Unknown
bin67	8.77	0	0.3	85	0	0	0	0.016	0.010	0.002	0	Delta 1	Unknown
bin77	4.17	0	0.2	70	0.006	0.008	0.002	0.001	0.001	0.001	0.005	Delta 1	Unknown
bin85	93.84	2.26	5.1	355	0.001	0.002	0.001	0.66	0.031	1.2	0.009	Delta 2	Order Desulfobacterales
bin5	78.02	1.75	4.8	738	0.36	0.21	0	0	0	0	0	Delta 2	Order Desulfobacterales
bin95	89.03	0.65	2.9	104	0.002	0	0	0.14	0.011	0.21	0.18	Delta 2	Smithella sp.
bin12	46.4	2.58	2.0	502	0.14	0.072	0	0	0	0	0	Delta 2	Order Desulfobacterales
bin26	18.73	0.65	0.9	215	0.012	0.027	0	0.015	0.002	0.048	0.027	Delta 2	Unknown
bin161	94.97	3.84	3.9	74	0.0002	0.0002	0	0.023	0.85	0.024	0.013	<i>Clostridia</i>	Class Clostridia
bin112	80.96	5.03	4.5	404	0	0	0	0.056	0.29	0.083	0.082	<i>Clostridia</i>	Class Clostridia
bin81	92.16	6.21	2.6	152	0	0	0	0.024	0.14	0.068	0.20	<i>Clostridia</i>	Class Clostridia
bin97	96.73	3.27	3.0	206	0.001	0.003	0	0.23	0.040	0.46	0.041	<i>Methanomicrobia</i>	Family Methanoregulaceae
bin144	95.45	5.98	8.0	171	0	0.014	0	2.0	1.3	7.3	0.83	Unknown	Ca. phylum Aminicenantes
bin93	90.03	1.55	2.7	104	0.003	0.007	0.013	0.17	0.081	1.4	1.9	Unknown	Phylum Kiritimatiellaeota
bin55_r1	76.41	7.35	3.1	321	0.009	0.009	0.50	0	0	0	0	Unknown	Phylum Kiritimatiellaeota
bin129	44.37	1.46	2.6	299	0.007	0.001	0.054	0	0.004	0.072	0.41	Unknown	Phylum Lentisphaerae
bin159	97.6	2.67	3.3	101	0	0	0	0.003	0.009	0.35	0.72	Unknown	Spirochaetes
bin43	92.3	1.15	4.1	383	0.044	0.039	0.007	0.11	0.71	0.078	0.10	Unknown	Spirochaetes
bin152	96.63	6.84	4.4	127	0	0	0	0.032	0.29	0.19	0.21	Unknown	Spirochaetes
bin180	78.31	1.2	3.1	136	0	0.002	0	0.001	0	0.094	0.40	Unknown	Spirochaetes
bin75	44.44	1.75	2.3	336	0.002	0	0	0.074	0.087	0.057	0.037	Unknown	Spirochaetes
bin143	98.88	2.25	3.7	162	0.004	0.006	0.001	0.12	0.030	0.37	1.1	Unknown	Unknown

^aBins are grouped by taxonomy and ordered in decreasing relative abundance within those groups^b% Complete and contamination determined using checkm. Additional statistics can be found in Supplementary Tables S3 and S4^cTotal reads recruited per bin/total reads per sample. Relative abundance values of >0.1% are displayed in bold font^dSeparate metagenomes were constructed from replicate cores

Fig. 3 Maximum likelihood analysis of full-length *hgcAB* sequences from isolates and metagenomes, with partial sequences from bins 61, 67, 75, 89, and 158 added using the EPA algorithm [66]. Numbers at nodes indicate bootstrap values >50. The clade of fused *hgcAB*-like sequences were used as an outgroup after Podar et al. [18]

D. desulfuricans ND132 [56]. C93 is preceded by a tryptophan residue in all metagenomic HgcAB (Supplementary Figure S7). All of the HgcB homologs from the

metagenomes have two 4Fe-4S regions (CXXCXXCXXC motifs), and all have C73 except for bins 77 and 111, which have a stop codon at that position (Supplementary

Figure S7). All metagenomic HgcAB include the trans-membrane region.

Phylogenetic analysis of HgcAB peptide sequences resulted in several clades (Fig. 3). Consistent with previous analyses [16, 18], HgcAB from the *Deltaproteobacteria* grouped into three clades, one with sequences from *Geobacter* spp. (clade “*Geobacter*-like” in Fig. 3), one with sequences from the genus *Desulfovibrio* and several other sulfate-reducing genera (“*Deltaproteobacteria* clade 1”), and one that includes sequences with the genera *Desulfomonile*, *Syntrophus*, and *Syntrophorhabdus* (“*Deltaproteobacteria* clade 2”) (Fig. 3). HgcAB from the *Methanomicrobia* form a monophyletic clade, as do the fused HgcAB sequences identified by Podar et al. [18]. HgcAB from the *Clostridia* occur in a clade with low bootstrap support that includes some *Spirochaetes* sequences from the metagenomes (discussed below). Several HgcAB from Manganika and McQuade metagenomes, including sequences from the *Spirochaetes*, PVC superphylum, and *Aminicenantes*-like bins, fall outside of the five major clades described above.

Deltaproteobacteria

The majority of the metagenomic *hgcAB* sequences belong to members of the *Deltaproteobacteria*. Seven metagenomic *hgcAB* occur in “*Deltaproteobacteria* clade 1”, five in “*Deltaproteobacteria* clade 2”, and one in the “*Geobacter*-like” clade. The *deltaproteobacteria* bins have lower contiguity compared to other bins, and many are incomplete. Six *deltaproteobacteria* bins are >75% complete and were therefore submitted to the IMG system for annotation. Four of these bins—bins 5, 47, 111, and 158—have dissimilatory sulfite reductase (*dsrAB*), confirming the capacity for sulfate reduction in some members of *Deltaproteobacteria* clades 1 and 2. Bins 85 and 95 in *Deltaproteobacteria* clade 2 do not have *dsrAB*. Bin85 contains other genes involved in the reduction of inorganic sulfur compounds including polysulfide reductase (*psrA*) and adenosine-5'-phosphosulfate reductase (*aprA*), and the bin is 93.8% complete (Table 3), so it is possible that it is capable of sulfate reduction but *dsr* genes are missing due to an incomplete bin. Bin 95 belongs to the genus *Smithella* (Figures S8, S9), and it likely has a fermentative or syntrophic lifestyle like other members of this genus [70]. Bins 12 and 26 also encode *dsrA* and therefore represent sulfate-reducing populations. The most abundant *deltaproteobacterium*, bin111, lacks C73 in its HgcB, indicating that it might not be capable of mercury methylation.

Clostridia

Three metagenomic bins (81, 112, and 161) represent members of the class *Clostridia*. These bins primarily occur

in the Manganika water column (Table 3). The HgcAB sequences from these bins do not group with known sulfate-reducing members of the *Clostridia*, and are instead either most closely related to HgcAB from the fermentative or acetogenic genera *Acetivibrio*, *Syntrophobotulus*, *Ethanoli-genens* and *Acetonema*. The populations represented by these three bins are also likely fermenters and acetogens. Bin 81 encodes a complete Wood–Ljungdahl pathway, as well as phosphate acetyltransferase and acetate kinase genes, indicating that this bin represents an acetogen. Bin 112 contains a complete glycolysis pathway, three acetate kinases, aldehyde and alcohol dehydrogenases, two lactate dehydrogenases, and an incomplete TCA cycle, consistent with a fermentative lifestyle like that of its closest relatives in the genera *Clostridium* and *Roseburia* (Figures S8, S9). Bin 161 also likely has a fermentative lifestyle, based on the presence of genes for a complete glycolysis pathway, phosphate acetyltransferase and acetate kinase (conversion of acetyl-CoA to acetate by acetylphosphate), and alcohol and aldehyde dehydrogenases. This bin also encodes a complete TCA cycle, a F-type ATPase, subunits of complex 1 and 2 (*nuoEFG* for NADH dehydrogenase and *sdhAB* and *frdA* for succinate dehydrogenase/fumarate reductase), and cytochrome BD, indicating potential respiratory capabilities.

Methanomicrobia

Bin 97 represents a member of the *Methanomicrobia* with *hgcAB*. Consistent with its classification in the family *Methanoregulaceae*, Bin97 contains a complete pathway for hydrogenotrophic methanogenesis. This bin was most abundant in Manganika sediments.

Aminicenantes-like population

Bin 144 is either a member or a close relative of candidate phylum *Aminicenantes* (formerly OP8; [71]) (Figures S8, S9). Unlike previously identified *hgcAB*-containing *Aminicenantes* [18], the *hgcA* and *hgcB* genes are not fused in this population. This bin represents the most abundant potential methylator in Manganika overall, but is absent or rare in McQuade.

Bin 144 appears to have diverse metabolic capabilities. It has a complete genetic complement for glycolysis and the TCA cycle, and encodes respiratory complexes cytochrome C oxidase (*coxABCD*), succinate/fumarate dehydrogenase (*sdhABCD*) and an F-type ATPase. The presence of *cox-ABCD* and a second terminal oxygenase, *cytBD*, indicate that Bin 144 might represent a facultative aerobe, or it could use these oxygenases for detoxification purposes under microoxic conditions. It also appears to have other respiratory capabilities. While Bin 144 does not appear to

have the potential for dissimilatory reduction of sulfate or nitrogen compounds, it can likely reduce intermediate sulfur compounds based on the presence of genes for thiosulfate reductase and tetrathionate reductase. In the anoxic lake environment, these organisms could be metabolizing organic carbon by fermentation or coupled to the reduction of intermediate sulfur compounds.

PVC superphylum

One of the other most abundant *hgcAB*-containing bins, Bin 93, is a member of the PVC superphylum (Table 3, Figure S8). Its closest relative is *Kiritimatiella glycovorans*, a member of the newly proposed phylum *Kiritimatiellaeota* that was formerly *Verrucomicrobia* subdivision 5 [72] (Figures S8, S9). The *hgcAB* from Bin 93 clusters with two other members of the PVC superphylum: Bin 55_r1, which is another member of *Kiritimatiellaeota*, and Bin 129, which appears to be a member of the *Lentisphaerae* (Figure S9). Bin 93 was most abundant in the late summer Manganika metagenomes and rare in McQuade samples, while Bin 55_r1 was present in McQuade and absent from Manganika.

These *Kiritimatiellaeota* bins have genes for the degradation of complex polysaccharides, such as cellulases and other abundant glycosyl hydrolases (12 in Bin 93, 21 in Bin 55_r1). These features, along with a lack of electron transport chain components and a complete glycolysis pathway in the case of Bin 93 (bin 55_r1 is missing two steps but is only 76% complete), indicate that Bins 93 and 55_r1 share the fermentative lifestyle of *K. glycovorans* and that they may degrade complex polysaccharides like other members of the PVC superphylum.

The *hgcA* and *hgcB* from Bin 93 are fused into a single open reading frame (Figure S10), but the *hgcAB* from Bin 93 is not phylogenetically related to previously discovered sequences with fused *hgcAB* (Fig. 3) [18]. Unlike those other fused *hgcAB* sequences, it also contains residue C73 that is crucial for methylation (Figure S7) [56], and read mapping confirms that this apparent fusion is not an assembly error (Figure S11). However, the transmembrane region of this gene only contains two helices (Figure S12), in contrast to the 5 and 4 helices of the true HgcAB and fused HgcAB-like sequences. The *hgcAB* in the other *Kiritimatiellaeota* bin, Bin 55_r1, are not fused and otherwise resemble true *hgcAB* (Figs. 3, S12).

Spirochaetes

Bins 43, 75, 152, 159, and 180 are members of the *Spirochaetes*. In RpoB and RecA phylogenies, Bin 152 groups with strong bootstrap support with sequences from the genus *Leptospira*, *Turneriella*, and environmental

Spirochaetes [73]. A partial RecA sequence from Bin75 clusters with that from Bin 152. Bins 180 and 159 group in a larger clade of *Spirochaetes* and are most closely related to a group of environmental spirochaetes from Rifle, CO [73] in RpoB and RecA phylogenies, and Bin 43 is related to *Treponema* spp. based on RecA. The *hgcAB* from Bin 180 is divergent, and occurs as a basal lineage between the clade of fused *hgcAB* and the true *hgcAB* (Fig. 3). Despite its phylogenetic distance from other *hgcAB*, however, the *hgcAB* from Bin 180 contains the key amino acid residues (C93, C73), five transmembrane helices in HgcA (Figure S12) and 4Fe-4S motifs in HgcB [56]. Based on fermentation capabilities and an absence of respiratory complexes and genes for dissimilatory N or S metabolism, the spirochaetes with *hgcAB* in Manganika and McQuade likely have a saprophytic role like that common to other environmental spirochaetes (e.g., ref. [74]).

Unclassified Bin 143

We could not classify the population that is represented by Bin 143. RecA and RpoB sequences from this bin are only distantly related to sequences from public databases and have low bootstrap support. This bin encodes a complete glycolysis pathway and contains phosphate acetyltransferase, acylphosphatase, and acetate kinase, but lacks genes for a complete TCA cycle, major respiratory complexes, and dissimilatory N or S metabolism, indicating that it likely represents another member of the community that degrades organic compounds by fermentation.

Discussion

Novel potential mercury methylators from Lakes Manganika and McQuade

Although most known mercury methylators are from the classes *Deltaproteobacteria*, *Clostridia*, and *Methanomicrobia*, they also occur sporadically in other clades across the tree of life. In Lakes Manganika and McQuade, some of the most abundant populations with *hgcAB* include close relatives or members of the phyla *Aminicenantetes*, *Kiritimatiellaeota*, and *Spirochaetes*, showing that potential methylators from these “other” clades can dominate the methylating community in certain circumstances. Because most of these novel *hgcAB*-containing populations are from clades of uncultivated organisms, their specific contribution for MeHg production remains to be explored. However, this finding shows that diverse populations from outside the *Deltaproteobacteria*, *Clostridia*, and *Methanomicrobia* are likely important for MeHg production in the environment, and raises the potential for mercury methylation in diverse

geochemical niches such as environmental conditions conducive to intermediate sulfur cycling, fermentative, and syntrophic lifestyles. Potential methylators were especially abundant and diverse in sulfidic Lake Manganika (discussed below). Hypereutrophic, sulfate-rich freshwaters like Manganika provide complex organic substrates and abundant electron acceptors that sustain diverse geochemical niches in which a higher diversity of potential methylators may be realized, and may be important settings to search for novel, environmentally relevant methylating organisms.

The *hgcAB*-containing bins identified here also expand the diversity of potential methylators within classes of known methylators. For example, the *hgcAB* from bins 81, 112, and 161 are not closely related to *hgcAB* from other *Clostridia* (Fig. 3), which suggests additional methylator diversity among acetogenic and fermentative *Clostridia* in addition to those identified by Gilmour et al. [15]. Likewise, the spirochaete *hgcAB* identified here are distantly related to *hgcAB* from other taxa, and either form in their own clade or cluster with members of the *Firmicutes* (Fig. 3). *hgcAB* have been previously identified in one member of the spirochaetes, *Spirochaeta* sp. JC202 [17], which has an *hgcA* that clusters with the *Deltaproteobacteria* in phylogenetic analyses and is very similar to that of *Desulfovibrio aespoensis* Aspo-2 [17]. This suggests that *Spirochaetes* acquired the ability to methylate mercury from multiple sources by horizontal gene transfer.

We also identified a sequence representing a potentially novel *hgcAB*-like structure in Bin 93. Unlike other fused *hgcAB*-like sequences [18], the fused *hgcAB*-like gene from Bin 93 has all the key cysteine residues in its *hgcA* and *hgcB*-like regions (Figure S7), and in phylogenetic analyses, clusters with the true *hgcAB* (Fig. 3). However, unlike the true *hgcAB* and other fused *hgcAB*-like sequences, it only contains two predicted transmembrane helices in its HgcA-like region (Figure S12). Although we do not know for certain, we expect that, like other fused *hgcAB*-like sequences [18, 33], the fused *hgcAB*-like gene in Bin 93 may not retain a mercury methylating function.

Methods for detecting potential mercury methylators in the environment

There were notable differences between the clone libraries and metagenomes. While most of the same groups of *hgcA* were represented in clone libraries and metagenomes from the same sample, relative abundances differed (Figure S13). For example, *Geobacter*-like *hgcA* clones were noticeably more abundant in clone libraries compared to metagenomes (Figure S13). Likewise, sequences represented by MNG clone group 1 (Figs. 2, S5) were not present in the metagenomic bins, indicating that these sequences were not as abundant in the environment as cloning had appeared to

indicate (Figure S13). The presence of these and other potential methylators in the clone libraries that were not represented in the metagenomic bins (Figure S5) shows that additional *hgcA* diversity exists in these environments than we captured with our metagenomics approach. Differences between clone libraries and metagenomes also likely reflect PCR and cloning bias. More inclusive *hgcAB* primers have been developed since the cloning work reported in this study [19], which would likely reduce but not eliminate these biases.

We caution against inferring methylation potential based on 16S rRNA sequences alone. Although certain classes contain abundant species and strains capable of MeHg production (the *Deltaproteobacteria*, *Clostridia*, and *Methanomicrobia*), many or most of the organisms within those groups cannot methylate mercury (e.g., refs. [16, 18, 32]). For example, *Deltaproteobacteria* were 4–18% of the rRNA gene libraries, but *Deltaproteobacteria* with *hgcAB* only represented 0.4–4.4% of the metagenomes.

Implications for MeHg production in the environment

Lakes Manganika and McQuade are potential hotspots for MeHg production in the St. Louis River watershed. Frequent measurements over 2012–2013 found that McQuade sediments had consistently higher methylation potentials and MeHg accumulation compared with Manganika sediments [34, 36]. However, MeHg accumulated to over 3 ng/L in Lake Manganika bottom waters in 2012 [40], and MeHg concentrations in Manganika outflows of over 2 ng/L have been reported in previous years, indicating that high MeHg production and export can occur under certain conditions [75].

The microbial community analyses provided here may shed some light on the differences in MeHg production between the two lakes. Although potential methylators were more abundant in Manganika compared to McQuade, *Deltaproteobacteria* make up a much larger proportion of the *hgcAB* community in McQuade—85–90% of the potential methylators we identified in McQuade sediments are *Deltaproteobacteria* (0.45–0.60% of total metagenome sequences), most of which are capable of sulfate reduction, compared with only 17–27% in Manganika sediments (1.1–2.2% of total sequences). *Deltaproteobacteria* with *hgcAB* were also relatively more abundant in McQuade bottom waters (62.5% in McQuade versus 10.9–42.3% in Manganika). Many *Deltaproteobacteria* are capable of efficient MeHg conversion compared with methylators from other clades [15, 33]. It seems noteworthy that while we do not see a clear relationship between k_{meth} and the relative abundance of all potential methylating populations (Fig. 4a), we do see that in the bottom waters, the relative

Fig. 4 Relationship between potential methylation rate constants and the relative abundance of all bins (**a, c**) and *deltaproteobacterial* bins (**b, d**) with *hgcA* across the seven metagenomes. **a, b** Bottom water samples; **c, d** sediment samples. Numbers associated with each point refer to the dissolved sulfide concentration for that sample

abundance of *Deltaproteobacteria* with *hgcAB* increased with increasing k_{meth} (Fig. 4b). However, these relationships with k_{meth} were not observed in the sediments (Fig. 4d), and we further caution that these trends are each based on only three samples.

Nevertheless, potential methylators were substantially more abundant in sulfidic Lake Manganika compared to McQuade, and Manganika populations with *hgcAB* are distributed across diverse clades and functional guilds (Fig. 5). The more diverse *hgcAB* community at Manganika would seem to suggest a higher potential for MeHg production because there are more potential methylators and different potential methylators that could be active under diverse geochemical conditions. However, rates of methylation scale not only with the presence and abundance of methylating populations, but also with the bioavailability of inorganic mercury. While many factors impact the speciation of inorganic mercury and its availability for uptake and methylation (and the potential for active uptake mechanisms have not been ruled out), sulfide is known to play an important role, and high sulfide concentrations in Manganika likely reduced mercury uptake [30, 76]. Perhaps Manganika communities have a high potential for mercury methylation that can result in rapid MeHg production at those times that geochemical conditions shift to become more conducive to cellular uptake of inorganic mercury,

and perhaps this phenomenon might help explain the episodes of unusually efficient methylation that were initially observed in Manganika (e.g., [75]).

Many of the same potential methylators occur in the water column and the sediments. There are some differences—for example, methanogens were more abundant in the sediments, *Clostridia* were more abundant in the water—but this observation suggests that broadly, similar geochemical niches sustain methylating populations in both the anoxic bottom waters and sediments. MeHg production in the water column has potentially more severe consequences for downstream environments than diffusion from sediment pore waters because mixing across the thermocline or advective MeHg export from the bottom water would occur immediately following seasonal turnover.

The relative contribution of these novel potential methylators to MeHg production is not yet known. Increased MeHg production associated with sulfate-reducing microorganisms has been observed repeatedly based on sulfate amendments to field, laboratory, and pure culture experiments. In contrast, the significance of iron-reducing, methanogenic, and fermentative methylators for MeHg production in the environment is not well understood. MeHg production measured *in vitro* is variable [15, 33], making it difficult to surmise the importance of different populations for MeHg production in nature. This

Fig. 5 Relative abundance of bins with *hgcA*, classified by metabolic guild. **a** Shows the abundance of *hgcA*-containing bins relative to all metagenomic sequences, and **b** shows the same information normalized to the relative abundance of *hgcA*-containing bins

and other metagenomic studies have shown that there is substantial diversity of uncultivated potential methylators in the environment, and we do not have information on the rates and methylation capabilities of these uncultivated populations. Additional analysis of *hgcAB* sequences across a wider range of environmental conditions is needed to ascertain the relevance and persistence of these novel potential methylators under different environmental conditions and their contributions to overall rates of mercury methylation.

Acknowledgements We thank D. Engstrom, M. Kelly, and M. Berndt for insightful discussion and assistance with sample collection. Thanks to C. Nguyen, L. Thomas, the University of Minnesota College of Science and Engineering Systems staff, and the Minnesota Supercomputing Institute for computing support and the use of facilities. Thanks to H. Huang for providing mercury analytical support. Sequencing was performed at the University of Minnesota Genomics Center (UMGC). We thank the University of Minnesota Undergraduate Research Fellowship Program (UROP) for supporting G.M. W. This work was supported by University of Minnesota MnDRIVE initiative and the University of Minnesota BioTechnology Institute. The metagenomic sequencing was funded by a grant from the

University of Minnesota Duluth's Natural Resources Research Institute Permanent University Trust Fund and by an MnDRIVE-supported research grant from the University of Minnesota Duluth's Swenson College of Science and Engineering.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

- Amos HM, Jacob DJ, Streets DG, Sunderland EM. Legacy impacts of all-time anthropogenic emissions on the global mercury cycle. *Global Biogeochem Cycles*. 2013;27:410–21.
- Morel FM, Kraepiel AM, Amyot M. The chemical cycle and bioaccumulation of mercury. *Annu Rev Ecol Syst*. 1998;29:543–66.
- Obrist D, Kirk JL, Zhang L, Sunderland EM, Jiskra M, Selin NE. A review of global environmental mercury processes in response to human and natural perturbations: changes of emissions, climate, and land use. *Ambio*. 2018;47:116–40.
- Compeau G, Bartha R. Sulfate-reducing bacteria: principal methylators of mercury in anoxic estuarine sediment. *Appl Environ Microbiol*. 1985;50:498–502.
- Gilmour CC, Henry EA, Mitchell R. Sulfate stimulation of mercury methylation in freshwater sediments. *Environ Sci Technol*. 1992;26:2281–7.
- Branfireun BA, Roulet NT, Kelly CA, Rudd JW. In situ sulphate stimulation of mercury methylation in a boreal peatland: toward a link between acid rain and methylmercury contamination in remote environments. *Global Biogeochem Cycles*. 1999;13:743–50.
- Coleman Wasik JK, Mitchell CP, Engstrom DR, Swain EB, Monson BA, Balogh SJ, et al. Methylmercury declines in a boreal peatland when experimental sulfate deposition decreases. *Environ Sci Technol*. 2012;46:6663–71.
- Jeremiason JD, Engstrom DR, Swain EB, Nater EA, Johnson BM, Almendinger JE, et al. Sulfate addition increases methylmercury production in an experimental wetland. *Environ Sci Technol*. 2006;40:3800–6.
- Strickman R, Fulthorpe R, Wasik JC, Engstrom D, Mitchell C. Experimental sulfate amendment alters peatland bacterial community structure. *Sci Total Environ*. 2016;566:1289–96.
- Mitchell CP, Branfireun BA, Kolka RK. Assessing sulfate and carbon controls on net methylmercury production in peatlands: an in situ mesocosm approach. *Appl Geochem*. 2008;23:503–18.
- Fleming EJ, Mack EE, Green PG, Nelson DC. Mercury methylation from unexpected sources: molybdate-inhibited freshwater sediments and an iron-reducing bacterium. *Appl Environ Microbiol*. 2006;72:457–64.
- Kerin EJ, Gilmour C, Roden E, Suzuki M, Coates J, Mason R. Mercury methylation by dissimilatory iron-reducing bacteria. *Appl Environ Microbiol*. 2006;72:7919–21.
- Mitchell CP, Gilmour CC. Methylmercury production in a Chesapeake Bay salt marsh. *J Geophys Res Biogeosci*. 2008;113: G00C04.
- Hamelin S, Amyot M, Barkay T, Wang Y, Planas D. Methanogens: principal methylators of mercury in lake periphyton. *Environ Sci Technol*. 2011;45:7693–700.
- Gilmour CC, Podar M, Bullock AL, Graham AM, Brown SD, Somenahally AC, et al. Mercury methylation by novel

- microorganisms from new environments. *Environ Sci Technol*. 2013;47:11810–20.
16. Parks JM, Johs A, Podar M, Bridou R, Hurt RA, Smith SD, et al. The genetic basis for bacterial mercury methylation. *Science*. 2013;339:1332–5.
 17. Gionfriddo CM, Tate MT, Wick RR, Schultz MB, Zemla A, Thelen MP, et al. Microbial mercury methylation in Antarctic sea ice. *Nat Microbiol*. 2016;1:16127.
 18. Podar M, Gilmour CC, Brandt CC, Soren A, Brown SD, Crable BR, et al. Global prevalence and distribution of genes and microorganisms involved in mercury methylation. *Sci Adv*. 2015;1:e1500675.
 19. Christensen GA, Wymore AM, King AJ, Podar M, Hurt RA, Santillan EU, et al. Development and validation of broad-range qualitative and clade-specific quantitative molecular probes for assessing mercury methylation in the environment. *Appl Environ Microbiol*. 2016;82:6068–78.
 20. Liu Y-R, Yu R-Q, Zheng Y-M, He J-Z. Analysis of community structure of Hg methylation gene (*hgcA*) in paddy soils along an Hg gradient. *Appl Environ Microbiol*. 2014;80:2874–9.
 21. Su Y-B, Chang W-C, Hsi H-C, Lin C-C. Investigation of biogeochemical controls on the formation, uptake and accumulation of methylmercury in rice paddies in the vicinity of a coal-fired power plant and a municipal solid waste incinerator in Taiwan. *Chemosphere*. 2016;154:375–84.
 22. Bae H-S, Dierberg FE, Ogram A. Syntrophs dominate sequences associated with the mercury-methylating gene *hgcA* in the water conservation areas of the Florida Everglades. *Appl Environ Microbiol*. 2014;80:6517–26.
 23. Schaefer J, Kronberg RM, Morel F, Skjellberg U. Detection of a key Hg methylation gene, *hgcA*, in wetland soils. *Environ Microbiol Rep*. 2014;6:441–7.
 24. Du H, Ma M, Sun T, Dai X, Yang C, Luo F, et al. Mercury-methylating genes *dsrB* and *hgcA* in soils/sediments of the Three Gorges Reservoir. *Environ Sci Pollut Res*. 2017;24:5001–11.
 25. Bravo AG, Zopfi J, Buck M, Xu J, Bertilsson S, Schaefer JK, et al. *Geobacteraceae* are important members of mercury-methylating microbial communities of sediments impacted by waste water releases. *ISME J*. 2018;12:802.
 26. Schwartz GE, Redfern LK, Ikuma K, Gunsch CK, Ruhl LS, Vengosh A, et al. Impacts of coal ash on methylmercury production and the methylating microbial community in anaerobic sediment slurries. *Environ Sci Process Impacts*. 2016;18:1427–39.
 27. Bravo AG, Loizeau J-L, Dranguet P, Makri S, Björn E, Ungur-eanu VG, et al. Persistent Hg contamination and occurrence of Hg-methylating transcript (*hgcA*) downstream of a chlor-alkali plant in the Olt River (Romania). *Environ Sci Pollut Res*. 2016;23:10529–41.
 28. Dranguet P, Le Faucheur S, Cosio C, Slaveykova VI. Influence of chemical speciation and biofilm composition on mercury accumulation by freshwater biofilms. *Environ Sci Process Impacts*. 2017;19:38–49.
 29. Bravo AG, Bouchet S, Guédron S, Amouroux D, Dominik J, Zopfi J. High methylmercury production under ferruginous conditions in sediments impacted by sewage treatment plant discharges. *Water Res*. 2015;80:245–55.
 30. Hsu-Kim H, Kucharzyk KH, Zhang T, Deshusses MA. Mechanisms regulating mercury bioavailability for methylating microorganisms in the aquatic environment: a critical review. *Environ Sci Technol*. 2013;47:2441–56.
 31. Graham AM, Aiken GR, Gilmour CC. Dissolved organic matter enhances microbial mercury methylation under sulfidic conditions. *Environ Sci Technol*. 2012;46:2715–23.
 32. Gilmour CC, Elias DA, Kucken AM, Brown SD, Palumbo AV, Schadt CW, et al. The sulfate-reducing bacterium *Desulfovibrio desulfuricans* ND132 as a model for understanding bacterial mercury methylation. *Appl Environ Microbiol*. 2011;77:3938–51.
 33. Gilmour CC, Bullock AL, McBurney A, Podar M, Elias DA. Robust mercury methylation across diverse methanogenic *Archaea*. *mBio*. 2018;9:e02403–17.
 34. Bailey LT, Mitchell CP, Engstrom DR, Berndt ME, Wasik JKC, Johnson NW. Influence of porewater sulfide on methylmercury production and partitioning in sulfate-impacted lake sediments. *Sci Total Environ*. 2017;580:1197–204.
 35. Berndt ME, Bavin TK. Methylmercury and dissolved organic carbon relationships in a wetland-rich watershed impacted by elevated sulfate from mining. *Environ Pollut*. 2012;161:321–7.
 36. Johnson NW, Mitchell CP, Engstrom DR, Bailey LT, Wasik JKC, Berndt ME. Methylmercury production in a chronically sulfate-impacted sub-boreal wetland. *Environ Sci Process Impacts*. 2016;18:725–34.
 37. Graham EB, Gabor RS, Schooler S, McKnight DM, Nemerugut DR, Knelman JE. Oligotrophic wetland sediments susceptible to shifts in microbiomes and mercury cycling with dissolved organic matter addition. *PeerJ*. 2018;6:e4575.
 38. Hintelmann H, Keppel-Jones K, Evans RD. Constants of mercury methylation and demethylation rates in sediments and comparison of tracer and ambient mercury availability. *Environ Toxicol Chem*. 2000;19:2204–11.
 39. Eaton AD, Clesceri LS, Rice EW, Greenberg AE, Franson MAH. Standard methods for the examination of water and wastewater. Centennial edition. Twenty-first edition. Washington, D.C., USA: American Public Health Association; 2005.
 40. Bailey L, Johnson N, Mitchell C, Engstrom D, Coleman-Wasik J, Kelly M, et al. Seasonal and spatial variations in methylmercury in the water column of sulfate-impacted lakes: a Report to Iron Ore Cooperative Research Program. St. Paul, MN: Minnesota Department of Natural Resources Division of Lands and Minerals; 2014. 64 p.
 41. Gohl DM, Vangay P, Garbe J, MacLean A, Hauge A, Becker A, et al. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat Biotechnol*. 2016;34:942.
 42. Jones DS, Lapakko KA, Wenz ZJ, Olson MC, Roepke EW, Sadowsky MJ, et al. Novel microbial assemblages dominate weathered sulfide-bearing rock from copper-nickel deposits in the Duluth Complex, Minnesota, USA. *Appl Environ Microbiol*. 2017;83:e00909–17.
 43. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res*. 2007;35:7188–96.
 44. R Core Development Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2016.
 45. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. *vegan*: Community Ecology Package. R package version 2.4-0 edn (2016).
 46. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Yadhukumar, et al. ARB: a software environment for sequence data. *Nucleic Acids Res*. 2004;32:1363–71.
 47. Huber T, Faulkner G, Hugenholtz P. Bellerophon; a program to detect chimeric sequences in multiple sequence alignments. *Bioinformatics*. 2004;20:2317–9.
 48. Joshi N, Fass J. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software] (2011).
 49. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17:10–2.
 50. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics*. 2011;27:863–4.

51. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 1997;25:3389–402.
52. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 2017;27:824–34.
53. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9:357.
54. Wu Y-W, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics.* 2015;32:605–7.
55. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30:2068–9.
56. Smith SD, Bridou R, Johs A, Parks JM, Elias DA, Hurt RA, et al. Site-directed mutagenesis of HgcA and HgcB reveals amino acid residues important for mercury methylation. *Appl Environ Microbiol.* 2015;81:3205–17.
57. Krogh A, Larsson B, Von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol.* 2001;305:567–80.
58. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ.* 2015;3:e1319.
59. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 2015;25:1043–55.
60. Chen I-MA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, et al. IMG/M: integrated genome and metagenome comparative data analysis system. *Nucleic Acids Res.* 2017;45: D507–16.
61. Wallace IM, O'sullivan O, Higgins DG, Notredame C. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* 2006;34:1692–9.
62. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics.* 2006;22:2688–90.
63. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol.* 2008;25:1307–20.
64. Darriba D, Taboada GL, Doallo R, Posada D. ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics.* 2011;27:1164–5.
65. Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol.* 2003;52:696–704.
66. Berger SA, Krompass D, Stamatakis A. Performance, accuracy, and web server for evolutionary placement of short sequence reads under maximum likelihood. *Syst Biol.* 2011;60(3):291–302.
67. Jones DT, Taylor WR, Thornton JM. The rapid generation of mutation data matrices from protein sequences. *Bioinformatics.* 1992;8:275–82.
68. Phylip FJ. Phylogeny Interference Package. *Cladistics.* 1989;5:164–6.
69. Orlygsson J, Kristjansson JK. The Family *Hydrogenophilaceae*. In: Rosenberg E, DeLong E, Lory S, Stackebrandt E, Thompson F, editors. *The Prokaryotes—Alphaproteobacteria and Betaproteobacteria*. 4th ed. Berlin: Springer; 2014. p. 859–68.
70. Liu Y, Balkwill DL, Aldrich HC, Drake GR, Boone DR. Characterization of the anaerobic propionate-degrading syntrophs *Smithella propionica* gen. nov., sp. nov. and *Syntrophobacter wolnii*. *Int J Syst Evol Microbiol.* 1999;49:545–56.
71. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature.* 2013;499:431.
72. Spring S, Bunk B, Spröer C, Schumann P, Rohde M, Tindall BJ, et al. Characterization of the first cultured representative of *Verucomicrobia* subdivision 5 indicates the proposal of a novel phylum. *ISME J.* 2016;10:2801–16.
73. Anantharaman K, Brown CT, Hug LA, Sharon I, Castelle CJ, Probst AJ, et al. Thousands of microbial genomes shed light on interconnected biogeochemical processes in an aquifer system. *Nat Commun.* 2016;7:13219.
74. Picardeau M. The Family *Leptospiraceae*. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F. (eds.). *The Prokaryotes*. (Springer, Berlin, Heidelberg, 2014). p. 711–29.
75. Berndt M, Bavin T. Sulfate and mercury cycling in five wetlands and a lake receiving sulfate from taconite mines in Northeastern Minnesota: a Report to Iron Ore Cooperative Research Program. St. Paul, MN: Minnesota Department of Natural Resources Division of Lands and Minerals; 2011. 77 p.
76. Benoit JM, Gilmour CC, Mason RP, Heyes A. Sulfide controls on mercury speciation and bioavailability to methylating bacteria in sediment pore waters. *Environ Sci Technol.* 1999;33: 951–7.