2021

Metagenomic Identification and Classification of the Mercury-Methylating Gene hgcA in Response to Water Table and Plant Functional Group Manipulations in Peat Soil

Madeline Peterson
Michigan Technological University, mrpeter1@mtu.edu

Recommended Citation

Peterson, Madeline, "Metagenomic Identification and Classification of the Mercury-Methylating Gene hgcA in Response to Water Table and Plant Functional Group Manipulations in Peat Soil", Open Access Master's Thesis, Michigan Technological University, 2021.
https://doi.org/10.37099/mtu.dc.etdr/1238

METAGENOMIC IDENTIFICATION AND CLASSIFICATION OF THE MERCURY-METHYLATING GENE *hgcA* IN RESPONSE TO WATER TABLE AND PLANT FUNCTIONAL GROUP MANIPULATIONS IN PEAT SOIL

By

Madeline R. Peterson

A THESIS

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In Applied Ecology

MICHIGAN TECHNOLOGICAL UNIVERSITY

2021

This thesis has been approved in partial fulfillment of the requirements for the Degree of MASTER OF SCIENCE in Applied Ecology.

College of Forest Resources and Environmental Sciences

Thesis Co-Advisor:	*Evan Kane*
Thesis Co-Advisor:	*Erik Lilleskov*
Committee Member:	*Stephen Techtmann*
College Dean:	*Andrew Storer*

Table of Contents

Author Contribution Statement

This body of work includes multi-authored papers in various stages of publishing. Details regarding publish status, copyright, and author contributions are detailed below.

Peatland microbial community responses to plant functional group and drought are depth-dependent, has been accepted for publication in Molecular Ecology. The manuscript ID is ID_MEC201469.

Evan Kane, Erik Lilleskov, Rod Chimner, Randy Kolka, and Tom Pypker designed and developed the experimental system PEATcosm.

Co-author contributions: Madeline Peterson, Evan Kane, and Erik Lilleskov conceived and designed this study. Kristine Haynes, Evan Kane, Erik Lilleskov, Lynette Potvin, and Carl Mitchell contributed to previous MeHg research utilized in this project.

Evan Kane and Karl Romanowicz collected peat soil samples. Jamie Lamit processed and extracted DNA from samples. DNA sequencing was provided by the Joint Genome Institute (JGI), with special help from Tijana Glavina del Rio. John Stanovick provided advice on statistical analysis of data. Madeline Peterson analyzed the data, created figures, and wrote the manuscript; committee members provided statistical and editorial advice.

Acknowledgements

I would like to thank immensely both of my advisors, Dr. Evan Kane and Dr. Erik Lilleskov, for this opportunity, as well as their constant guidance and support throughout this project. I would also like to thank my committee member, Dr. Stephen Techtmann, for his time and patience in helping me resolve the technical difficulties that come with coding and bioinformatics. None of this work would have been possible without the generous help and advice from this committee.

I'd like to thank my family that assisted with my move to Michigan and for helping me make this place a home as I completed this research.

To all my friends I've made along the way, I'd like to thank you all for your loving support that helped me complete this project from beginning to end– here's to you all!

Lastly, I'd like to thank my road dog, Seth, for being there for me and encouraging me through the last couple of years.

List of Abbreviations

WT water table

PFG plant functional group

MeHg methyl-mercury

IRB iron-reducing bacteria

SRB sulfur-reducing bacteria

redox reduction-oxidation

HMM hidden Markov model

Abstract

Methyl-mercury (MeHg) is a potent neurotoxin that threatens the environment and the health of humans and wildlife alike. Exposure to the bioaccumulating toxin will become more prevalent as the climate continues to warm and alter northern ecosystems. In addition to storing a large portion of the planet's carbon, boreal peatlands also act as reservoirs of atmospherically deposited inorganic mercury that can be converted into methyl-mercury (MeHg). The mercury-methylating genes responsible for this activity, the obligatory gene pair *hgcA* and *hgcB*, are currently the only genes recognized as a requirement for the process of anoxic mercury methylation. Mercury deposition paired with anoxic conditions creates perfect environments for anaerobic prokaryotes to methylate mercury. Peatlands are ranked among the leading hotspots for such activity, yet little is known about the community composition or functional relationship of mercury-methylating microbes in response to varying environmental conditions. In the PEATcosm experiment, water tables and plant functional groups were manipulated to determine their impact on peatland biogeochemistry. Metagenomic data from this project was obtained to examine the effects of treatment variables on the abundance and functional composition of *hgcA*-containing organisms and to contextualize these findings within another larger dataset. We hypothesized that *hgcA* occurrence would be in accordance with measured methyl-mercury concentrations among treatments, with a predicted higher abundance of *hgcA* genes found in lowered water table and/or sedge treatments as compared to raised water table and/or shrub treatments. We found significant effects of water table, plant functional group, and depth on *hgcA* gene abundance, with evidence for increased abundance under high water table treatments– especially at the 30 cm depth. Methanogens dominated as the most abundant functional assignment for gene hits that were recognized, suggesting that methanogens have potential to be leading mercury methylators in some systems.

1 Introduction

Peatlands are environments defined by saturated, anoxic conditions that support large volumes of stored soil carbon and unique communities of anaerobic microbiota, some of which convert inorganic forms of mercury (Hg) into methylated mercury (Podar et al. 2015; Schaeffer et al. 2020). The potential for mercury methylation is present in a diversity of organisms and environments (Podar et al.2015, McDaniel et al. 2020, Villar et al. 2020), including peatlands, which are hotspots for methyl-mercury (MeHg) production (Haynes et al. 2019, Hu et al. 2020, Schaeffer et al. 2020). Inorganic mercury is naturally present in all environments at low concentrations via atmospheric deposition (Engle et al. 2010; Schuster et al. 2002) which, when methylated, can be carried to other ecosystems via water transport through hydrologically connected systems (Gordon et al. 2016; Mitchell et al. 2008, Schuster 2011) where it will have the chance to further bioaccumulate. As a bio-accumulating potent neurotoxin, MeHg threatens a variety of ecosystems that could negatively impact the health of both humans and wildlife (Watanabe & Satho 1996; Trasande et al. 2005; Evers 2018). In addition, climate change threatens to shift regional precipitation and climate regimes (Loisel et al. 2020) which could destabilize northern peatlands. These changes could include altered water table height and the indirect shift of plant functional groups. While peatland destabilization risks elevated greenhouse gas emissions through carbon degradation, it could also mean a further increase in Hg release and methyl-mercury production in peatlands (Haynes et al. 2017b, 2019).

Boreal wetlands are widely distributed, storing an estimated 30% of the planet's total carbon stores (Gorham 1991; Dixon 1994). Their historic stability and low rates of decomposition allow these wetlands to act as storage for atmospheric inputs such as carbon and Hg. Over the past several decades, long-range transport of anthropogenic Hg in the northern hemisphere has expanded, leading to increased levels of Hg deposition and elevated concentrations of Hg in wetlands soils (Fitzgerald et al. 1998; Grigal 2002). The paired effects of climactic change and increased Hg deposition in peatlands will likely push peatlands, once atmospheric carbon and Hg sinks, to become sources through

the activities of peat fire and peat degradation (Turetsky et al. 2006; Haynes 2019). Furthermore, climate warming trends show acceleration of inorganic Hg release from melting permafrost landscapes (Schuster et al. 2011), increasing the supply of available inorganic Hg for conversion to MeHg in downstream environments. For these reasons, understanding the identities and functional links of mercury-methylating communities in relation to altered environmental conditions is central to understanding methyl-mercury production and accumulation in the environment.

All microbes that contain the gene pair *hgcAB* are currently believed to be capable of methylating Hg, as has been shown in culture by Graham et al. 2013. So far, only anaerobic microbes have been found to produce MeHg via the gene pair *hgcAB* (Parks et al. 2013; Ma et al. 2019). A very limited number of studies have suggested aerobic Hg-methylation via alternative pathways (Cao et al. 2021), but these pathways currently lack support, with no alternative genetic pathways identified to date.

Since the discovery of the gene pair *hgcAB* and its necessity for mercury-methylation (Parks et al. 2013), numerous studies have examined the diversity, global prevalence, and individual abilities of microbes to methylate mercury. Mercury-methylation has been found within methanogenic, fermentative, acetogenic, and cellulolytic microbes (Gilmour et al. 2013). The biogeochemical setting, redox conditions, and availability of quality carbon and nutrients determine the community of microbes that can exist within a local environment, while further competition amongst microbes is instigated by environmental settings that determine which microbial groups dominate a niche.

It is possible that microbial species methylate Hg at differing rates and produce varying concentrations of MeHg compared to one another (Graham et al. 2012; Gilmour et al. 2013); but measurement comparisons can be difficult and lacking precision due to differences in growth rates, culture densities, and growth medium chemistries. To make predictions on MeHg production and rates, we must first understand what organisms are present and participating in peat systems. Ideally, we would like to understand the

community effects of environmental conditions on Hg-methylating microbes in order to make informed predictions of MeHg in peatlands. In another study of peatland Hg-methylators, iron- and sulfate-reducing bacteria (IRB and SRB respectively) were found to dominate in younger or richer peatlands, while as peatlands age they transition to host higher proportions of syntrophic and methanogenic microbiota (Hu et al. 2020). This transition effect on the functional community reflects the influence of longer time-scales on peat characteristics, but those age classifications also include the influences of local effects such as plant community composition and the soil redox environment.

We would predict that at smaller time-scales, changes in precipitation and subsequent changes in water table position would shape the microbial community (Lamit et al. accepted). Water table (WT) has shown to be a more powerful predictor of Hg concentration than plant functional group, as vertical movement of the WT can aerate and expose previously submerged peat to O_2 (Haynes et al. 2017a), altering peat soil chemistry and microbial activities. The highest concentration of MeHg in the full-factorial peatland mesocosm experiment PEATcosm was found at the 30-40 cm mark in lowered water table treatments and was in-line with the position of the drawn-down water table (Haynes et al. 2019). This pool of MeHg could be partially explained by oxidative recharge of the local environment and availability of electron acceptors for syntrophic organisms (Hu et al. 2020), which allowed the sequential activity of anaerobic microbial groups with increasing distance from these more aerated zones (Agethen et al. 2018). This pattern of shifting microbial function to a "hotspot" depth below the moving WT boundary also aligns to the largest shift in the overall microbial community in response to WT alterations (Lamit et al. accepted), showing that these effects on microbe abundance are greatest at newly formed aerobic/anaerobic boundaries.

Though not as powerful a predictor as WT, PFG influences still have effects on microbial community composition and function, especially when analyzed concurrently with WT effects. Additional influences on increased overall microbial activity include responses to carbon, oxygen, and nutrient exudates released from sedge roots, which can then have an

indirect effect on anaerobic processes. Aerenchymas sedge roots create an environment adjacent to the roots that mimic the sharp oxic/anoxic barrier of the WT surface. Lowered WT conditions and transport of atmospheric oxygen by sedge aerenchyma allows for oxidative "recharge" of electron acceptors, including organic matter (OM). Organic matter is a terminal electron acceptor that mediates anaerobic metabolism under a range of redox conditions, which includes supporting the respiratory processes for IRB, SRB, fermenters, and methanogens (Klupfel 2014). The availability of Hg to microbes appears to be dependent on both the existing relationship between Hg(II) and its carbon complex, but also the availability of usable carbon substrates required for metabolism (Mazrui et al. 2016), thus anaerobic mercury-methylating microbes can benefit from the by-products produced from other microbes. Decomposition of OM in response to aeration is also expected to increase mobility of both C substrates and inorganic Hg from bulk peat to the pore water phase (Haynes et al. 2019; Martínez et al. 2007), resulting in more accessible and/or more abundant Hg in porewater where microbes would then be exposed to elevated Hg and methylate Hg at a higher rate.

We used metagenomics to identify the mercury-methylating gene, *hgcA,* in peat soils as it was related to experimental treatments within PEATcosm (Peatland Climate Change Experiment at the Houghton Mesocosm). PEATcosm was a full-factorial peatland mesocosm experiment that manipulated aboveground plant community composition and water table levels (Potvin et al. 2015). During the course of this experiment, additional data were collected that were yet to be analyzed. The available metagenomic data from PEATcosm were put into context within a larger global *hgcAB*-containing community as has been defined by previous metagenomic work (McDaniel et al. 2020). In addition to a bioinformatics approach, further analysis of these findings combined additional soil and porewater Hg concentration data to identify patterns and correlations between methylation rates, concentrations, and experimental treatment effects.

We tested the following hypotheses:

H1: *We assumed there would be a higher relative abundance of hgcA genes in lowered WT and sedge treatments as compared to high WT and ericaceous shrub treatments. Within this hypothesis we would also expect high WT and ericaceous treatments to possess a lower relative abundance of hgcA genes.*

Justification for H1: Prior data suggested water table and above-ground plant communities shaped total mercury (THg) and methyl-mercury concentrations in both peat and porewater in this system (Haynes et al. 2019). The findings of Haynes et al. 2019 found elevated levels of MeHg in the porewater of low water table manipulations of PEATcosm, as well as elevated levels of MeHg in the peat fractions of sedge treatment manipulations.

H2: *We predicted that SRB and IRB would have the highest methylating potential in lowered WT and sedge treatments.*

Justification for H2: Methylating potential within the context of this paper is defined as the predicted ability of a microbial functional group (e.g., SRB, IRB, methanogens, fermenters) to methylate available Hg, as measured by the abundance of *hgcA* genes in the metagenomic data. Within the PEATcosm system, we assumed sedge treatments would exhibit characteristics more like a younger peatland (as classified in Hu et al. 2020) in that they would host a higher proportion of SRB and IRB. This is assumed to be due in part through increased decomposition of organic matter due to aeration by lowered WT and sedge root aerenchyma, increased availability of electron acceptors, increased C sources, and inorganic Hg should then allow for elevated rates of Hg methylation under these treatments.

H3: *We predict that methanogen-associated hgcA genes will have a larger proportional contribution to overall hgcA abundance in high WT treatments and ericaceous treatments compared to the other treatments.*

Justification for H3: W predicted that systems containing ericaceous shrubs would behave more like that of peatland classified as older in that it would support higher proportion of methanogen-associated Hg-methylators. High water tables and ericaceous shrub treatments promote an increasingly anoxic environment that is more supportive of anaerobic microbes. Evidence of this pattern for methanogens in PEATcosm is supported by the work of Lamit et al. accepted, which found a positive shift by methanogens in high water table treatments.

H4: *We hypothesized that Hg methylation would be reflected in the metagenomic data where highest measured methylation rates should align with higher hgcA gene abundance.*

2 Materials and Methods

2.1 Mesocosm setup and experimental design

All samples for this research were part of the data collection in support of PEATcosm, a multifactorial experiment that examined the effects of plant community composition and water table depth on peatland microbial processes, highlighting the potential impacts of climate change. The project consisted of 24 mesocosm chambers measuring $1m^3$ located at the Forest Service Northern Research Station Experimental in Houghton, Michigan. Intact peat blocks were extracted in 2010 from a Minnesota peatland and allowed to stabilize for 1 year before manipulation experiments began in 2011. Experimental manipulation treatments consisted of water table (WT) maintained at two heights (low and high), and 3 plant functional group (PFG) treatments (unmanipulated, sedge-only, and Ericaceae-only). For this research, soil samples were available only for the ericoid- and sedge-only treatments. The sedge and ericoid manipulations were obtained by actively clipping out sedges or Ericaceae from the target treatments. A full description of experimental design and manipulation methods is described in depth in Potvin et al. 2015.

2.2 Chemical and physical soil properties

Methyl-mercury measurements used in this research were taken the summer of 2014, the final year of experimental treatment in PEATcosm. Methyl-mercury concentrations were measured in both solid phase peat and porewater from the peatland mesocosm bins. For the research described in this paper, MeHg data that were sampled most closely in time to the date of soil sample collection for metagenomics were utilized. Peat fraction Hg and MeHg data was collected in mid-July. Methyl-mercury concentrations that were collected at the same depth as those in the soil sample collection were utilized, at the 10-20 cm and 30-40 cm depths. Entries for the 60-70 cm peat MeHg data were absent and were supplemented with data from the 50-60 cm depth for those depths only. Porewater Hg and MeHg data were collected at the three specified depths in mid-July of 2014. A more

detailed description of data collection can be found in Haynes et al. 2017b and Haynes et al. 2019.

2.3 DNA extraction, sequencing, and data processing

In late July of 2014, the final full year of experimental treatment, a subset of bins were selected that consisted of sedge-only and Ericaceae-only mesocosm bins factorial with both WT treatments, with 2 replicates each for a total of 8 bins. Sub-samples of each soil were collected at the target depths of 10-20cm, 30-40cm, and 60-70cm (for a total of 24 metagenomes), hereby referred to as the 10cm, 30cm, and 60cm depth respectively. Immediately following extraction, peat cores were flash-frozen in liquid nitrogen and stored frozen at -80°C until processing. Sub-samples of soil samples were later thawed and processed for genomic analysis.

Soil samples were homogenized and ground using mortar and pestle followed by further pulverization by a clean coffee grinder. DNA was extracted from 0.5 g of ground peat from each sample using an RNA PowerSoil Total RNA Isolation Kit with a DNA co-elution accessory Kit, followed by cleaning with a PowerClean DNA Clean-Up kit (MoBio Laboratories; now Qiagen, Germantown, MD, USA). Cleaned DNA was quantified with a Qubit Fluorometer (Invitrogen, Life Technologies, Carlsbad, CA, USA).

2.4 Identification of *hgcA* genes from assembled genomic data

All extracted soil samples were sent to and processed by the Joint Genome Institute (JGI). Shotgun metagenomics were performed for all 24 samples. Plate-based DNA library preparation for Illumina sequencing was performed on the PerkinElmer Sciclone NGS robotic liquid handling system using Kapa Biosystems library preparation kit. 200 ng of sample DNA was sheared to 300 bp using a Covaris LE220 focused-ultrasonicator. The sheared DNA fragments were size selected by double-SPRI and then the selected fragments were end-repaired, A-tailed, and ligated with Illumina compatible sequencing adaptors from IDT containing a unique molecular index barcode for each sample library.

The prepared libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit and run on a Roche LightCycler 480 real-time PCR instrument. The quantified libraries were then prepared 7 for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v4. Sequencing of the flowcell was performed on the Illumina HiSeq 2500 sequencer using HiSeq TruSeq SBS sequencing kits, v4, following a 2x150 indexed run recipe

Post-processing, resulting read data became accessible through the JGI metagenome portal, where the files with the translated gene sequences (.faa files) were individually downloaded for each sample. All genomic .faa files were concatenated and converted to single-line format. McDaniel et al. 2020 provided an expanded range of globally diverse *hgcAB* gene clusters from available metagenomic data. We therefore utilized this data in an approach to identify our sequences against this broad collection of *hgcAB* sequences. The available hidden markov model (HMM) code introduced in McDaniel et al. 2020 was utilized to conservatively identify and compile *hgcA* sequences from each of our sample .faa files.

Running the HMM on individual soil samples versus running it on the concatenated list of all files yielded a slightly different number of resulting gene hits. To maximize the number of total *hgcA* gene hits, the outputs of the two methods were combined resulting in the 181 unique genomic hits. Sequence IDs for the *hgcA* gene hits were then used to search concatenated .faa files to extract a list of corresponding sequences for the gene hits. The resulting .faa file contained only the sequences of the *hgcA* gene hits. The resulting output from the HMMs produced tables of sequence IDs matched to *hgcA* gene hits. Gene sequences compiled in McDaniel et al. 2020 were downloaded and renamed MCD001 through MCD904. All unique *hgcA* gene hits from the present study were renamed PEAT001 through PEAT181.

Not all contigs were long enough to contain both the A and the B gene, and only 70 paired *hgcB* genes were identified out of the 181 *hgcA* gene hits. Under the assumption

that an A gene would be co-occur with the complementary B gene, just the *hgcA* gene was used for continued analysis, as has been done in other studies with moderate reliability (Christensen et al. 2019; Bae et al. 2014).

2.5 Clustering gene hit sequences into OTUs and normalizing gene count abundance

The combined list of 181 PEAT and 904 MCD *hgcA* sequences were clustered at 80% identity similarity (Liu, Yu, & He 2018), which allowed for integration of the McDaniel and PEAT sequences and assisted in identification of the PEAT sequences. Clustering of all PEAT and MCD sequences resulted in 591 clusters. Of the 591 clusters resulting from the combined MCD and PEAT *hgcA* sequences, 58 of them contained at least one PEAT genomic *hgcA* sequence. The 58 clusters containing all 181 unique genomic PEAT sequences were utilized for further analysis.

To determine the abundance of the *hgcA* genes in the metagenomic datasets, coverage files for all PEAT *hgcA* scaffolds were downloaded via JGI. Using the DESeq2 package in R (Love et al. 2014), all 181 *hgcA* gene coverage values were normalized using the median of ratios method per individual sample using the coverage data. Scaffolds were matched to normalized count data manually using the contig portion of the *hgcA* gene hit ID as a reference to locate the correct scaffolds.

2.6 Taxonomic identification of *hgcA*-containing sequences

To determine the most conservative taxonomic assignments for the PEAT *hgcA* sequences, a set of rules were followed to select representative taxonomy for each cluster and PEAT sequence. Hierarchical rules for selection of representative taxonomy for each cluster were as follows: A cluster that contained a sequence derived from McDaniel et al. 2020 was assigned the taxonomy given to the McDaniel-assigned sequence with the longest length and highest taxonomic resolution; Clusters that contained only PEAT sequences were assigned to the JGI-assigned taxonomy of the longest *hgcA* sequence and/or the sequence with the most genes on the original contig (these were typically the

same sequence). Taxonomic assignment of *hgcA*-containing PEAT scaffolds were determined according to the JGI methods for taxonomic assignment of metagenomic scaffolds. In brief, scaffolds were assigned to a taxonomy if more than 50% of the genes on the scaffold were assigned to the same taxon.

2.7 Functional assignment and analysis of *hgcA* gene hits

The functional assignment program FAPROTAX (Louca et al. 2016) was used to understand and group the identified *hgcA* gene clusters in our dataset into functional categories. Following taxonomic identification of the gene hits, an OTU table was prepared as input for FAPROTAX (Louca et al. 2016) using the taxonomies of assigned gene hits/clusters and the normalized *hgcA* gene counts. Zero values were supplemented in samples where there were no gene hits identified. Genes with taxonomic names that were recognized in the FAPROTAX database were assigned to a functional classification. For the purposes of this research, only functional assignments listed under potential methanogens, SRB, IRB, or fermenters were utilized for analytical purposes. These functional groups were selected as they pertained to our original hypotheses and because these groups have been listed as having members that contain the *hgcA* gene and are therefore potential methylators. Roughly a third (37.7%) of all gene counts were assigned to as least one functional group, with only around 3 of those gene counts accounting for a functional assignment outside of the four listed previously. Functional assignments for the OTU clusters were matched back to their original samples and in accordance with corresponding gene abundances.

2.8 Constructing the *hgcA* phylogenetic tree

All 181 PEAT and 904 MCD *hgcA* sequences were used as input for the sequence alignment program MUSCLE (Edgar 2004). Model tests were performed using ModelTest-NG Darriba et al. 2020) to select the best model for tree building. The program FastTree (Price et al. 2009) was used to create the phylogenetic tree using the model LG (Le & Gascuel 2008) with the CAT approximation for evolution rate followed

by optimization of branch length rescaling. Resulting alignment files were used to construct phylogenetic trees for *hgcA* in iTOL (Letunic&Bork, 2021).

2.9 Statistical Analysis

An overall analysis of total abundance data by depth and treatments was performed with a mixed model ANOVA in JMP (JMP® Pro 15.2.0). Parameters for this analysis were set for repeated effects covariance with the data structure defined by a transformation of log10(n+1). The parameters for the ANOVA contained the class variables of bin, depth, water table (WT), and plant functional group (PFG), with WT and PFG nested within the random effect of bin. Post hoc analysis using Tukey's test analyzed data for contrasts between depth and treatment combinations.

Analyses for treatment effects on functional assignment and taxonomy derived from the *hgcA* gene reads were performed using Primer v.6 (Clarke&Gorley, 2006) PERMANOVA+ (Anderson, Clarke, &Gorely 2008) for permutational multivariate analysis of variance and visualized using non-metric multidimensional scaling (NMDS). PERMANOVA utilized normalized *hgcA* gene count data, WT, PFG, depth, and all interactions. Bin was set as a random effect nested within WT and PFG. A NMDS analysis using the best 2-D configuration (stress= 0.06) was created to visualize both functional and taxonomic distribution of *hgcA* read hits, with parameters for standardization of samples by total, a square root transformation of normalized *hgcA* gene count data, and resemblance using Bray Curtis similarity.

Analysis of the relationships between measured concentrations of total mercury (THg), methyl-mercury (MeHg), percent methyl mercury (%MeHg) and *hgcA* gene count were performed using the least square method.

2.10 Data availability

JGI accession numbers, data, and working code is available at https://osf.io/pfesn/

3 Results

We found that depth of sample in addition to the experimental treatments adjusting water table (WT) and plant functional group (PFG) had significant effects on *hgcA* gene abundance in peat soil. Multiple permutations of PERMANOVA and ANOVA analyses agreed, showing the singular effects of depth, WT, PFG and the combined effects of depth*WT and the three-way interaction of depth*WT*PFG to be significant in explaining experimental and depth effect differences on *hgcA* gene abundance. Construction of the phylogenetic tree showed closely grouped branches of PEAT *hgcA* sequences clustered with MCD sequences, most clearly along branches classified as Firmicutes, Geobacteraceae, Methanoregula, and Syntrophobacteraceae. We also found weak relationships between select-measured methyl-mercury concentrations and *hgcA* gene abundance, especially at the 30 cm depth, showcasing the variability of the environment at this level when environmental conditions are changed.

3.1 Output of the HMM for identification of *hgcA* genes

The total number of metagenomic *hgcA* gene hits identified by the HMM on individual sample .faa files produced 174 genomic gene hit results, while running the HMM on the concatenated metagenomic .faa file produced 177 metagenomic hits. Comparison of the metagenomic outputs revealed both output lists shared the 170 genomic gene hits with the additional 7 gene hits being unique to the concatenated output and 4 being unique to the combined HMM runs done on individual metagenomic .faa samples, for a total of 181 unique metagenomic gene sequences and 1737 total *hgcA* gene counts when normalized for each sample. Clustering of all 181 unique genomic PEAT *hgcA* hits resulted in 58 unique OTUs.

3.2 Treatment effects shape *hgcA* gene abundance

We hypothesized (H1) a higher relative abundance of *hgcA* gene hits from lowered WT and sedge treatments compared to high water table and ericaceous treatments, which was not supported. Depth, WT, and depth*WT were the most significant predictors of *hgcA* gene count, followed by significant PFG effects and a three-way interaction effect of

depth*WT*PFG (Table 1). The strong depth effects were dependent on the PFG and WT treatments (Fig. 1). Specifically, at the 30 cm depth we saw the highest *hgcA* gene abundance, but only in the high WT treatments.

Table 1. Full-factorial ANOVA for treatment effects on *hgcA* gene count.

Source	DFNum	DFDen	F Ratio	Prob > F
Depth	**2**	**8**	**55.79476**	**<.0001**
WT	**1**	**4**	**152.6986**	**0.0002**
Depth*WT	**2**	**8**	**21.34051**	**0.0006**
PFG	**1**	**4**	**15.49459**	**0.017**
Depth*PFG	2	8	2.167935	0.1769
WT*PFG	1	4	2.310449	0.2031
Depth*WT*PFG	**2**	**8**	**5.336516**	**0.0337**

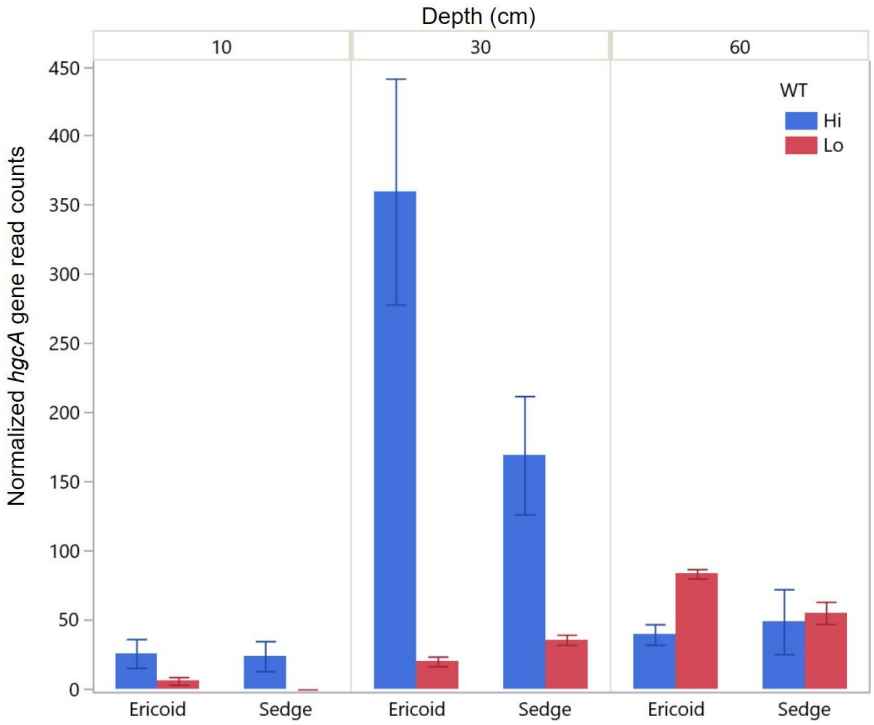

Figure 1. Normalized (untransformed) *hgcA* read counts sorted by treatment. LS means from ANOVA analysis. Each error bar is constructed using 1 standard error from the mean Bars with the same lettering are not statistically different.

3.3 Methanogens as potentially important Hg methylators

Of the 58 genomic PEAT clusters, 34 received a higher resolution assignment below the taxonomic level of order after sequence clustering described in section 2.8. The resulting output from FAPROTAX identified functional assignments for 44 out of the total 181 unique genomic *hgcA* hits. Of the 58 genomic clusters, 20 received at least one functional assignment via FAPROTAX, which is equivalent to 44 of the total 181 unique genomic *hgcA* hits, or 37.7% of the normalized gene counts. The remaining 137 unique hits, representing 62.3% of normalized sequence reads, were left unassigned (supplemental Fig. S1). No *hgcA* genes were identified for two low WT sedge 10 cm depth samples. Both treatments (WT and PFG), depth, and the interactive effects of WT and PFG with depth were significant in shaping the functional communities of potential methylators (Table 2).

Methanogens were the most abundant functionally-assigned group, especially at the 30 cm depth under Ericaceae treatment, but only under high WT conditions (Fig. 2). Other functional group assignments, IRB, SRB, and fermenters, did not contain enough data to show distribution patterns influenced by treatments/depth, and were dwarfed in comparison to methanogen assignments and the unassigned.

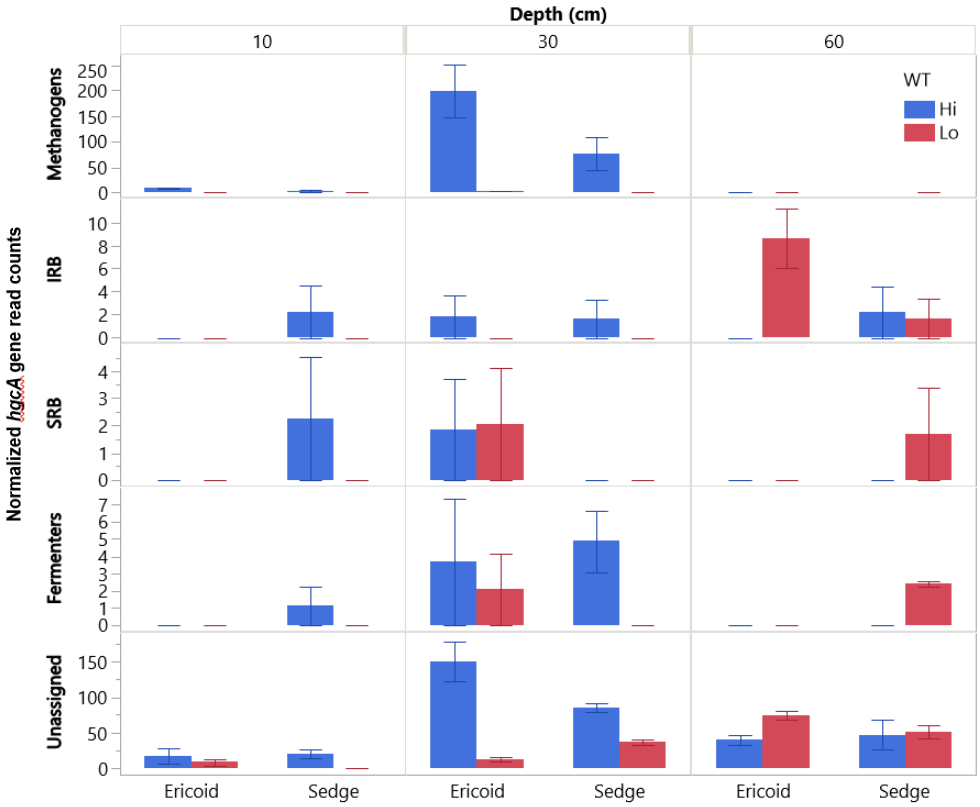

Figure 2. Normalized gene count by functional assignments as they relate to the treatment combinations.

Contrary to our prediction (H2) that there would be a greater number of gene counts of SRB and IRB in lowered water table and sedge treatments as compared to other functional assignments, we found little supporting data that these groups contribute much to *hgcA* gene abundance. However, in support of H3, the relative number of assignments suggest that methanogens are a large part of the functional community (33.65% of normalized gene hits) in these high WT peatland systems (Fig. 2, Fig. 3) compared to SRB and IRB (1.4%–2.1% and >1.14% respectively). The first axis of the functional NMDS (Fig. 3) is driven more strongly by WT and depth, as shown by the left and right separation of WT treatments, gradient of 10 cm depth on the left to 60 cm on the right. Our results here must be considered provisional, given the proportion of unassigned functional taxa, as can be seen as a clustering of overlapping and uninformative data points shown to the right in Fig. 3.

Table 2. PERMANOVA results of treatment effects on read counts for functional groups.

Source	df	SS	MS	Pseudo-F	P(perm)	Unique perms
WT	**1**	**2460.5**	**2460.5**	**21.096**	**0.0191**	**1229**
PFG	**1**	**1393.7**	**1393.7**	**11.95**	**0.018**	**1227**
Depth	**2**	**4296.9**	**2148.4**	**14.813**	**0.0004**	**9951**
WTxPFG	*1*	*504.7*	*504.7*	*4.3272*	*0.0614*	*1228*
WTxDepth	**2**	**2129.8**	**1064.9**	**7.3425**	**0.0016**	**9958**
PFGxDepth	**2**	**867.49**	**433.74**	**2.9906**	**0.0361**	**9962**
Bin(WTxPFG)	4	423.93	105.98	0.73075	0.6985	9945
WTxPFGxDepth	*1*	*419.86*	*419.86*	*2.8949*	*0.0743*	*9969*
Res	7	1015.2	145.03			
Total	21	11393				

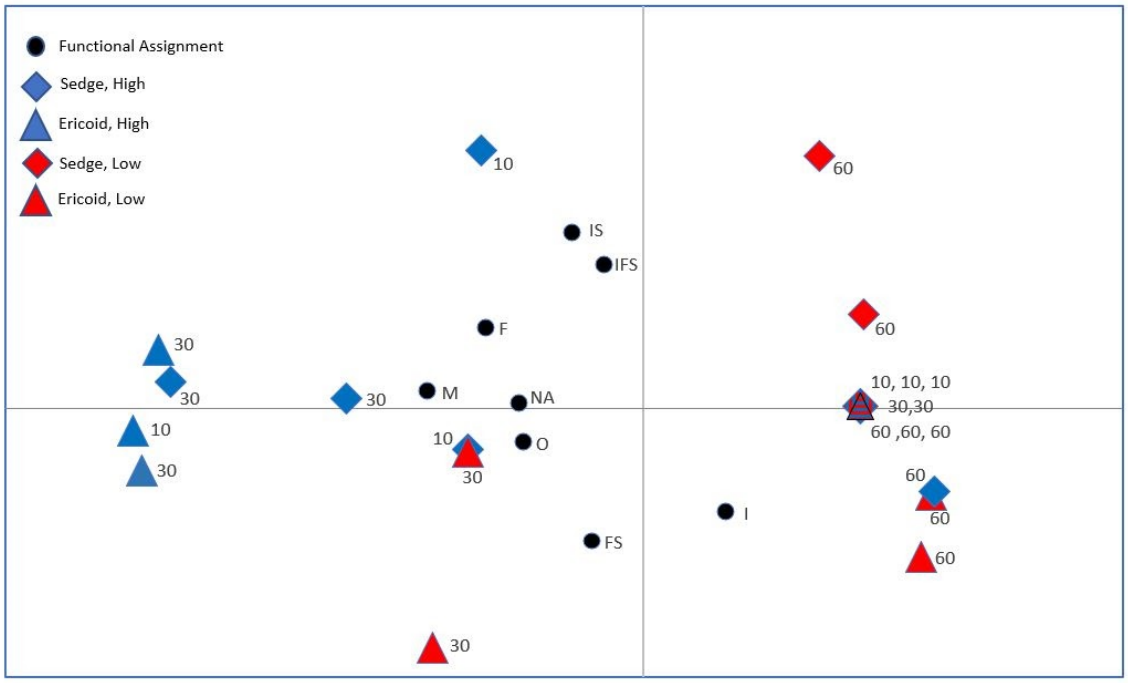

Figure 3**.** NMDS of functional assignments given to microbes containing *hgcA* gene hits. Functional assignments (black dot symbols) correspond to vector direction and magnitude of functional assignments. I= iron-reducing; F= fermenters; M= methanogens; S= sulfur-reducing; O= other; NA= no functional assignment provided. Stress value is equal to .06.

3.4 Phylogeny of *hgcA* gene hits

The construction of the *hgcA* phylogenetic tree shows clear branch integration of our PEAT sequences within the larger MCD sequence library (Fig. 4). Some of the largest clades were integrated within the branch belonging to the order Syntrophobacterales. A large branch section dominated by the family Methanoregulaceae had many branches containing PEAT sequences. Another dominant set of branches hosted a tight clustering of PEAT sequences under the order Geobacterales. Some of the more unresolved gene hits are more scattered throughout the tree, with some nested within branches classified at the order Thermodesulfovibrionales and more unresolved branches located amongst sequences classified under the phylum Spirochaetota and Firmicutes.

Within the tree, there was one clade that contained two OTUs but nearly half of all the PEAT gene counts. These two OTUs were functionally assigned as methanogens, making up a majority (84.6%) of all functionally assigned gene counts and roughly half (46.9%) of all gene counts. Other PEAT sequences classified as Archaea but not functionally resolved were placed on the tree much further from the larger Methanomicrobiales clade, and contained less than 2% of all gene counts. The largest unresolved and functionally unassigned OTU was cluster 213 which contained 13.7% of all gene counts and was present in the upper 10 and 30 cm of the high WT treatments. Including OTU 213, 38 more of the OTUs do not have enough taxonomic resolution for functional assignments, which left 62% of our gene counts unidentified.

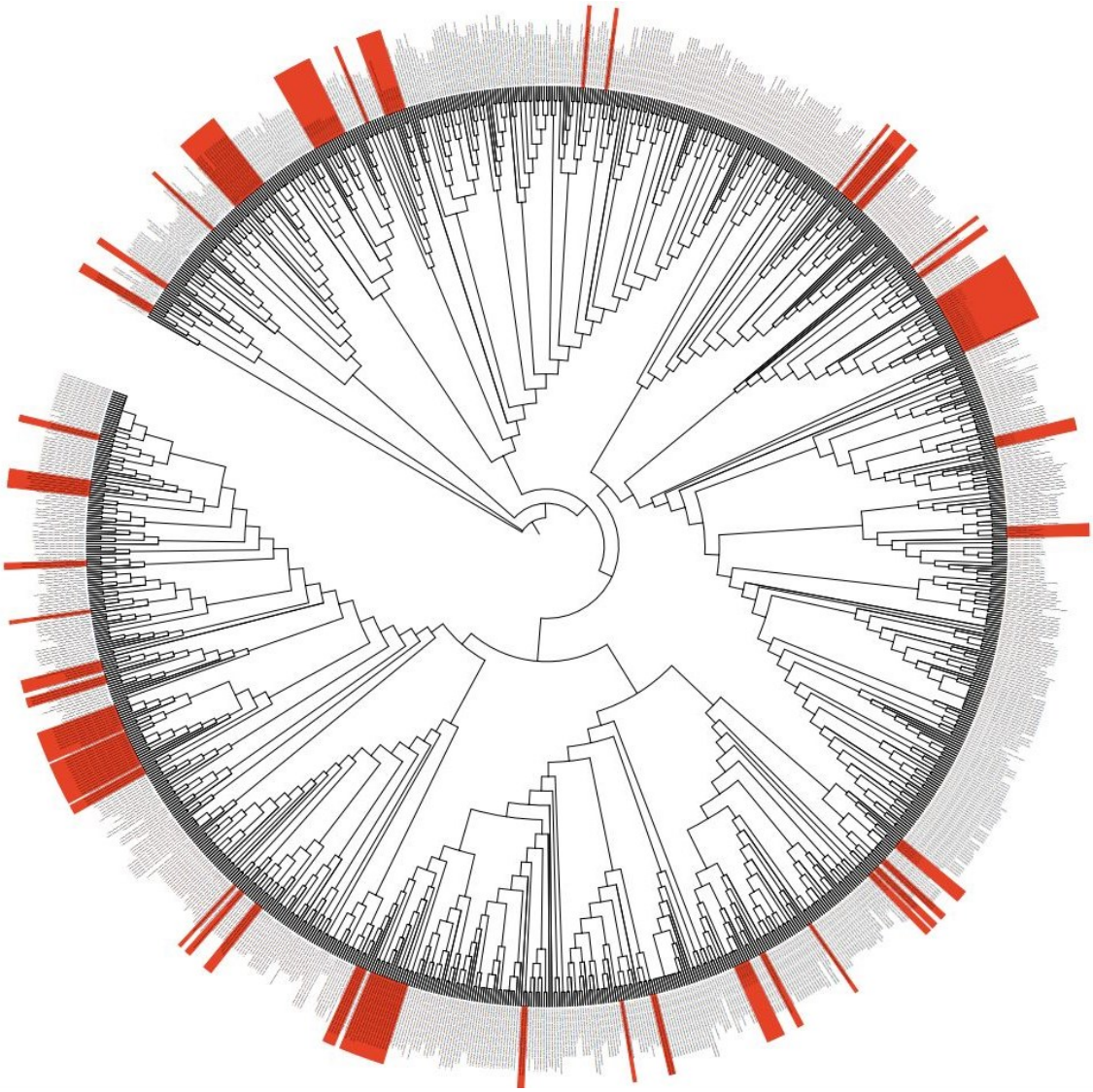

Figure 4. Phylogenetic tree of *hgcA* sequences, both PEAT and MCD. Branch length is ignored. PEAT sequences are highlighted in red.

Water table, depth, and the WT*depth interaction were significant predictors of OTU community composition in the peat, while PFG effects were not significant (Table 3). Visualization of community similarity of the taxonomically defined OTUs showed the first axis of the NMDS was strongly driven by depth, especially separating the samples at 60 cm from the upper two (10 cm and 30 cm) depths (Fig. 5). A weaker, more diffused grouping separated the gene hits at the 10 and 30 cm depths.

Table 3. PERMANOVA of treatment effects on read counts for OTUs.

Source	df	SS	MS	Psuedo-F	P(perm)	Unique Perms
PFG	1	2584.8	2584.8	1.4598	0.2082	1229
WT	1	7013.4	7013.4	3.9609	**0.0149**	1229
Depth	2	27742	13871	7.659	**0.0001**	9923
PFGxWT	1	1615.2	1615.2	0.9122	0.5532	1229
PFGxDepth	2	3609.2	1804.6	0.99641	0.481	9929
WTxDepth	2	12744	6371.8	3.5182	**0.0007**	9932
Bin(WTxPFG)	4	7021.9	1755.5	0.9693	0.5511	9866
Res	8	14489	1811.1			
Total	21	80017				

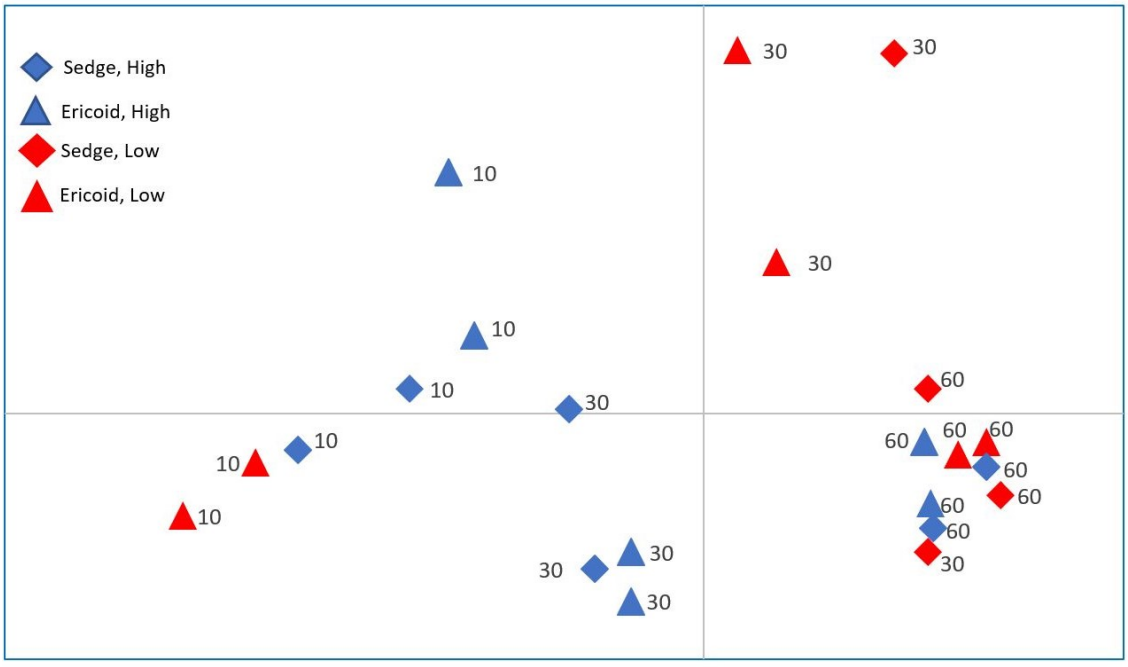

Figure 5**.** NMDS of taxonomic assignments given to microbes containing the *hgcA* gene hits. Stress value is equal to .06.

3.5 Depth effects shape weak relationships between *hgcA* gene count and MeHg and THg

In contrast to our prediction (H4) that there would be a positive relationship between measured methyl-mercury concentrations (in both peat and porewater) and the abundance of the mercury-methylating gene *hgcA,* we found minimal to no evidence of such a relationship. Analysis for overall trends when including all gene count data showed no significant relationships between *hgcA* gene count and any of the pore water nor peat fraction measurements (Table S1), including pore water total mercury (THg), porewater methyl-mercury (MeHg), or pore water percent methyl-mercury (%MeHg).

Additionally, there were no significant relationships between pore water fraction THg, MeHg, %MeHg and gene counts when separated by depth (Table S2) or either treatment (PFG, Table S3; WT, Table S4). However, there were a few significant relationships found within the peat fraction when broken down further. When separated by depth, there were significant negative relationships at the 30 cm depth between gene count and measured peat THg and MeHg, as well as a negative relationship between gene count and THg at the 60 cm depth (Table S1). Data transformation (log10(n+1)) did not influence statistical significant or goodness of fit of all data except for peat %MeHg when separated into the 30 cm depth, where log transformation found a significant negative relationship (Prob >F = 0.03) between peat %MeHg and log10 transformed gene count at 30 cm.

The severity of low WT treatments resulted in an unavailability of pore water data at the 10 cm depth. Limited data at this depth prevented full analysis of pore water relationships to Hg and MeHg measurements and *hgcA* gene counts.

4 Discussion

We used metagenomic analysis on soil core samples to identify the mercury-methylating gene *hgcA* from the novel peatland mesocosm experiment, PEATcosm, where we identified significant patterns and predictors of gene abundance in peat soil. The significant three-way interactive effects of depth, water table, and plant functional group on *hgcA* gene abundance would indicate a complex relationship that shapes the soil microbial community and the potential of that system to methylate mercury. Methanogens were the dominant functionally-assigned group recognized overall as well as within the treatments that were most *hgcA*-abundant. The phylogenetic analysis showed a distribution of gene hits nested throughout diverse lineages. When taken together, this research provides evidence to indicate the real effects of our treatments on the abundance of the mercury-methylating gene *hgcA*.

4.1 Treatment and depth effects on gene abundance

The significant interactive effects of our treatments were most pronounced at the 30 cm depth, and especially under ericaceous, high WT treatments– this suggests that the distance to the surface of the WT is more influential on the methylating community than any other factor, but were modified by PFG. Depth effects on the complete microbial community in this system have previously been analyzed, finding significant effects of PFG and WT experimental treatments on community composition and abundance (Lamit et al. 2021). Targeted metagenomics revealed distinct vertical stratification of microbial communities controlled by water table, plant functional group and depth, where low WT treatments favored aerobic bacteria while more anoxic conditions under high WT and ericoid treatments supported increased anaerobic populations. Mercury-methylation via *hgcAB* is only known to occur under anoxic conditions (Ma et al. 2019), which could explain some of the disparity in gene abundance between the high and low WT treatments as there is a greater oxic zone within the low WT peat. At the 30 cm sampling depth, the water table was measured at 23 to 30 cm above the soil sample location in high WT treatments while in low WT treatments the soil sample location varied between around 5 cm above or below the WT boundary. This distinction could aid in explaining

the discrepancies between high and low WT treatments at the intermediate depth, as communities measured at the 10-20 cm depth under high WT, are more similar to low WT at 30-40 cm when it comes to total gene abundance (Fig. 1). The increase of gene abundance under the Ericaceae treatments, but only under high water table at 30 cm, suggests that the root systems of Ericaceae and sedges had significant impacts on microbial composition at that depth, but those effects depended strongly on the water table height. The significant influence of PFG treatments at the 30 cm depth could be explained by addressing the contrasting rooting strategies between shallow-rooting mycorrhizal ericoid shrubs and deeper-rooting non-mycorrhizal aerenchymous sedges. Removal of sedges effectively eliminated the transport of both oxygen and root exudates to deeper areas in the soil, creating a more anoxic environment and shifting carbon use pathways, which had a positive effect on *hgcA* gene abundance in Ericaceae treatments. Variability between treatments did not occur at the deeper depth at 60 cm, as deeper soil zones maintain more consistent conditions that are less dependent on WT fluctuation and surface communities, at least within the time scale of this experiment.

4.2 Functional and taxonomic OTUs

Of the genes assigned to a functional group, methanogens were most dominant when compared to other functionally assigned groups, especially at the 30 cm depth. We found supporting evidence (H3) that methanogens had the largest proportional contribution to *hgcA* abundance in high WT and Ericaceae treatments (as compared to other treatments). Here, methanogen assignments outnumbered the unassigned genes (Fig. S1), and may be a dominant mercury methylator in these systems. The dominance of methanogens at the 30 cm depth for the high WT Ericaceae treatment was also reflected within community 16S analysis of overall composition which showed a relatively large positive shift of methanogens relative to other treatments (Lamit et al accepted). Our finding is similar to another study that used PCR primers that identified more sequences that clustered with methanogens than other groups, such as SRB, in a temperate swamp in Sweden (Schaefer et al. 2014). Similarly, these findings of methanogen functional dominance agree with patterns found by Hu et al. 2020, in which “older” peatlands, described as more nutrient poor and containing less sedge, were dominated by mercury-methylating methanogens

and syntrophy-driven methanogens. The case for syntrophy-driven Hg-methylation would be supported by the dominance of methanogens in this system and notable lack of *hgcA*-containing SRB or other groups.

We also identified potential methylators that were most closely clustered to families with known Hg methylators such as the families Syntrophaceae, Syntrophobacteraceae, Geobacteraceae, and Ruminococcaceae. One clade within the tree contains a rich OTU containing many PEAT sequences all classified under the Order Sytrophobacterales, whose function could not be assigned. Two OTUs that contained many of the methanogen-assigned PEAT *hgcA* sequences were intermixed with one another and showed diverse branching within the clade. The phylogenetic tree also places currently taxonomically and functionally unclassified *hgcaA* genes within branches that contained known Hg methylators of higher resolution. Taken together, these results could be an artifact of differing lengths of the PEAT sequences but could also indicate undiscovered diversity of Hg-methylating microbes.

4.3 Relationships between *hgcA* and Hg species and pools

Contrary to our hypotheses (H4), there was not support for any positive relationships of all data between gene abundance and either peat or pore water THg, MeHg, or %MeHg. Additionally, in contrast to our hypothesis (H1), we did not find the highest gene abundance under the predicted low WT sedge treatments where methyl-mercury pools were highest, instead we saw the highest abundance of *hgcA* genes under high WT ericoid treatments. The availability of inorganic Hg is reflected in the concentration of THg within the pore water portion, with THg within the solid phase peat acting more as a Hg reservoir (Haynes et al. 2017; Skyllberg et al. 2003). The largest pools of pore water THg and MeHg within this system were located at the 30-40 cm zone under low WT conditions (Haynes et al. 2019), which implicates additional factors as driving soil organisms to methylate mercury that is not based singularly on available Hg or *hgcA* gene abundance. The lack of an overall relationship between gene abundance and measured mercury concentrations may be due to inherent issues using mercury cycling genes as a

proxy for methyl-mercury concentrations, as data from other systems besides peatlands have shown (Christensen et al. 2019). At this time, it is unknown what controls Hg uptake in these organisms, but it is suspected that biogeochemical factors such as carbon quality or the availability of sulfates (Mitchell et al. 2008 AG) could fuel syntrophic SRB that likely support mercury-methylation in methanogens (Hu et al. 2020). This theory agrees with findings that show boreal fen wetlands of intermediate nutrient status to support the highest levels of MeHg production (Tjerngren et al. 2012).

Since MeHg concentration is a net result from the combined processes of methylation and demethylation, both are considered when attempting to explain patterns. The mismatch between treatments that were measured with elevated MeHg in the pore water at (low WT sedge) and the treatments with the highest gene abundance (high WT Ericaceae) could also be explained through the activity of demethylating microbial processes which are also performed by methanogens, SRB, IRB (Du et al. 2019) and methanotrophs (Lu et al. 2017), as well as abiotic processes (Sellers et al. 1996; Hu et al. 2020). Demethylation has been reported as elevated in older peatlands (Hu et al. 2020), which were classified as having increased levels of stored THg and decreased levels of available or incoming nutrients– similar to our system. Methanotrophs were top indicators of the upper 10-20 cm depths in high WT treatments (Lamit et al. accepted), and perhaps limited net MeHg production through demethylation activity, whose environmental controls are also not yet fully understood. However, this explanation does not fully agree with Haynes et al. 2019, which found little support of a de-methylation relationship to MeHg concentration, and instead offered the suggestion that Hg and MeHg mobility and re-adsorption may redistribute this compound throughout the vertical peat profile. The proposed vertical movement of Hg species through the peat profile may also help explain our finding that there was a significant negative relationship between gene abundance and both THg and MeHg at the 30 cm depth, as well as a significant negative relationship between gene abundance and THg at 60 cm.

4.4 Conclusions, limitations, and outlooks

Projections of regional temperature and precipitation shifts are predicted to influence water table heights and plant community compositions. In this novel finding, both plant functional group and water table treatments had significant effects on *hgcA* gene abundance. The setup of this experiment differs from the other publications in that PEATcosm was an experiment that controlled for environmental variables and measured the response of the microbes to these alterations over time. To our knowledge this is the first work to report on the effects of such treatments on mercury-methylating gene abundance in peatlands.

High water table treatments resulted in the most elevated *hgcA* gene abundance, especially at the 30 cm depth and under Ericaceae plant functional group treatments. These results contrasted with our initial hypotheses that gene abundance would be highest within treatments that had the highest reported MeHg concentrations. We instead found no correlation between total gene abundance and both THg and MeHg.

Methanogens stood out as the dominant functional group within the treatment combination that reported the highest overall gene abundance (high WT ericoid). To our knowledge, this is the first report of this pattern in peatlands, highlighting the potential for methanogens to be a dominant mercury methylator in these systems. The output of a shotgun metagenomics approach for identifying *hgcAB* has been found to be comparable to PCR-based methods, as were the qualitative outputs of both methodologies (Christensen et al. 2019), providing support for our findings using these methods. In the future, the addition of targeted transcript sequencing combined with metagenomic analysis would add to the Hg cycling story by differentiating gene presence with measurable activity.

Concerns about climate change are directly linked to peatland health; alterations to these peatland ecosystems alter decomposition rates and processes and thus their emitted products, like the neruotoxin MeHg. There are still many questions and black boxes

within this research area. Which microbes are participating in Hg methylation in response to the WT and PFG treatments? Which pathways/cycles are important to Hg-methylation and which are connected or run parallel with Hg-methylation in response to the treatments?

Finally, the data collected here represent single time points in a very dynamic system. It is possible that discrepancies in timing (days to a couple weeks) between the collection of different data sets could have missed some of the subtleties of rapid responses of microbial communities to sudden short-term environmental alterations. To fully understand these systems in detail, timeline measurements and samples of the mercury-methylating communities should be taken simultaneously to better document microbial responses to both short- and long-term variation in environmental factors. Future work that further examines peatland mercury cycling and the response of microbes to experimental manipulations like this one will aid in elucidating microbial identities and the connections between MeHg and Hg concentration, *hgcA* presence, and *hgcA* activity.

5 Literature Cited

Agethen, S, M, S., C, W., & KH, K. (2018). Plant rhizosphere oxidation reduces methane production and emission in rewetted peatlands. *Soil Biology & Biochemistry*, *125*, 125–135. https://doi.org/10.1016/J.SOILBIO.2018.07.006

Anderson, M., Clarke, K., & Gorley, R. (2008). PERMANOVA+ for Primer. Guide to Software and Statistical. *Methods.*, 214.

Bae, H.-S., Dierberg, F. E., & Ogram, A. (2014). Syntrophs Dominate Sequences Associated with the Mercury Methylation-Related Gene hgcA in the Water Conservation Areas of the Florida Everglades. *Applied and Environmental Microbiology*, *80*(20), 6517. https://doi.org/10.1128/AEM.01666-14

Cao, D., Chen, W., Xiang, Y., Mi, Q., Liu, H., Feng, P. Y., Shen, H., Zhang, C., Wang, Y., & Wang, D. (2021). The efficiencies of inorganic mercury bio-methylation by aerobic bacteria under different oxygen concentrations. *Ecotoxicology and Environmental Safety*, *207*, 111538. https://doi.org/10.1016/J.ECOENV.2020.111538

Christensen, G. A., Gionfriddo, C. M., King, A. J., Moberly, J. G., Miller, C. L., Somenahally, A. C., Callister, S. J., Brewer, H., Podar, M., Brown, S. D., Palumbo, A. V., Brandt, C. C., Wymore, A. M., Brooks, S. C., Hwang, C., Fields, M. W., Wall, J. D., Gilmour, C. C., & Elias, D. A. (2019). Determining the Reliability of Measuring Mercury Cycling Gene Abundance with Correlations with Mercury and Methylmercury Concentrations. *Environmental Science and Technology*, *53*(15), 8649–8663. https://doi.org/10.1021/acs.est.8b06389

Clarke, K. R., & Gorley, R. N. (2005). PRIMER: Getting started with v6. *PRIMER-E Ltd: Plymouth, UK*, *931*, 932.

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A New and Scalable Tool for the Selection of DNA and Protein Evolutionary Models. *Molecular Biology and Evolution*, *37*(1), 291–294. https://doi.org/10.1093/MOLBEV/MSZ189

Du, H., Ma, M., Igarashi, Y., & Wang, D. (2019). Biotic and Abiotic Degradation of Methylmercury in Aquatic Ecosystems: A Review. *Bulletin of Environmental Contamination and Toxicology 2019 102:5*, *102*(5), 605–611. https://doi.org/10.1007/S00128-018-2530-2

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. https://doi.org/10.1093/NAR/GKH340

Engle, M. A., Tate, M. T., Krabbenhoft, D. P., Schauer, J. J., Kolker, A., Shanley, J. B., & Bothner, M. H. (2010). Comparison of atmospheric mercury speciation and deposition at nine sites across central and eastern North America. *Journal of Geophysical Research: Atmospheres*, *115*(D18), 18306. https://doi.org/10.1029/2010JD014064

Evers, D. (2018). *The Effects of Methylmercury on Wildlife: A Comprehensive Review and Approach for Interpretation*. *5*, 181–194. https://doi.org/10.1016/B978-0-12-809665-9.09985-7

Fitzgerald*, W. F., Engstrom, D. R., Mason, R. P., & Nater, E. A. (1998). The Case for Atmospheric Mercury Contamination in Remote Areas. *Environmental Science and Technology*, *32*(1), 1–7. https://doi.org/10.1021/ES970284W

Gilmour, C. C., Podar, M., Bullock, A. L., Graham, A. M., Brown, S. D., Somenahally, A. C., Johs, A., Hurt, R. A., Bailey, K. L., & Elias, D. A. (2013). Mercury methylation by novel microorganisms from new environments. *Environmental Science and Technology*, *47*(20), 11810–11820. https://doi.org/10.1021/es403075t

Gordon, J., Quinton, W., Branfireun, B. A., & Olefeldt, D. (2016). Mercury and methylmercury biogeochemistry in a thawing permafrost wetland complex, Northwest Territories, Canada. *Hydrological Processes*, *30*(20), 3627–3638. https://doi.org/10.1002/hyp.10911

Gorham, E. (1991). Northern Peatlands: Role in the Carbon Cycle and Probable Responses to Climatic Warming. *Ecological Applications*, *1*(2), 182–195. https://doi.org/10.2307/1941811

Graham, A. M., Aiken, G. R., & Gilmour, C. C. (2012). Dissolved Organic Matter Enhances Microbial Mercury Methylation Under Sulfidic Conditions. *Environmental Science and Technology*, *46*(5), 2715–2723. https://doi.org/10.1021/ES203658F

Graham, A. M., Bullock, A. L., Maizel, A. C., Elias, D. A., & Gilmour, C. C. (2012). Detailed Assessment of the Kinetics of Hg-Cell Association, Hg Methylation, and Methylmercury Degradation in Several Desulfovibrio Species. *Applied and Environmental Microbiology*, *78*(20), 7337. https://doi.org/10.1128/AEM.01792-12

Grigal, D. F. (2002). Inputs and outputs of mercury from terrestrial watersheds: A review. *Environmental Reviews*, *10*(1), 1–39. https://doi.org/10.1139/A01-013

Haynes, K. M., Kane, E. S., Potvin, L., Lilleskov, E. A., Kolka, R. K., & Mitchell, C. P. J. (2019). Impacts of experimental alteration of water table regime and vascular plant community composition on peat mercury profiles and methylmercury production. *Science of the Total Environment*, *682*, 611–622. https://doi.org/10.1016/j.scitotenv.2019.05.072

Haynes, K. M., Kane, E. S., Potvin, L., Lilleskov, E. A., Kolka, R. K., & Mitchell, C. P. J. (2017a). Mobility and transport of mercury and methylmercury in peat as a function of changes in water table regime and plant functional groups. *Global Biogeochemical Cycles*, *31*(2), 233–244. https://doi.org/10.1002/2016GB005471

Haynes, K. M., Kane, E. S., Potvin, L., Lilleskov, E. A., Kolka, R. K., & Mitchell, C. P. J. (2017b). Gaseous mercury fluxes in peatlands and the potential influence of climate change. *Atmospheric Environment*, *154*, 247–259. https://doi.org/10.1016/j.atmosenv.2017.01.049

Hu, H., Wang, B., Bravo, A. G., Björn, E., Skyllberg, U., Amouroux, D., Tessier, E., Zopfi, J., Feng, X., Bishop, K., Nilsson, M. B., & Bertilsson, S. (2020). Shifts in mercury methylation across a peatland chronosequence: From sulfate reduction to methanogenesis and syntrophy. *Journal of Hazardous Materials*, *387*(October 2019), 121967. https://doi.org/10.1016/j.jhazmat.2019.121967

JMP® Pro, Version 15.2.0. SAS Institute Inc., Cary, NC, 1989–2021.

Klüpfel, L., Piepenbrock, A., Kappler, A., & Sander, M. (2014). Humic substances as fully regenerable electron acceptors in recurrently anoxic environments. *Nature Geoscience*, *7*(3), 195–200. https://doi.org/10.1038/ngeo2084

Le, S. Q., & Gascuel, O. (2008). An Improved General Amino Acid Replacement Matrix. *Molecular Biology and Evolution*, *25*(7), 1307–1320. https://doi.org/10.1093/MOLBEV/MSN067

Letunic, I., & Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research*, *47*(W1). https://doi.org/10.1093/NAR/GKZ239

Loisel, J., Gallego-Sala, A. V., Amesbury, M. J., Magnan, G., Anshari, G., Beilman, D. W., Benavides, J. C., Blewett, J., Camill, P., Charman, D. J., Chawchai, S., Hedgpeth, A., Kleinen, T., Korhola, A., Large, D., Mansilla, C. A., Müller, J., van Bellen, S., West, J. B., … Wu, J. (2020). Expert assessment of future vulnerability of the global peatland carbon sink. *Nature Climate Change*, *11*(January). https://doi.org/10.1038/s41558-020-00944-0

Louca, S., Hawley, A. K., Katsev, S., Torres-Beltran, M., Bhatia, M. P., Kheirandish, S., Michiels, C. C., Capelle, D., Lavik, G., Doebeli, M., Crowe, S. A., & Hallam, S. J. (2016). Integrating biogeochemistry with multiomic sequence information in a model oxygen minimum zone. *Proceedings of the National Academy of Sciences*, *113*(40), E5925–E5933. https://doi.org/10.1073/PNAS.1602897113

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology 2014 15:12*, *15*(12), 1–21. https://doi.org/10.1186/S13059-014-0550-8

Lu, X., Gu, W., Zhao, L., Ul Haque, M. F., DiSpirito, A. A., Semrau, J. D., & Gu, B. (2017). Methylmercury uptake and degradation by methanotrophs. *Science Advances*, *3*(5). https://doi.org/10.1126/sciadv.1700041

Ma, M., Du, H., & Wang, D. (2019). Mercury methylation by anaerobic microorganisms: A review. *Critical Reviews in Environmental Science and Technology*, *49*(20), 1893–1936. https://doi.org/10.1080/10643389.2019.1594517

Martínez Cortizas, A., Biester, H., Mighall, T., & Bindler, R. (2007). Climate-driven enrichment of pollutants in peatlands. Biogeosciences, 4(5), 905–911. https://doi.org/10.5194/BG-4-905-2007

Mazrui, N. M., Jonsson, S., Thota, S., Zhao, J., & Mason, R. P. (2016). Enhanced availability of mercury bound to dissolved organic matter for methylation in marine sediments. *Geochimica et Cosmochimica Acta*, *194*, 153–162. https://doi.org/10.1016/J.GCA.2016.08.019

McDaniel, E. A., Peterson, B. D., Stevens, S. L. R., Tran, P. Q., Anantharaman, K., & McMahon, K. D. (2020). *Expanded Phylogenetic Diversity and Metabolic Flexibility of Mercury-Methylating Microogranisms*. *August*, 1–21.

Mitchell, C. P. J., Branfireun, B. A., & Kolka, R. K. (2008). Total mercury and methylmercury dynamics in upland–peatland watersheds during snowmelt. *Biogeochemistry 2008 90:3*, *90*(3), 225–241. https://doi.org/10.1007/S10533-008-9246-Z

Mitchell, C. P. J., Branfireun, B. A., & Kolka, R. K. (2008). Assessing sulfate and carbon controls on net methylmercury production in peatlands: An in situ mesocosm approach. *Applied Geochemistry*, *23*(3), 503–518. https://doi.org/10.1016/J.APGEOCHEM.2007.12.020

Podar, M., Gilmour, C. C., Brandt, C. C., Soren, A., Brown, S. D., Crable, B. R., Palumbo, A. V., Somenahally, A. C., & Elias, D. A. (2015). Global prevalence and distribution of genes and microorganisms involved in mercury methylation. *Science Advances*, *1*(9), 1–13. https://doi.org/10.1126/sciadv.1500675

Potvin, L. R., Kane, E. S., Chimner, R. A., Kolka, R. K., & Lilleskov, E. A. (2014). Effects of water table position and plant functional group on plant community, aboveground production, and peat properties in a peatland mesocosm experiment (PEATcosm). *Plant and Soil*, *387*(1–2), 277–294. https://doi.org/10.1007/s11104-014-2301-8

Price, M. N., Dehal, P. S., & Arkin, A. P. (2009). FastTree: Computing Large Minimum Evolution Trees with Profiles instead of a Distance Matrix. *Molecular Biology and Evolution*, *26*(7), 1641–1650. https://doi.org/10.1093/MOLBEV/MSP077

Schaefer, J. K., Kronberg, R. M., Morel, F. M. M., & Skyllberg, U. (2014). Detection of a key Hg methylation gene, hgcA, in wetland soils. *Environmental Microbiology Reports*, *6*(5), 441–447. https://doi.org/10.1111/1758-2229.12136

Schaefer, J. K., Kronberg, R. M., Björn, E., & Skyllberg, U. (2020). Anaerobic guilds responsible for mercury methylation in boreal wetlands of varied trophic status serving as either a methylmercury source or sink. *Environmental Microbiology*, *22*(9), 3685–3699. https://doi.org/10.1111/1462-2920.15134

Schuster, P. F., Striegl, R. G., Aiken, G. R., Krabbenhoft, D. P., Dewild, J. F., Butler, K., Kamark, B., & Dornblaser, M. (2011). Mercury export from the Yukon River Basin and potential response to a changing climate. *Environmental Science and Technology*, *45*(21), 9262–9267. https://doi.org/10.1021/es202068b

Seller, P., Kelly, C. A., Rudd, J. W. M., & Mac Hutchon, A. R. (1996). Photodegradation of methylmercury in lakes. *Nature*, *380*(6576), 694–697. https://doi.org/10.1038/380694A0

Skyllberg, U., Qian, J., Frech, W., Xia, K., & Bleam, W. F. (2003). Distribution of mercury, methyl mercury and organic sulphur species in soil, soil solution and stream of a boreal forest catchment. *Biogeochemistry 2003 64:1*, *64*(1), 53–76. https://doi.org/10.1023/A:1024904502633

Tjerngren, I, T, K., E, B., & U, S. (2012). Potential Hg methylation and MeHg demethylation rates related to the nutrient status of different boreal wetlands. *Biogeochemistry.*, *108*(1–3), 335–350. https://doi.org/10.1007/S10533-011-9603-1

Turetsky, M. R., Harden, J. W., Friedli, H. R., Flannigan, M., Payne, N., Crock, J., & Radke, L. (2006). Wildfires threaten mercury stocks in northern soils. *Geophysical Research Letters*, *33*(16). https://doi.org/10.1029/2005GL025595

Villar, E., Cabrol, L., & Heimbürger-Boavida, L. E. (2020). Widespread microbial mercury methylation genes in the global ocean. *Environmental Microbiology Reports*, *12*(3), 277–287. https://doi.org/10.1111/1758-2229.12829

Watanabe,C & Satoh, H. (1996). Evolution of our understanding of methylmercury as a health threat. Environmental Health Perspectives, 104 Suppl 2(Suppl 2), 367–379. https://doi.org/10.1289/EHP.96104S2367

6 Supplemental Materials

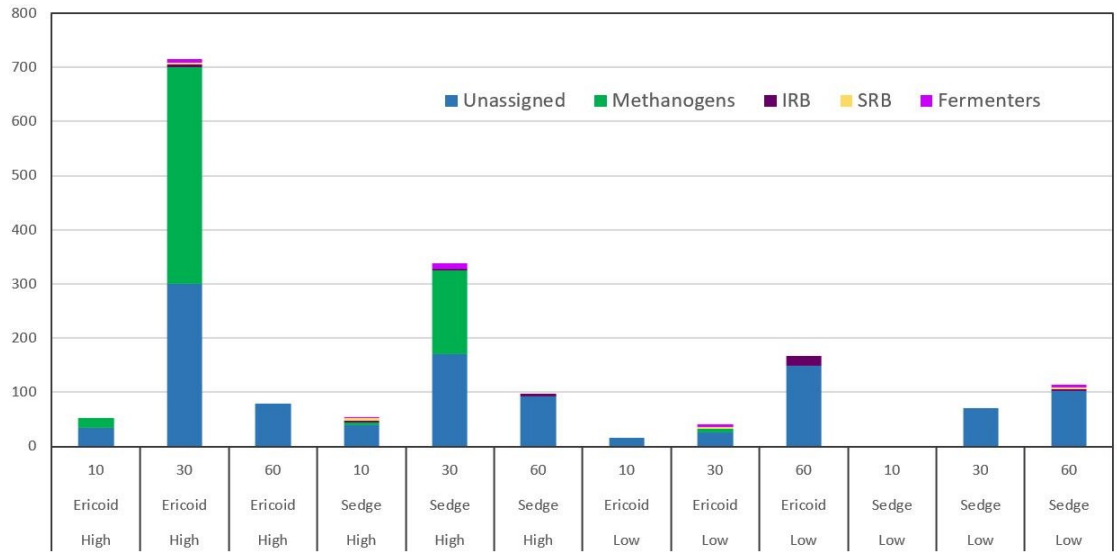

Figure S1. Normalized (un-transformed) *hgcA* gene counts for the different functional groups by treatment combination. Stacked colored bars represent conservative functional assignments.

Table S1. Results from one-way ANOVA of all normalized and log transformed (log10(n+1)) gene count data versus measured Hg concentrations (Haynes et al. 2019). P-values are for analysis of variance. RSquare value corresponds to line of fit value for data distribution.

	Normalized Gene Count		Log transformed Gene Count	
	Prob > F	Rsquare	Prob > F	Rsquare
Peat THg	0.65	0.009	0.31	0.047
Porewater THg	0.81	0.004	0.89	0.001
Peat MeHg	0.48	0.023	0.54	0.018
Porewater MeHg	0.8	0.004	0.95	<0.001
Peat %MeHg	0.68	0.008	0.4	0.032
Porewater %MeHg	0.39	0.053	0.79	0.005

Table S2. Results of one-way ANOVA of normalized *hgcA* gene counts versus measured Hg concentrations by depth (Haynes et al. 2019). P-values are for analysis of variance. RSquare value corresponds to line of fit value for data distribution.

	10 cm		30 cm		60 cm	
	Prob > F	Rsquare	Prob > F	Rsquare	Prob > F	Rsquare
Peat THg	0.61	0.047	**0.006**	0.74	**0.005**	0.752
Porewater THg	0.99	<0.001	0.69	0.061	0.42	0.134
Peat MeHg	0.86	0.006	**0.04**	0.527	0.052	0.493
Porewater MeHg	0.62	0.146	0.8	0.017	0.47	0.091
Peat %MeHg	0.75	0.017	0.08	0.431	0.17	0.288
Porewater %MeHg	0.75	0.062	0.36	0.279	0.7	0.033

Table S3. Results of one-way ANOVA of normalized *hgcA* gene counts versus measured Hg concentrations by PFG treatment (Haynes et al. 2019). P-values are for analysis of variance. RSquare value corresponds to line of fit value for data distribution.

	Ericoid		Sedge	
	Prob > F	Rsquare	Prob > F	Rsquare
Peat THg	0.53	0.041	0.61	0.027
Porewater THg	0.74	0.095	0.62	0.052
Peat MeHg	0.49	0.049	0.96	<0.001
Porewater MeHg	0.15	0.242	>0.99	<0.001
Peat %MeHg	0.68	0.017	0.95	<0.001
Porewater %MeHg	0.25	0.187	0.6	0.058

Table S4. Results of one-way ANOVA of normalized *hgcA* gene counts versus measured Hg concentrations by WT treatment (Haynes et al. 2019). P-values are for analysis of variance. RSquare value corresponds to line of fit value for data distribution.

	High WT		Low WT	
	Prob > F	Rsquare	Prob > F	Rsquare
Peat THg	0.67	0.002	0.73	0.012
Porewater THg	0.56	0.035	0.82	0.031
Peat MeHg	0.49	0.05	0.92	0.001
Porewater MeHg	0.74	0.012	0.24	0.326
Peat %MeHg	0.64	0.022	0.85	0.004
Porewater %MeHg	0.57	0.033	0.38	0.381