

**Artificial wetlands in the Saint Maurice River, Québec, Canada, a hotspot for mercury  
methylating microorganisms**

Makayla Catherine Harrison

A Thesis  
in  
The Department  
of  
Biology

Presented in Partial Fulfillment of the Requirements  
For the Degree of Master of Science (Biology) at  
Concordia University  
Montreal, Quebec, Canada

September 2019

© Makayla Catherine Harrison, 2019



**CONCORDIA UNIVERSITY**  
**School of Graduate Studies**

This is to certify that the thesis prepared

By: **Makayla Catherine Harrison**

Entitled: **Artificial wetlands in the Saint Maurice River, Québec, Canada, a hotspot for mercury methylating microorganisms**

and submitted in partial fulfillment of the requirements for the degree of

**Master of Science (Biology)**

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

\_\_\_\_\_ External Examiner  
*Dr. Carly Ziter*

\_\_\_\_\_ Examiner  
*Dr. Dolores Planas (UQAM)*

\_\_\_\_\_ Examiner  
*Dr. James Grant*

\_\_\_\_\_ Examiner, Thesis co-supervisor  
*Dr. Alexandre Poulain (University of Ottawa)*

\_\_\_\_\_ Thesis supervisor  
*Dr. David Walsh*

Approved by \_\_\_\_\_

Chair of Department or Graduate Program Director

\_\_\_\_\_ Date \_\_\_\_\_

Dean,

## Abstract

Artificial wetlands in the Saint Maurice River, Québec, Canada, a hotspot for mercury methylating microorganisms

Makayla Catherine Harrison

Mercury (Hg) is a nonessential trace metal that occurs in aquatic environments through both natural and anthropogenic deposition. When mercury undergoes methylation, it is transformed into methylmercury (MeHg), a neurotoxin capable of bioaccumulating in food webs. This process is facilitated by methylating microorganisms (e.g. sulfate-reducing bacteria, iron-reducing bacteria, and methanogens), all of which contain a *hgcAB* gene pair. High levels of methylmercury have been recorded in local fish populations near run-of-river hydroelectric dams along the St. Maurice River in Québec, Canada. In order to investigate the source MeHg levels, periphyton biofilms were collected from three sampling sites upstream of the dam, a natural, flooded and artificial wetland site. Using periphyton biofilm samples, we compared the microbial community composition within and between periphyton biofilms, assessed the natural diversity of mercury methylating microorganisms within periphyton biofilms, and quantified the abundance of mercury methylating microorganisms at each sampling site in relation to mercury methylation and demethylation. Different rates of mercury methylation were observed in all three sampling sites, the highest rates seen in the artificial wetlands. This was supported by 16S rRNA analyses, which revealed lineages associated with mercury methylating microorganisms being present at all sites, with the greatest abundance observed in the wetlands. Additionally, the presence of mercury methylating microorganisms was later confirmed through the presence of potential *hgcA* and *hgcB* genes found within the metagenomes at the wetland sampling site. Overall, these findings highlight the importance of the artificial wetlands and subsequent flooding following the construction of the Chute-Allard hydroelectric generating station, as being important contributors to increased MeHg levels and its bioaccumulation in local fish populations.

## **Acknowledgements**

I would like to thank my supervisor, Dr. David Walsh, for his continued guidance and support throughout my Master's degree. I would also like to thank the members of the Walsh Lab, specifically Dr. Susanne Kraemer and Rebecca Garner, for their help in processing the 16S rRNA and metagenomic data sets. Thank you to the 2016 Wemotaci field team, specifically Maxime Leclerc, for conducting the *in-situ* incubations and collection of the periphyton biofilms from the St. Maurice River. I would also like to thank my co-supervisor, Dr. Alexandre Poulain and my committee members Dr. Dolores Planas and Dr. James Grant as well as my external examiner, Dr. Carly Ziter.

## **Contributions of authors**

This thesis written by Makayla Harrison entitled “Artificial wetlands in the Saint Maurice River, Québec, Canada, a hotspot for mercury methylating microorganisms”, includes the work and collaboration of Maxime Leclerc, Dominic E. Ponton, Marc Amyot, Alexandre J. Poulain, and David A. Walsh. Makayla Harrison and David Walsh developed the research objectives and methodologies implemented in this thesis. Maxime Leclerc and Dominic Ponton performed the *in-situ* experiments and the subsequent collection of periphyton biofilms from the St. Maurice River in Québec, Canada in the summer of 2017. Maxime Leclerc performed the mercury methylation and demethylation rate measurements as part of his thesis at the Université de Montréal. The DNA and RNA from the periphyton biofilms used for all further 16S rRNA gene and transcript analyses as well as the metagenomic analyses was extracted by Makayla Harrison.

## Table of Contents

<b>List of Figures .....</b>	<b>1</b>
<b>List of Supplementary Figures .....</b>	<b>4</b>
<b>List of Supplementary Tables .....</b>	<b>5</b>
<b>1. Introduction .....</b>	<b>7</b>
1.1 Mercury .....	7
1.2 Methylmercury in the environment .....	8
1.3 Genetic basis for mercury methylation.....	10
1.4 Altered landscapes and their impact on mercury methylation .....	12
1.5 Chute-Allard run-of-river hydroelectric generating station .....	13
<b>2. Materials and Methods .....</b>	<b>15</b>
2.1 Field sampling in the St. Maurice River, QC, Canada .....	15
Figure 1 .....	17
2.2 Methylation and demethylation rates .....	18
2.3 DNA and RNA extractions.....	19
2.4 16S rRNA gene and transcript analysis.....	19
2.4.1 <i>16S rRNA amplification</i> .....	19
2.4.2 <i>16S rRNA gene and transcript amplicon sequencing</i> .....	21
2.4.3 <i>16S rRNA alpha diversity</i> .....	22
2.4.4 <i>Diversity, abundance and variability of microorganisms</i> .....	22
2.4.5 <i>Taxonomic analysis of 16S rRNA amplicons</i> .....	23
2.4.6 <i>Indicator species analysis</i> .....	23
2.6 Metagenomic analysis .....	24
2.6.1 <i>Metagenomic sequencing</i> .....	24
2.6.2 <i>Taxonomic analysis</i> .....	24
2.6.3 <i>Potential hgcA and hgcB genes in assembled metagenomes</i> .....	25
2.6.4 <i>Potential hgcA genes in the unassembled metagenome</i> .....	26
2.6.5 <i>Functional gene markers associated with mercury methylation</i> .....	26
<b>3. Results.....</b>	<b>26</b>
3.1 Methylation and demethylation rate results .....	26
Figure 2.....	<b>Error! Bookmark not defined.</b>
3.2 16S rDNA and rRNA analysis .....	28
3.2.1 <i>Alpha diversity analysis</i> .....	29

Figure 3.....	30
3.2.2 Diversity, abundance, and variability of microorganisms .....	30
Figure 4.....	32
Figure 5.....	34
3.2.3 Taxonomic analysis of 16S rRNA amplicons.....	34
Figure 6.....	35
3.2.4 Indicator species analysis.....	36
Figure 7.....	38
3.3 Metagenomic analysis .....	39
3.3.1 Taxonomic analysis of potential mercury methylating taxa.....	40
Figure 8.....	41
3.3.2 Potential <i>hgcA</i> and <i>hgcB</i> genes in the assembled metagenome .....	42
Figure 9.....	43
Figure 10.....	45
Figure 11.....	46
3.3.3 Potential <i>hgcA</i> genes in the unassembled metagenome .....	47
Figure 12.....	48
Figure 13.....	50
3.3.4. Functional gene markers associated with mercury methylation .....	51
Figure 14.....	53
<b>4. Discussion .....</b>	<b>54</b>
4.1 Diversity and abundance of microbial communities from 16S rRNA amplicons.....	54
4.2 Diversity and abundance of microbial communities from metagenomic data .....	56
4.3 Potential <i>hgcA</i> and <i>hgcB</i> genes in the artificial wetlands.....	57
4.4 Conclusions and future work.....	59
<b>References .....</b>	<b>61</b>
<b>5. Supplementary Material and Methods.....</b>	<b>73</b>
5.1 <i>hgcAB</i> gene pair analysis.....	73
5.2 Dilution series.....	74
5.3 Clade-Specific primer development .....	74
<b>6. Supplementary Results .....</b>	<b>74</b>
6.1 <i>hgcAB</i> gene pair analysis .....	74
<b>Supplementary Figures and Tables .....</b>	<b>76</b>
Figure S1 .....	76

Figure S2 .....	77
Figure S3 .....	78
Table S1 .....	79
Table S2.....	80
Table S3.....	81
Table S4.....	82
Table S5.....	83
Table S6.....	83
Table S7.....	84

## List of Figures

**Figure 1.** Map of sampling sites upstream of the Chute-Allard hydroelectric generating station, QC, Canada. Sampling site one is a portion of the natural river where periphyton was collected from natural substrates (Natural). Sampling site two is a part of the river that flooded following the construction of the Chute-Allard hydroelectric generating station and periphyton samples were collected from artificial substrates left the year prior (Flooded). Sampling site three is in an artificial wetland channel made prior to the construction of the dam for water flow from the flooding following the construction of the dam (Wetland). The periphyton in the wetland site was collected from artificial substrate left during the sampling season of the previous year.

**Figure 2A.** Mean and standard error of mercury methylation rates per day at each sampling site. Methylation rates between all three sampling sites are not statistically significant from each other, however there is an observed difference between the natural and flooded sampling sites and the wetland sampling site.

**Figure 2B.** Mean and standard error of mercury demethylation rates per day for each sampling site. The natural sampling site is statistically different from both the flooded and wetland sites; however, the flooded and wetland sites are not statistically significant from each other.

**Figure 3.** Alpha diversity analysis for all 16S rRNA amplicons sequenced at each sampling sites using the Simpson Index. rDNA is represented by red circles and rRNA by blue.

**Figure 4.** Heatmap illustrating the abundance of ASVs of all phyla using the 16S rDNA and rRNA amplicons from all three sampling sites. Higher abundances were observed in 16S rDNA amplicons rather than rRNA, with the greatest diversity of abundant phyla of both rDNA and rRNA observed in the artificial wetland sampling site.

**Figure 5A.** Principal coordinate analysis plot using the Bray-Curtis dissimilarity measure of the 16S rDNA. Samples grouped by sampling site are statistically significant from one another but are not statistically significant based on treatments (i.e., replicates). Sampling sites are

represented by colours and treatments or replicates by solid shapes. The natural, flooded and wetland sites are green, red and blue respectively, with the replicates; time zero (t0), 48 hr incubation (E), and 48 hr isotope incubation (F) represented by squares, circles and triangles respectively.

**Figure 5B.** Principal coordinate analysis plot using the Bray-Curtis dissimilarity measure of the 16S rRNA. Samples grouped by sampling site are statistically significant from one another but are not statistically significant based on treatments (i.e., replicates). Sampling sites are represented by colours and treatments by outlined shapes. The natural, flooded and wetland sites are green, red and blue respectively, with the replicates; time zero (t0), 48 hr incubation (E), and 48 hr isotope incubation (F) outlined by squares, circles and triangles respectively.

**Figure 5C.** Principal coordinate analysis plot using the Bray-Curtis dissimilarity measure of the 16S rRNA amplicon types. Samples grouped by sampling site and amplicon type. Sampling sites and treatments from 16S rRNA are represented by colours and outlined shapes. The natural, flooded and wetland sites are green, red and blue respectively, with the replicates; time zero (t0), 48 hr incubation (E), and 48 hr isotope incubation (F) represented by squares, circles and triangles respectively.

**Figure 6A.** Mean percent read counts from 16S rRNA amplicons of families associated with mercury methylation based on the main groups of mercury methylators (i.e. Methanogens, Firmicutes and Chloroflexi). Bars are broken down into rDNA and rRNA for all three sampling sites. *Methanoregulaceae* are the most abundant family seen across all sampling sites with their greatest abundance observed in the wetland site.

**Figure 6B.** Mean percent read counts from 16S rRNA amplicons of families associated with mercury methylation based on the main classes of mercury methylators (i.e. *Delta-Proteobacteria*). Bars are broken down into rDNA and rRNA for all three sampling sites. *Geobacteraceae* are the most abundant family seen across all sampling sites with their greatest abundance observed in the wetland site.

**Figure 7.** Indicator species analysis broken down into important families of mercury methylating microorganisms found in known mercury methylating groups e.g., *Delta-Proteobacteria*, *Firmicutes*, *Methanogens*, and *Chloroflexi*. Each plot contains the data from both 16S rDNA (red) and rRNA (blue) for each sampling site.

**Figure 8.** Plot illustrating the number of assigned taxa at each taxonomic rank for each replicate at every sampling site using the metagenomic data.

**Figure 9.** The mean and standard error of potential *hgcA* (A) and *hgcB* (B) genes per sampling site found in the assembled metagenomic data at various stringency levels. The three sampling sites; natural, flooded and wetland are distinguished by the colours; red, green and blue respectively.

**Figure 10.** Sequence alignments comparing a subset of known *hgcA* genes from *Delta-Proteobacteria*, *Firmicutes*, and *Euryarchaeota* to blastp outputs from the assembled metagenomic data. The red boxes indicate highly conserved regions in the putative cobalamin-binding domain of the *hgcA* gene. Abbreviations for amino acids; A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

**Figure 11.** Phylogenetic analysis comparing known *hgcA* sequences to potential *hgcA* genes from the replicates (e-value  $1 \times 10^{-40}$ ) of assembled protein-coding genes from the metagenomes at the artificial wetland sampling site. The replicates are time zero (t0), 48 hr incubation (E), and 48 hr isotope incubation (F) represented by the colours blue, green, and red respectively.

**Figure 12.** The mean and standard error of potential *hgcA* genes per sampling site found in the unassembled portion of the metagenomic data at various stringency levels. The three sampling sites; natural, flooded and wetland are distinguished by the colours; red, green and blue respectively.

**Figure 13.** Sequence alignments comparing a subset of known *hgcA* genes from *Delta-Proteobacteria*, *Firmicutes*, and *Euryarchaeota* to tblastn outputs from the unassembled metagenomic data. The red boxes indicate highly conserved regions in the putative cobalamin-binding domain of the *hgcA* gene. Abbreviations for amino acids; A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

**Figure 14.** Bar plot indicating the number of functional gene markers found at each sampling site. *dsrA* (sulfite reductase alpha subunit), *merA* (mercuric reductase), and *mcrA* (methyl-coenzyme M reductase alpha subunit) are represented by the colours blue, red, and green respectively.

### List of Supplementary Figures

**Figure S1.** Bar plot representing the mean relative percent abundance from 16S rDNA and rRNA for each sampling site at the kingdom level. Over 90% of all samples are Bacteria, with less than 5% being Eukaryotes and Archaea. Archaea were only detected at the wetland site.

**Figure S2.** Bar plot representing the mean relative percent abundance from 16S rDNA and rRNA for each sampling site at the phylum level from the bacterial kingdom. The most abundant phyla across all sampling sites for both 16S rDNA and rRNA are cyanobacteria, proteobacteria, verrucomicrobia, chloroflexi, actinobacteria, acidobacteria and planctomycetes.

**Figure S3.** Plot illustrating the number of assigned taxa at the phylum rank for each replicate at every sampling site of the metagenomic data. Proteobacteria and cyanobacteria are the most abundant phyla across all sampling sites.

## List of Supplementary Tables

**Table S1A.** 16S rRNA amplicon sequencing data showing the number of unique ASVs and read counts for each sampling site. Samples ending in -C are 16S rRNA amplicons and those ending in -G are 16S rDNA amplicons.

**Table S1B.** Average number of unique ASVs and read counts based on 16S rDNA sequences for each treatment at all three sampling sites.

**Table S1C.** Average number of unique ASVs and read counts based on 16S rRNA sequences for each treatment at all three sampling sites.

**Table S2.** Mean relative percent abundance of each kingdom for the 16S rDNA and rRNA at each sampling site. Values used to create Figure S1.

**Table S3.** Mean relative percent abundance of each phyla for the 16S rDNA and rRNA at each sampling site. Values used to create Figure S2.

**Table S4.** Mean relative percent abundance of each family previously implicated in mercury methylation found in the 16S rDNA and rRNA dataset at each sampling site. Values used to create Figure 6.

**Table S5.** Breakdown of metagenomic data after each step from sequencing to assigning taxonomy.

**Table S6A.** Number of potential *hgcA* genes found in unassembled metagenome data at various levels of stringency for each sampling site and treatment.

**Table S6B.** Average number of potential *hgcA* genes found in the unassembled metagenome data at various levels of stringency (e-values) at each sampling site.

**Table S6C.** Standard variance of potential *hgcA* genes found in the unassembled metagenome data at various levels of stringency at each sampling site.

**Table S7A.** Table of primer sets used to amplify *hgcA* and *hgcB* found in the literature. Primers are sorted into broad-range or clade-specific. Red represents degenerate bases; Y = C/T, R = A/G, M = A/C, S = C/G, W = A/T, V = A/C/G, B = C/G/T, D = A/G/T, H = A/C/T, N = A/T/C/G.

**Table S7B.** Table of less degenerate and more specific primer sets used to amplify *hgcA* and *hgcB* modified from primers found in literature. Primers are sorted into clades. Red represents degenerate bases; Y = C/T, R = A/G, M = A/C, S = C/G, W = A/T, V = A/C/G, B = C/G/T, D = A/G/T, H = A/C/T, N = A/T/C/G.

## 1. Introduction

### 1.1 Mercury

Mercury is a naturally occurring element that can be found in environments all around the world. The main method of transport of elemental mercury (Hg(0)) is through atmospheric deposition following emissions. This elemental mercury can be transported long distances, with a residence time of several months to up to a year in the atmosphere (Pirrone et al., 2010; Selin et al., 2009). Hg(0) will eventually be oxidized to Hg(II) and removed by particle and gas-phase dry deposition or through scavenging by precipitation from rain and snow (Driscoll et al., 2013). However, since Hg(0) can be transported large distances and has a long residence time, it may travel to remote locations and be deposited into various environments such as the Arctic and Antarctic before being oxidized (Durnford et al., 2010; Sprovieri et al., 2002). Other forms of mercury such as reactive gaseous mercury (RGM) and particulate bound ionic mercury (Hg(II)) only have an atmospheric residence time of a few hours to a couple of days, and can therefore not travel as far, and are usually deposited locally (Driscoll et al., 2013). For this reason, Hg inputs to ecosystems are primarily in the form of Hg(II) and are of the greatest concern for the production of the most dangerous form of Hg, monomethylmercury (MMHg or MeHg), a potent neurotoxin (Driscoll et al., 2013). Hg(II) is transformed into MeHg by biotic processes mediated in anoxic environments (Gilmour et al., 2013; Hamelin et al., 2011).

Primary emission originates from both natural and anthropogenic processes and have the ability to transfer Hg from lithospheric reservoirs to the atmosphere and eventually to the land and oceans (Driscoll et al., 2013). Some of these natural processes include the outgassing of granite rock, volcanoes, and forest wildfires, while some anthropogenic sources include gold mining, industrial processes, and commercial purposes such as coal combustion (Bravo et al., 2018; Broussard et al., 2002; Parks et al., 2013). Anthropogenic primary sources of Hg are responsible for the increase in the global pool of Hg in surface reservoirs by approximately 1900 to 2900 Mg yr<sup>-1</sup> globally (Driscoll et al., 2013). The Hg deposited there can be reduced to Hg(0) and re-emitted into the atmosphere, known as secondary emission, providing an exchange of Hg between terrestrial and ocean reservoirs via the atmosphere (Driscoll et al., 2013). These

secondary sources are responsible for the redistribution of the original global Hg pools among and within ecosystems (Driscoll et al., 2013).

As Hg can be spread across many ecosystems, the amount of Hg(II) available to be transformed into MeHg, the most toxic form of Hg, is increasing and causing severe health concerns for the public. Of the most known accounts of mercury poisoning, and their symptoms, was discovered following the release of chemical waste in Minamata Bay, Japan, resulting in the deaths of hundreds of people, later called Minamata Disease (Tamashiro et al., 1984). Another account of mercury poisoning occurring in Canada affected first nation communities in Northwestern Ontario between the 1960s and 1970s. Individuals within these communities contracted Minamata disease due to the dumping of chemical waste much like those in Minamata Bay, Japan. The toxic effects of MeHg are produced through enzyme inhibition, protein precipitation, and generalized corrosive action. Some symptoms caused from the inhalation or ingestion of MeHg include muscle weakness, speech impairment, odd sensations such as the feeling of pins and needles in one's extremities, and could result in both breathing and kidney problems, and may even result in death (Broussard et al., 2002). Another major concern regarding the transformation of Hg(II) into MeHg, is its ability to bioaccumulate and biomagnify within the food web. MeHg accumulates in fish populations and acts as the main exposure route to human populations. MeHg is able to bioaccumulate so effectively because of efficient uptake, but poor elimination from organisms (Holmes et al., 2009).

## **1.2 Methylmercury in the environment**

The net concentration of MeHg in ecosystems is largely dependent on biotic and abiotic factors affecting mercury methylation ( $K_m$ ) and demethylation ( $K_d$ ) rates. Mercury methylation is a biotic process mediated by anaerobic microorganisms that could potentially be controlled by abiotic factors such as pH, redox potential, salinity, light, temperature, water level, and the presence and abundance of sulfur and iron (Celo et al., 2006; Correia et al., 2012; Eckley & Hintelmann, 2006; Fredrik et al., 2014; Gilmour et al., 1992; Grégoire & Poulain, 2014; Hamelin et al., 2015; Johnson et al., 2016; Liu et al., 2014; Sunderland et al., 2009). Dissolved organic matter (DOM) and dissolved organic carbon (DOC) concentrations may also influence the

production of MeHg as they can serve as carriers of Hg during transport from watershed to a lake and as an important ligand in solutions (Aiken et al., 2011; Driscoll et al., 1995; Munthe et al., 2007). Abiotic factors may affect the supply of electron acceptors to mercury-methylating zones and therefore impact methylation rates (Todorova et al., 2009). For instance, increased inputs of  $\text{NO}_3^-$  and  $\text{O}_2$  suppress  $\text{SO}_4^{2-}$  reduction, decreasing net MeHg (Todorova et al., 2009), while an increase of  $\text{SO}_4^{2-}$  in environments with currently low concentrations of  $\text{SO}_4^{2-}$  increases the production of MeHg (Gilmour et al., 1992). However, MeHg levels will decrease in environments with currently high concentration of  $\text{SO}_4^{2-}$  due to the formation of mercuric sulfide complexes that are less available for reduction (Benoit et al., 2003). Biotic processes of mercury methylation mediated by anaerobic microorganisms are the primary forms of methylation in aquatic environments (Celo et al., 2006). However, mercury also has the potential to be methylated by abiotic transmethylation caused by ultraviolet radiation (Morel et al., 2002). Mercury methylation can also occur in oxic waters in oceans (Blum et al., 2013).

Net MeHg rates are influenced by both mercury methylation and demethylation. Mercury demethylation can occur in the environment at the same time, and has the potential to be mediated by anaerobic microorganisms and abiotic factors. Two forms of microbial mercury detoxification are described and include oxidative demethylation and reduction by mercury resistant *mer* operon-mediated pathways composed of the *merA* and *merB* genes (Barkay et al., 1989; Figueiredo et al., 2018; Ni Chadhain et al., 2006; Oremland et al., 1995; Parks et al., 2009; Robinson & Tuovinen, 1984; Schaefer et al., 2004), while abiotic methods include photodegradation by ultraviolet light, which has been shown to occur in water of the Everglades (Li et al., 2010).

Mercury methylation and demethylation rates have been studied in numerous environments in which MeHg is thought to be most abundant. Originally, sediments were thought to be the primary site of mercury methylation (Matilainen et al., 1991; Rudd & Furutani, 1980). However more recent research has implicated other ecosystems such as periphyton biofilms (Acha et al., 2005; Cleckner et al., 1999; Desrosiers et al., 2006; Hamelin et al., 2011; Mauro et al., 2002), and in the water column as important sites of mercury methylation (Mauro et al., 2002). Initial experiments conducted in sediments primarily focused on sulfate-reducing

bacteria (SRB), as they were initially thought to be the main anaerobic bacteria responsible for mercury methylation. Sulfate-reducing bacteria were first implicated in MeHg production through the addition of molybdate, an inhibitor of sulfate reducers, to sediment samples of an acid-stressed lake in the Muskoka-Haliburton area of Ontario, Canada, and to sediments from the Cheesequake estuary in Northern New Jersey. Both studies found that mercury methylation was inhibited following the addition of molybdate, up to as much as 95% in sediments (Compeau & Bartha, 1987; Kerry et al., 1991). However, sulfate-reducing bacteria only had a partial effect on methylation rates in periphyton from the Bolivian Amazon Region, indicating that there were potentially other mercury methylators (Achá et al., 2011). Iron-reducing bacteria (IRB) were not implicated in contributing to net MeHg levels until the early 2000's. One study isolated the iron-reducing bacterium *Geobacter* sp. Strain CLFeRB from sediment samples collected from Clear Lake, CA, USA, and discovered that it was able to methylate mercury at rates comparable to known sulfate-reducing bacteria (Fleming et al., 2006). Another study found that methylation rates were stimulated following the addition of low levels of amorphous iron(III) oxyhydroxide in river sediments of South River, VA, USA (Yu et al., 2012).

Another group of anaerobic microorganisms was implicated in mercury methylation following studies conducted on periphyton. A study conducted on periphyton biofilms from a shallow fluvial lake located along the St. Lawrence River, QC, Canada, found that methylation rates were completely inhibited following the addition of BESA (2-bromoethane sulfonic acid), a methanogenesis inhibitor, implicating methanogens as potentially the main methylators in periphyton biofilms (Hamelin et al., 2011). However, another study found that co-inhibition experiments of both SRB and methanogens was the most effective at inhibiting the production of MeHg in tropical environments, suggesting that mercury methylation may be a product of complex interactions of prokaryotic SRB and methanogens (Correia et al., 2012).

### **1.3 Genetic basis for mercury methylation**

Until recently, potential mercury methylating microorganisms were identified through metabolism inhibition experiments and the use of pure cultures. However, the discovery of a two-gene cluster responsible for mercury methylation led to the detection of many more taxa

involved. The *hgcA* and *hgcB* genes are key genetic elements responsible for the methylation of mercury (Parks et al., 2013). The *hgcA* gene encodes a corrinoid iron-sulfur protein (CFeSP), which is involved in the reductive acetyl coenzyme A (acetyl-CoA) carbon fixation pathways and is capable of transferring a methyl group to Hg(II), whereas *hgcB* encodes a small iron-sulfur cluster-containing ferredoxin protein, and is capable of returning HgcA to a redox state enabling it to receive a new methyl group (Obrist et al., 2018; Parks et al., 2013). The proteins encoded by both genes are essential for the methylation of mercury. Microorganisms in which either gene has been altered or deleted do not methylate mercury (Parks et al., 2013). Following the discovery of the *hgcAB* gene pair, studies were conducted to determine if the presence of this gene was a reliable predictor of mercury methylation capability. It was found that *hgcAB* could be used to accurately predict a microorganisms ability to methylate mercury (Gilmour et al., 2013). These genes were later identified in many organisms across various environments including previously unknown taxa as mercury methylating microorganisms through a metagenomic study (Podar et al., 2015).

The newly (through the discovery of the *hgcAB* gene pair) identified taxa were used to create clade-specific and broad-range probes to target the genes in environmental samples. One study developed probes specifically targeting the *hgcA* gene to determine the phyla present in the northern Everglades and discovered *Delta-Proteobacteria* to be the most dominate class of mercury methylating bacteria (Bae et al., 2014). Another study, also focusing on *hgcA*, reported *hgcA* diversity in environmental samples from various parts of Everglades wetlands, suggesting a role in mercury methylation of various taxonomic or metabolic groups (Schaefer et al., 2014). *hgcA* gene abundance was measured for rice paddy soils and a correlation was observed with the methylmercury levels, supporting the notion that microbes containing the *hgcA* genes contribute to mercury methylation (Liu et al., 2014). Later, a study used the known taxa as well as previously designed probes in order to create broad-range probes that could be used on environmental samples. These probes had the ability to amplify *hgcAB* from taxa across all phylogenetic groups currently known to methylate mercury (Christensen et al., 2016).

#### **1.4 Altered landscapes and their impact on mercury methylation**

Extensive research has been conducted on human activities and altered landscapes and their impact on mercury transport, methylation and demethylation rates and bioaccumulation of MeHg in the environment. These anthropogenic sources include industrial contamination, mining operations, urbanization, rice paddy fields, forestry operations, and wetland and reservoir creation and its watershed management. These sites are not only responsible for the re-mobilization and transport of Hg(II) between ecosystems, they can also be primary sites for Hg(II) methylation (Hsu-Kim et al., 2018). Historically, mining operations have been a major contributor to Hg emissions into the environment and account for downstream Hg inputs from practices such as artisanal small-scale gold mining (ASGM), which uses mercury for amalgamation and extraction of gold from ores (Hsu-Kim et al., 2018). ASGM has also been used in large scale mining operations to aid in the extraction of gold and silver from ores, and include sites such as the coastal range of California (Rytuba, 2000), large Hg mines in Almaden, Spain (Barre et al., 2016), Hg mining region of Idrija, Slovenia (Foucher et al., 2009), and large-scale cinnabar refining in the Andean highlands of Huancavelica, Peru (Hagan et al., 2015). Not only do these mining activities contribute to atmospheric Hg emissions, they can also transport Hg to local aquatic environments or from the waste of mining pits if flooded, when they are no longer in use, and provide an environment rich in sulfur and low organic carbon levels, creating a habitat conducive to mercury methylation (Gammons et al., 2013; Hsu-Kim et al., 2018).

Similarly, to mining operations, forestry operations or forest fires increase the amounts of Hg released into the environment. Logging activities increase soil erosion and water run-off that intensify Hg downstream to local aquatic systems (Béliveau et al., 2009; Comte et al., 2013; Garcia & Carignan, 2005; Garcia et al., 2007). The use of heavy machinery during these practices also compacts the soils, creating opportunities for flooding and the release of Hg into small pools of water which will eventually impact downstream environments (Hsu-Kim et al., 2018).

Agricultural wetlands are another primary site of mercury methylation (Ackerman et al., 2010). Wetlands such as rice paddy fields provide environments suitable for anaerobic

microorganisms facilitating the growth of mercury methylating microbes. These fields create a major public health concern as the MeHg bioaccumulates within the rice as well as at higher trophic levels such as fish found in the fields (Ackerman et al., 2010; Meng et al., 2011).

One of the most common anthropogenic manipulations of environments impacting MeHg is the impoundment of rivers and the construction of reservoirs (Hsu-Kim et al., 2018). Reservoirs may be constructed to provide flood control, water irrigation systems for agricultural purposes or used to generate electricity through the use of run-of-river dams (Hsu-Kim et al., 2018). Newly flooded reservoirs on land previously covered with vegetation and soils rich in organic matter, turn previously oxic soils into newly anoxic environments creating a habitat conducive to the growth of anaerobic microorganisms capable of methylating mercury, resulting in an increase of MeHg levels (Hall et al., 2009; Kelly et al., 1997). Mercury can also be released into the water column as sediments erode through the breakdown of organic matter (Eckley et al., 2017). However, in the main input of mercury from the watersheds to aquatic systems in boreal lakes is through dissolved organic carbon (Bravo et al., 2017). Fish present in reservoirs can have elevated Hg concentrations, especially large piscivorous fish, such as walleye (*Sander vitreus*) and northern pike (*Esox lucius*) in boreal Canadian reservoirs (*Aménagements hydroélectriques de la Chute-Allard et des Rapides-des-Cœurs*, 2014; Bodaly et al., 2007; Kamman et al., 2005). Hg levels in these fish could increase three- to six-fold following the flooding of the reservoir with the Hg levels remaining constant for several decades later (*Aménagements hydroélectriques de la Chute-Allard et des Rapides-des-Cœurs*, 2014; Bodaly et al., 2007; Kamman et al., 2005).

### **1.5 Chute-Allard run-of-river hydroelectric generating station**

In an attempt to reduce the impacts large hydroelectric reservoirs have on the environment, Hydro Québec began installing run-of-river hydroelectric generating stations. These new run-of-river dams do not require the impoundment or flooding of large landscape areas and were therefore not expected to have a severe impact on the amount of Hg being released into the environment and subsequently transformed into MeHg. One such station, at Chute-Allard (CA; 47.892910, -73.718247), constructed on the St. Maurice River (QC, Canada)

downstream of the Reserve of Wemotaci in 2008. Unexpectedly, following the construction of this dam, monitoring studies revealed that MeHg levels in fish were much higher than the MeHg levels in fish between 1990 and 1993, prior to the construction of the hydroelectric generating station (*Aménagements hydroélectriques de la Chute-Allard et des Rapides-des-Cœurs*, 2014). However, it is not known what specifically caused the increase of MeHg in the local fish populations. At the time of the construction of the Chute-Allard generating station, large forest fires decimated vegetation around the St. Maurice river near where the of construction, and logging operations were taking place in the area.

In an attempt to better understand the cause of the increased levels of MeHg, a multiscale interdisciplinary research project was organized to understand how the construction of the Chute-Allard run-of-river hydro facility could have altered the pathways responsible for the mobilization, transformation and transfer of mercury in the St. Maurice river, QC, Canada. In particular, this thesis focuses on mercury methylation associated with periphyton biofilms collected from three sampling sites upstream of the Chute-Allard run-of-river hydro facility. The three sampling sites include periphyton biofilms collected from a natural site, which is close to the original river path, the flooded site which is in an area that was flooded following the construction of the hydroelectric dam, and the artificial wetland site which was in a channel constructed prior to the construction of the run-of-river dam to compensate for the flooding and provide new spawning grounds for local fish communities. Periphyton is a complex community consisting of algae, bacteria, fungi, detrital particulate organic matter, and microinvertebrates (Correia et al., 2012; Desrosiers et al., 2006). Periphyton is thought to be a key site for mercury methylation since the biofilms have strong spatial and temporal redox conditions. Thus, it can support active microbial methylating bacteria (Olsen et al., 2016). Methylation occurs in oxic-anoxic interfaces within the periphyton (Olsen et al., 2016). These small pockets create the perfect habitat for anaerobic mercury methylating microorganisms such as sulfate-reducing bacteria (SRB), iron-reducing bacteria (IRB), and methanogens to grow (Fleming et al., 2006; Hamelin et al., 2011). It is important to understand MeHg levels in periphyton biofilms, as they are at the bottom of the food web, hence it has the ability to transfer Hg to organisms at higher trophic levels. Environments within the St. Maurice river conducive to the formation and growth

of periphyton biofilms on different substrata may have the greatest diversity of potential mercury methylating microorganisms, and therefore the highest potential for increased MeHg production.

In this thesis I hypothesized that: 1) mercury methylation will occur at highest rates in the artificial wetland channels due to a greater abundance of organic matter and anaerobic niches for methylation to occur, and 2) there will be greater abundance and diversity of microorganisms with the potential to methylate mercury in the artificial wetland channels due to the greater abundance of organic matter. To address these hypotheses, I 1) compared the microbial community composition within and between periphyton biofilms, 2) assessed the natural diversity of mercury methylating microorganisms within periphyton biofilms, and 3) quantified the abundance of mercury methylating microorganism at each sampling site in relation to mercury methylation and demethylation rates. In order to address these hypotheses and objectives, analyses were conducted using the 16S rRNA gene (rDNA) and transcripts (rRNA), functional gene markers, and metagenomics.

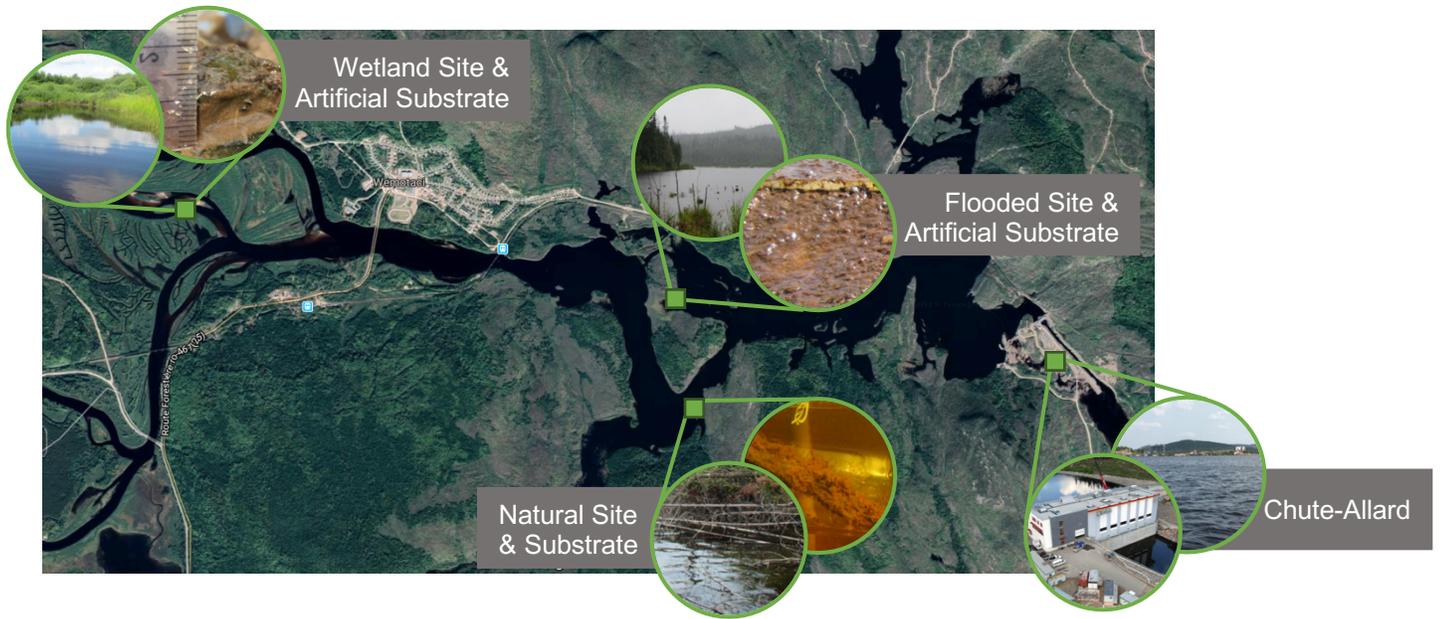
## **2. Materials and Methods**

### **2.1 Field sampling in the St. Maurice River, QC, Canada**

In order to investigate how the construction of an artificial wetland close to the Chute-Allard hydroelectric generating station and, subsequent flooding of the surrounding landscapes could have impacted MeHg levels, we measure methylation and demethylation rates. Hg methylation and demethylation rates were quantified using stable mercury isotope tracers. The tracers were added to enclosed periphyton biofilms and incubated *in situ*. At the end of the incubation, the periphyton biofilms were collected for methylation and demethylation rate analysis, as well as genetic and metagenomic analyses.

Periphyton biofilms were collected at three locations upstream of the Chute-Allard Run-of-River Hydro Electric Station along the St. Maurice River by Maxime Leclerc from Université de Montréal. The three locations are identified as the natural site (S1) (47.885464, -73.754885), the flooded site (S2) (47.893588, -73.756927), and the artificial wetland site (S3) (47.899882, -

73.808857) (Figure 1). Periphyton biofilms from natural substrate were collected from the natural site (S1), which is close to the original river path, while periphyton biofilms from artificial substrates colonized for one year were collected from both the flooded (S2) and wetland sites (S3) (Figure 1). The S2 site is in an area that was flooded following the construction of the hydroelectric dam, and S3 was in a channel constructed prior to the construction of the run-of-river dam to compensate for the flooding and provide new spawning grounds for local fish communities. The three sampling sites were chosen as they represent what occurs in a river following the construction of a dam. The natural sampling site is a control for the microbial community composition found in periphyton biofilms as they are the least affected by the flooding of the surrounding landscapes. The flooded sampling site allows us to study how the microbial community composition changes following the construction and subsequent flooding of the river and its potential impact on mercury methylation rates. The artificial wetland sampling site provides a location to study how changes in landscape compensating for loss of habitat increased available substrata for the growth of periphyton biofilms and how they could impact the microbial community diversity, mercury methylation and demethylation rates, and the cycle of mercury in the system.



**Figure 1.** Map of sampling sites upstream of the Chute-Allard hydroelectric generating station, QC, Canada. Sampling site one is a portion of the natural river where periphyton was collected from natural substrates (Natural). Sampling site two is a part of the river that flooded following the construction of the Chute-Allard hydroelectric generating station and periphyton samples were collected from artificial substrates left the year prior (Flooded). Sampling site three is in an artificial wetland channel made prior to the construction of the dam for water flow from the flooding following the construction of the dam (Wetland). The periphyton in the wetland site was collected from artificial substrate left during the sampling season of the previous year.

Following the collection of the periphyton biofilms, six replicates in triplicate of each of the three treatments were incubated *in situ* at each of the three sampling sites ( $n = 54$  samples). For the first manipulation the substrates were flash frozen at time zero with no incubation or addition of mercury isotopes ( $t_0$ ). For the second manipulation, the substrates were collected and then incubated *in situ* for 48 hours with no mercury isotope addition (E). The substrates in the third treatment were spiked with 4 ng/L of  $^{198}\text{Hg}$  and 4 ng/L of  $\text{Me}^{200}\text{Hg}$ , and then incubated *in situ* for 48 hours (F). Following the incubation period, the substrates were removed, filtered and flash frozen to preserve the samples for future analyses. Triplicate samples of each treatment from each sampling site ( $n = 27$ ) were acidified to stop reactions and brought to Université de Montréal to calculate methylation and demethylation rate results. The remaining samples ( $n = 27$ ) were brought to Concordia University for genomic and metagenomic analyses.

## 2.2 Methylation and demethylation rates

Total mercury (Hg) and methylmercury (MeHg) were measured using stable isotopes 198, and 200, from the 27 periphyton samples collected from the St. Maurice River in 2017. The measured concentrations of each isotope were corrected to remove background noise from natural isotopic distributions and make the different analytical runs comparable. The correction was applied using a matrix of the ratio of abundance of 198, 200 and 202 in both 198 and 200-enriched Hg solution and using the natural ratio abundance of mercury isotopes to reduce the natural (prior to spike) mercury isotope signal. The true inorganic concentrations were obtained through by  $\text{Inorganic Hg} = \text{Total Hg} - \text{MeHg}$ , as methylmercury is assumed to be the most abundant organic form of mercury in freshwater samples. The concentrations from the periphyton incubated for 48 hrs without the addition of Hg enriched isotope solution (E) (i.e. Hg concentration in the periphyton of this area before the addition of  $\text{Me}^{198}\text{Hg}$  and  $^{200}\text{Hg}$ ) were used for the base level of Hg for a given site to calculate methylation ( $K_m$ ) and demethylation rates ( $K_d$ ). In order to create the base level, the spiked Hg was added for the calculation of  $t_0$  concentrations. To do so, the amount of Hg added in each bottle has been considered to be evenly and totally absorbed and distributed in the biofilm matrix. Concentrations were always applied in function of the actual dry weight of a given bottle. Methylation rates ( $K_m$ ) and demethylation rates ( $K_d$ ) were calculated by:

$$K_m = -\ln(1 - t_x \text{Me}^{200}\text{Hg}/t_0^{200}\text{Hg})/t$$

$$K_d = -\ln(1 - t_x \text{Me}^{198}\text{Hg}/t_0^{198}\text{Hg})/t$$

### 2.3 DNA and RNA extractions

The remaining 27 periphyton samples were stored on dry ice, divided in half and separated into two 1.5 mL Eppendorf tubes. One Eppendorf tube designated for DNA extraction and analysis and the other for RNA extraction and analysis resulting in 27 samples of each gDNA and rRNA. Genomic DNA was extracted from each triplicate of each manipulation using the DNeasy PowerWater Kit (Qiagen) with a few modifications. Following the addition of 1 mL of solution PW1 to the PowerWater bead tube in step 5, the bead tubes were placed horizontally in a 65 °C water bath for 10 minutes. 1 µl of RNase was added to each sample and incubated at 37 °C for 30 minutes following the transfer of supernatant in step 9. Finally, the DNA was resuspended in 50 µl of solution PW6 rather than the prescribed 100 µl. Ribosomal RNA was extracted from each triplicate of each manipulation using the RNeasy PowerWater Kit (Qiagen) with a few modifications. Following the addition of solution PM1/β-ME in step 5, the bead tubes were placed horizontally in a 65 °C water bath for 10 minutes. The RNA was then resuspended in 50 µl of RNase-Free Water.

### 2.4 16S rRNA gene and transcript analysis

#### 2.4.1 16S rRNA amplification

To ensure the presence of DNA in the genomic samples, PCR amplification of the 16S rRNA gene was completed using forward primer 27F (5' AGAGTTTGATCCTGGCTCAG 3') and reverse primer 1492R (5' GGTTACCTTGTTACGACTT 3'). The 25 µL PCR reaction consisted of 15.5 µL autoclaved milliQ water, 5 µL of 5x Phusion Reaction HF Buffer, 0.5 µL dNTPs (10 mM), 0.5 µL Phusion HF DNA Polymerase (2000 U/mL), 1.25 µL Forward Primer, 1.25 µL Reverse Primer, and 1 µL of extracted periphyton DNA. 1 µL of autoclaved milliQ dH<sub>2</sub>O was used as a template for the negative reaction and 1 µL of isolated *Geobacter*

*sulfurreducens PCA* was used as a template for the positive control reaction. The reactions were incubated at 98°C for 1 minute, and cycled 35 times at 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 20 seconds, then incubated at 72°C for 5 minutes. 1 µL of loading dye was added to each sample and 5 µL was loaded and run on a 1 % agarose gel containing 0.03 µL/mL ethidium bromide for 45 minutes at 90 volts. A GeneRuler 1kb DNA Plus Ladder (Thermo Scientific) was used to compare the size of the PCR products against the known sizes of the bands within the ladder.

The extracted gDNA from the periphyton samples were then prepared for 16S rRNA gene sequencing of the V4 region targeting both bacterial and archaeal rRNA using the primers 515 FB (5' GTGYCAGCMGCCGCGGTAA 3') and 806 RB (5' GGACTACNVGGGTWTCTAAT 3') (Apprill, McNally, Parsons, & Weber, 2015; Parada, Needham, & Fuhrman, 2016). The 25 µL PCR reaction consisted of the same mixture as described previously and was cycled using the same conditions, with the addition of the new modified primers. Following the PCR, 1 µL of loading dye was added to each sample and 5 µL was loaded and run on a 1 % agarose gel containing 0.03 µL/mL ethidium bromide for 45 minutes at 90 volts. A GeneRuler 100 bp DNA Plus Ladder (ThermoFisher Scientific) was used to compare the size of the PCR products against the known sizes of the bands within the ladder. Once the proper band size was confirmed, the PCR was re-done, however this time with the CS1 (5' AACTGACGACATGGTTCTACA 3') and CS2 (5' TACGGTAGCAGAGACTTGGTCT 3') adapters added to the front of the forward and reverse primers respectively in preparation for 16S rRNA gene sequencing by Genome Québec (McGill University) resulting in CS1 515 BAF (5' AACTGACGACATGGTTCTACAGTGYCAGCMGCCGCGGTAA 3') and CS2 806 BAR (5' TACGGTAGCAGAGACTTGGTCTGGACTCANVGGGTWTCTAAT 3'). The PCR was completed and visualized following the same methods above, however the annealing temperature was 62 °C rather than 55 °C as before. The remaining 20 µL PCR product was stored at -20 °C until being sent for sequencing.

### 2.4.2 16S rRNA gene and transcript amplicon sequencing

The Extracted RNA from the 27 periphyton samples was reverse transcribed using sequence-specific TaqMan Reverse Transcription Reagents Kit (Applied Biosystems). A 10  $\mu\text{L}$  reaction was created for each sample consisting of 0.85  $\mu\text{L}$  RNase-free water, 1.0  $\mu\text{L}$  10X TaqMan RT Buffer, 2.2  $\mu\text{L}$  25 mM Magnesium Chloride, 2.0  $\mu\text{L}$  deoxyNTPs Mixture, 0.5  $\mu\text{L}$  806 RB (5' GGACTACNVGGGTWTCTAAT 3'), 0.2  $\mu\text{L}$  RNase Inhibitor, 0.25  $\mu\text{L}$  MultiScribe Reverse Transcriptase (50 U/ $\mu\text{L}$ ), and 3.0  $\mu\text{L}$  periphyton sample RNA. The reactions were then placed in the thermocycler at 48 °C for 30 minutes for the RT step and then the temperature was increased to 95 °C for 5 minutes to inactivate the reverse transcriptase. 1  $\mu\text{L}$  of the resulting cDNA was then amplified using CS1 515 BAF and CS2 806 BAR 16S rRNA primers targeting both prokaryotic and archaeal DNA from the V4 region. The 25  $\mu\text{L}$  reaction consisted of 15.5  $\mu\text{L}$  autoclaved milliQ water, 5  $\mu\text{L}$  of 5x Phusion Reaction HF Buffer, 0.5  $\mu\text{L}$  dNTPs (10 mM), 0.5  $\mu\text{L}$  Phusion HF DNA Polymerase (2000 U/mL), 1.25  $\mu\text{L}$  Forward Primer, 1.25  $\mu\text{L}$  Reverse Primer, and 1  $\mu\text{L}$  of periphyton cDNA. The blank created during reverse transcription was used as the negative control for the reaction. The reactions were incubated at 98°C for 1 minute, and cycled 35 times at 98°C for 10 seconds, 55°C for 30 seconds, and 72°C for 20 seconds, then incubated at 72°C for 5 minutes. 1  $\mu\text{L}$  of loading dye was added to each sample and 5  $\mu\text{L}$  was loaded and run on a 1 % agarose gel containing 0.03  $\mu\text{L}/\text{mL}$  ethidium bromide for 45 minutes at 90 volts. A GeneRuler 1kb DNA Plus Ladder (Thermo Scientific) was used to compare the size of the PCR products against the known sizes of the bands within the ladder. The remaining 20  $\mu\text{L}$  of PCR product was stored at -20 °C until being sent for sequencing.

Following the amplification of the V4 region of the 16S rRNA genes with the CS1 515 BAF and CS2 806 BAR primers and adapters from both the 27 gDNA and 27 cDNA templates for a total of 54 periphyton samples, the remaining 20  $\mu\text{L}$  PCR products were sent to Genome Québec (McGill University) for 16S rRNA gene and transcript amplicon sequencing using the Illumina MiSeq as a high-throughput sequencing platform producing 250 bp paired-end reads with a 10 million read depth of sequencing split across a 96 well plate. The fastq files were downloaded from Nanuq and imported into RStudio where the demultiplexed paired-end files were run through the DADA2 Pipeline 1.6 ([https://benjjneb.github.io/dada2/tutorial\\_1\\_6.html](https://benjjneb.github.io/dada2/tutorial_1_6.html))

(Callahan et al., 2016) to produce amplicon sequence variant (ASV) tables as well as a table containing the taxonomy of the output sequences (Callahan et al., 2016). Quality profiles were created for each forward and reverse fastq sequence and then trimmed at 220 and 180 for the forward and reverse reads respectively. The first 19 bp and the first 20 bp were also removed from the forward and reverse reads to remove the primers 515BAF and 806BAR primers respectively. The forward and reverse reads were then dereplicated to combine all identical sequences into “unique sequences” with the corresponding abundance. Real sequence variants were then identified by applying the core sequence-variant inference algorithm and the paired forward and reverse reads were merged. A sequence table was then constructed, and the chimeras were removed. Taxonomy was assigned down to the species level using the `silva_nr_v132_train_set.fa.gz` and `silva_species_assignment_v132.fa.gz` (Pruesse et al., 2007). The final products include an ASV table and a taxa table containing the information from all 54 periphyton samples, 27 from gDNA and 27 from cDNA.

Using Phyloseq, a package in R that analyses microbiome data, the ASV and taxa tables were combined to create a phyloseq object, a matrix containing all of the 16S rDNA and rRNA information (McMurdie & Holmes, 2013). The resulting matrices were used to analyze, and visualize 16S rRNA gene community diversity, abundance and community statistics of both the ribosomal and transcriptomic RNA of the periphyton communities.

### **2.4.3 16S rRNA alpha diversity**

To investigate the microbial species diversity of the periphyton biofilms communities, an alpha diversity plot was created using both 16S rDNA and rRNA amplicon data using the Simpson index measure from the phyloseq package in R (McMurdie & Holmes, 2013; Simpson, 1949).

### **2.4.4 Diversity, abundance and variability of microorganisms**

Microbial diversity and abundance of all 16S rDNA and rRNA ASVs were visualized at the phylum level using a heatmap created in ggplot2 (Figure 4) (Wickham, 2016). To further

explore the diversity and variability of microbial species present, a principal coordinate analysis (PCoA) using the Bray-Curtis dissimilarity matrix was used (McMurdie & Holmes, 2013). The Bray Curtis Principal Coordinate Analysis (PCoA) plot was constructed for both 16S rDNA and 16S rRNA data to compare the ASVs at each sampling site to the treatments using the phyloseq package in R (Figure 5) (McMurdie & Holmes, 2013). A Permutational Multivariate Analysis of Variance Using Distance Matrices or multivariate analysis of variance (MANOVA) was completed in conjunction with the principal coordinate analysis plots for both the 16S rDNA and rRNA amplicons from the periphyton biofilms, using the Adonis function from the vegan package in R (Oksanen et al., 2013).

#### **2.4.5 Taxonomic analysis of 16S rRNA amplicons**

Taxonomic analyses of the 16S rDNA and rRNA amplicons were conducted using the phyloseq package and then visualized using the ggplot2 package in R to determine abundance and diversity of the microbes present in the periphyton biofilm communities at each sampling site (McMurdie & Holmes, 2013; Wickham, 2016). Initially, the average percent read composition was calculated using the three replicates at each sampling site for both the 16S rDNA and rRNA amplicons at various taxonomic levels and visualized in bar plots (Figure S1, Figure S2). Specific families from groups known to methylate mercury (i.e., *Chloroflexi*, *Euryarchaeota*, *Firmicutes*, and *Delta-Proteobacteria*) were isolated and plotted in bar graphs to visualize the average percent read composition of potential mercury methylating families from the replicates at each sampling site for both the 16S rDNA and rRNA amplicons (Figure 6).

#### **2.4.6 Indicator species analysis**

An indicator species analysis was conducted using the multipatt function from the Indicspecies package in R on both 16S rDNA and rRNA amplicons to determine the presence and abundance of indicator species at each sampling site (De Cáceres & Legendre, 2009). The number of indicator species were calculated for each replicate at all three sampling sites and then averaged for each sampling site using the 16S rDNA and rRNA amplicons. The number of

indicator species at the sampling sites for both 16S rDNA and rRNA amplicons was plotted into various bar graphs separated by groups previously implicated in mercury methylation (Figure 7).

## **2.6 Metagenomic analysis**

### **2.6.1 Metagenomic sequencing**

150 ng of DNA from 9 periphyton samples (the first replicate from each manipulation of each sampling site) were sent to Genome Québec (McGill University) for high-throughput Shotgun Illumina NovaSeq 6000 S4 sequencing of the entire sample, resulting in 150 bp long paired-end reads. The Md5 sum and fastq files were downloaded for both the forward and reverse reads, and the Md5 sum files were checked to ensure proper downloading. The paired-end reads were filtered and trimmed using trimmomatic-0.38 (20:30:10 LEADING: 3 TRAILING: 3 SLIDINGWINDOW: 4:15 MINLEN: 36) (Bolger et al., 2014). The filtered files were then assembled individually and into coassemblies of each sampling site (Natural, Flooded, and Wetland) using MEGAHIT v1.0.6 (--k-list 23, 43, 63, 83, 103, 123), producing final contig fasta files (Li et al., 2015). The unassembled reads were then mapped to the assembled scaffolds using the Burrows-Wheeler Alignment Tool (bwa) to determine coverage (Li & Durbin, 2010). The assembled fasta files were indexed using bwa index (-a bwtsv, to accommodate large files), and then alignments were seeded with maximal exact matches (MEMs) using the bwa mem command (-h, to reduce the redundancy of output) (Li & Durbin, 2010). Samtools view was used to convert the files from sam to bam (using default settings) (Li et al., 2009).

The assembled fasta files submitted to the Joint Genome Institute Integrated Microbial Genomes and Microbiomes (JGI IMG/M) for annotation along with the .bam files to determine coverage.

### **2.6.2 Taxonomic analysis**

Taxonomic analysis of the assembled metagenomes was conducted using MEGAN6 (Community Edition, version 6.13.5, built 7 Jan 2019) (Huson et al., 2016) to visualize the

microbial community structure and abundance of each replicate at each sampling site. Prior to the analysis in MEGAN6, the protein files (.faa) produced by the JGI IMG/M were taxonomically assigned by conducting a blastp search using Diamond\_09 (Buchfink et al., 2014) against the NCBI refseq\_protein database (refseq\_nr\_december2018.dmnd) and taxonomy was assigned using the prot\_acc2tax-Nov2018X1.abin file. The resulting diamond files were imported into MEGAN6 (Community Edition, version 6.13.5, built 7 Jan 2019), and the absolute gene counts were compared for all replicates at each sampling site. Initially, the absolute number (square-root scale) of genes were visualized at the phylum level (Figure S3). The absolute number of protein coding genes at various taxonomic levels were visualized for groups previously implicated in mercury methylation for each sampling site (Figure 8).

### 2.6.3 Potential *hgcA* and *hgcB* genes in assembled metagenomes

To determine if there were any *hgcA* or *hgcB* genes present in the assembled metagenome data, blastp searches from the Basic Local Alignment Search Tool (BLAST) from NCBI were conducted using known *hgcA* and *hgcB* protein sequences from 60 species against the assembled protein sequences produced by the JGI IMG/M (Podar et al., 2015). blastp searches were conducted at various stringency levels for each replicate by adjusting the e-value;  $\times 10^{-5}$ ,  $\times 10^{-10}$ ,  $\times 10^{-15}$ ,  $\times 10^{-20}$ ,  $\times 10^{-25}$ ,  $\times 10^{-30}$ ,  $\times 10^{-35}$ ,  $\times 10^{-40}$ , and  $\times 10^{-45}$ , for both the *hgcA* and *hgcB* genes, and the number of unique hits were averaged for each sampling site and the mean and standard deviation was plotted in a bar graph (Figure 9). To better visualize the alignment of potential *hgcA* genes isolated from the assembled metagenomic data (e-value  $\times 10^{-40}$ ), the 17 potential *hgcA* protein sequences of the artificial wetland site from the blastp outputs were aligned to 16 reference *hgcA* protein sequences (Figure 10). Sequences that did not align to the conserved regions of the reference *hgcA* sequences in the putative cobalamin-binding domain were removed. The remaining sequences from the artificial wetlands were aligned with 60 known *hgcA* protein sequences using MEGA6 (aligned by muscle), and a maximum likelihood phylogenetic tree was produced using the Jones-Taylor-Thornton (JTT) model for amino acids to determine what group of mercury methylators the potential *hgcA* genes belong to (Figure 11).

#### **2.6.4 Potential *hgcA* genes in the unassembled metagenome**

To further identify potential *hgcA* genes that may not have been properly assembled, a tblastn search was conducted. A tblastn search aligns protein sequences to a nucleotide database translated in all six open reading frames. Known *hgcA* protein sequences were blasted against the replicates from all three sampling sites at various levels of stringency (i.e.,  $\times 10^{-10}$ ,  $\times 10^{-15}$ ,  $\times 10^{-20}$ ) and the number of unique hits were averaged for each sampling site with the mean and standard deviation plotted in a bar graph (Figure 12). Results of the replicates from the artificial wetland site from the tblastn search (e-value  $\times 10^{-20}$ ) were adjusted and aligned to a subset of known *hgcA* protein sequences, and the alignment was created in excel to visualize the alignment of potential *hgcA* genes to known *hgcA* protein sequences (Figure 13).

#### **2.6.5 Functional gene markers associated with mercury methylation**

Functional gene markers *drsA* (sulfite reductase alpha subunit), *merA* (mercuric reductase), and *mcrA* (methyl-coenzyme M reductase alpha subunit), were identified for each sample at the sampling sites from the JGI IMG/M data set. The number of functional gene markers were plotted in a bar graph (Figure 14).

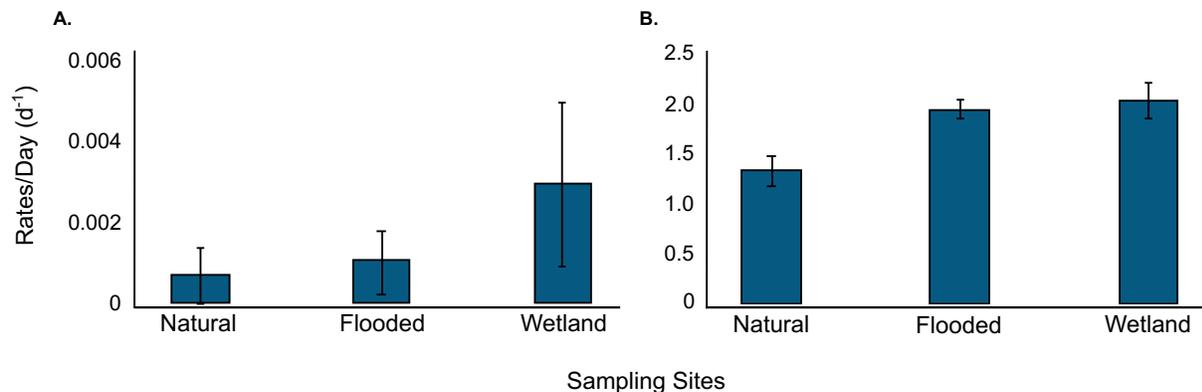
### **3. Results**

Periphyton biofilms were collected from three sampling sites upstream of the Chute-Allard hydroelectric dam. The first sampling site is located in the natural portion of the river, unaffected by construction of the dam, the second from a portion of the river that was flooded following the construction of the dam, and the final sampling site, an artificial wetland channel was created by Hydro Québec in order to promote fish spawning.

#### **3.1 Methylation and demethylation rate results**

Average mean and standard error were calculated for the F1 (Incubated *in situ* for 48 hrs following mercury isotope addition) for each replicate of each sampling site in a study conducted

by Maxime Leclerc. The average mean and standard error were plotted for both the mercury methylation ( $K_m$ ) (Figure 2a) and mercury demethylation ( $K_d$ ) (Figure 2b). There is a significant difference between the  $K_m$  rates for Natural (S1) and Flooded (S2) compared to the wetland sampling site (S3) however, the natural and flooded sites are not significantly different from each other (Figure 2a). For the  $K_d$  rates, there is a significant difference based on standard error (ANOVA followed by two-sample t-test) for the natural site compared to the flooded and wetland sites, however, there is no observed significant difference between the flooded and wetland sites (Figure 2b).



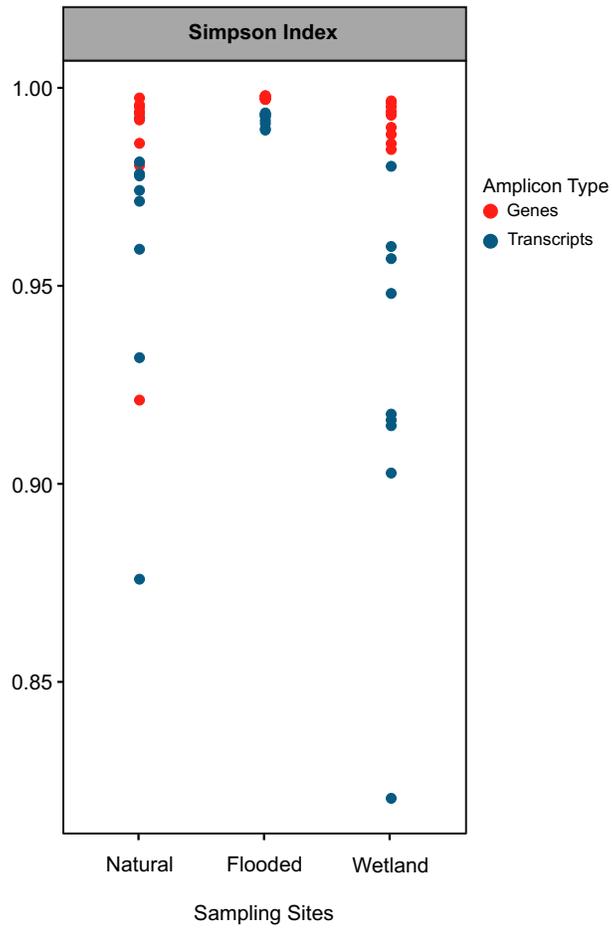
**Figure 2A.** Mean and standard error of mercury methylation rates per day at each sampling site. Methylation rates between all three sampling sites are not statistically significant from each other, however there is an observed difference between the natural and flooded sampling sites and the wetland sampling site. **B.** Mean and standard error of mercury demethylation rates per day for each sampling site. The natural sampling site is statistically different from both the flooded and wetland sites; however, the flooded and wetland sites are not statistically significant from each other.

## **3.2 16S rDNA and rRNA analysis**

The composition and variability of microbial communities present in and between periphyton biofilms were investigated through the sequencing of 16S ribosomal genes and transcripts. A total of 3,739,141 reads, 2,227,245 from the 16S ribosomal genes and 1,511,896 from the 16S ribosomal transcripts, totalling 19,044 unique ASVs, 14,221 genes and 10,083 transcripts. The amount of unique ASVs associated with 16S rDNA analysis averaged 1,938 across all three sampling sites with an average read count of 82,490. There was an average 1,177 unique ASVs from the rRNA data across the three sampling sites and an average read count of 55,996.

### **3.2.1 Alpha diversity analysis**

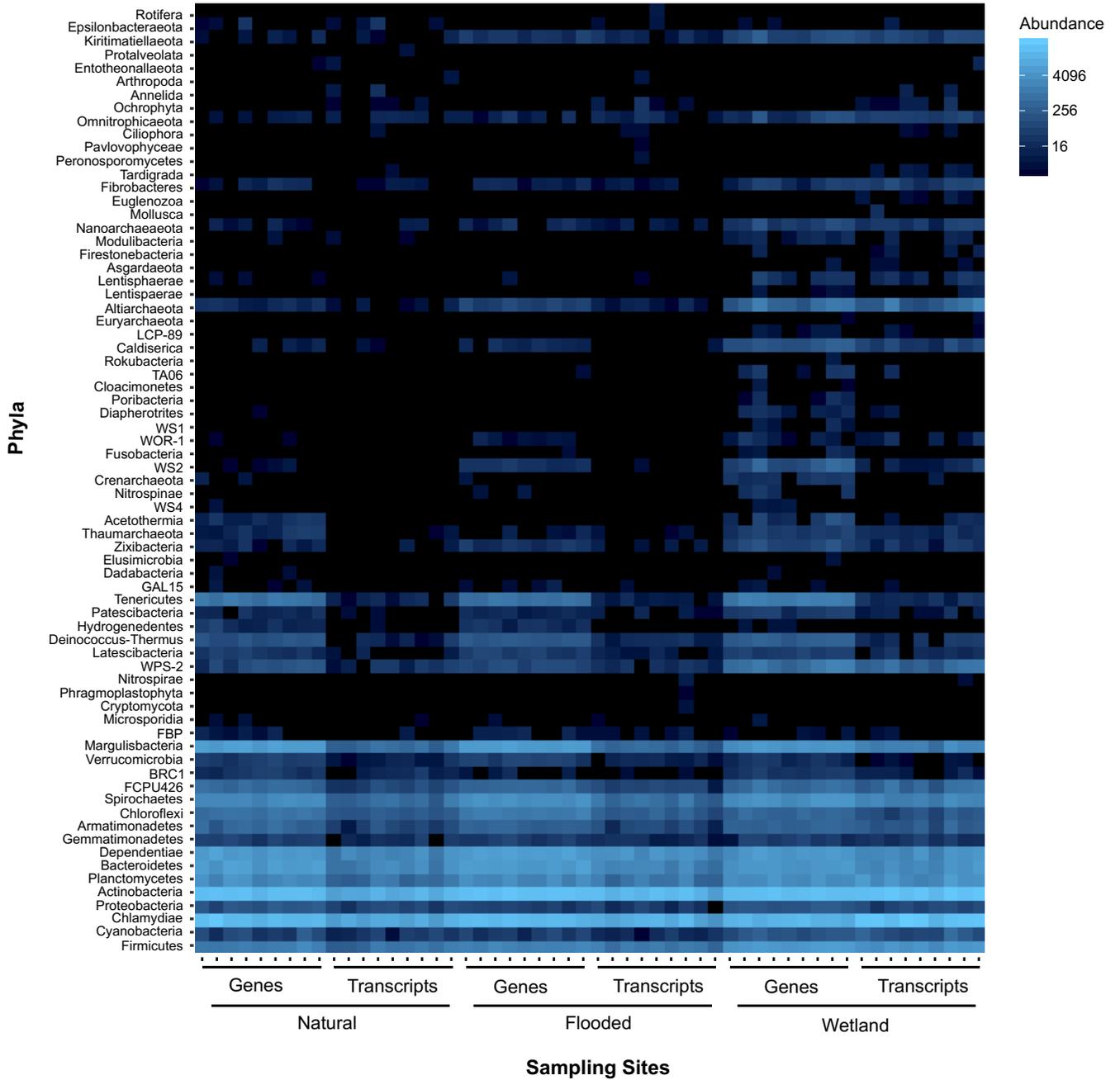
The species abundance and distribution were used to measure the microbial species diversity of periphyton biofilm communities at each sampling site using the Simpson index (1-dominance) (Magurran, 2005; Oksanen et al., 2013). All replicates from both 16S ribosomal genes and transcripts were plotted and designated by sampling site (Figure 3). Majority of samples range between 0.95 - 1.0, meaning that the microbial communities within the samples are very diverse. Samples collected from the flooded site are tightly clustered near 1.0, while the samples from both the natural and wetland sampling sites range from 0.85 - 1.0 and 0.80 - 1.0 respectively (Figure 3).



**Figure 3.** Alpha diversity analysis for all 16S rRNA amplicons sequenced at each sampling sites using the Simpson Index. rDNA is represented by red circles and rRNA by blue.

### **3.2.2 Diversity, abundance, and variability of microorganisms**

Microbial diversity and abundance of all 16S ribosomal genes and transcripts were visualized at the phylum level using a heatmap (Figure 4). A clear pattern can be observed between 16S rDNA and 16S rRNA. In all three sampling sites, 16S rDNA has a greater abundance and diversity of phyla across every replicate than 16S rRNA. The greatest diversity of phyla was observed in both the rDNA and rRNA amplicons of the wetland site compared to both the natural and flooded sites (Figure 4).

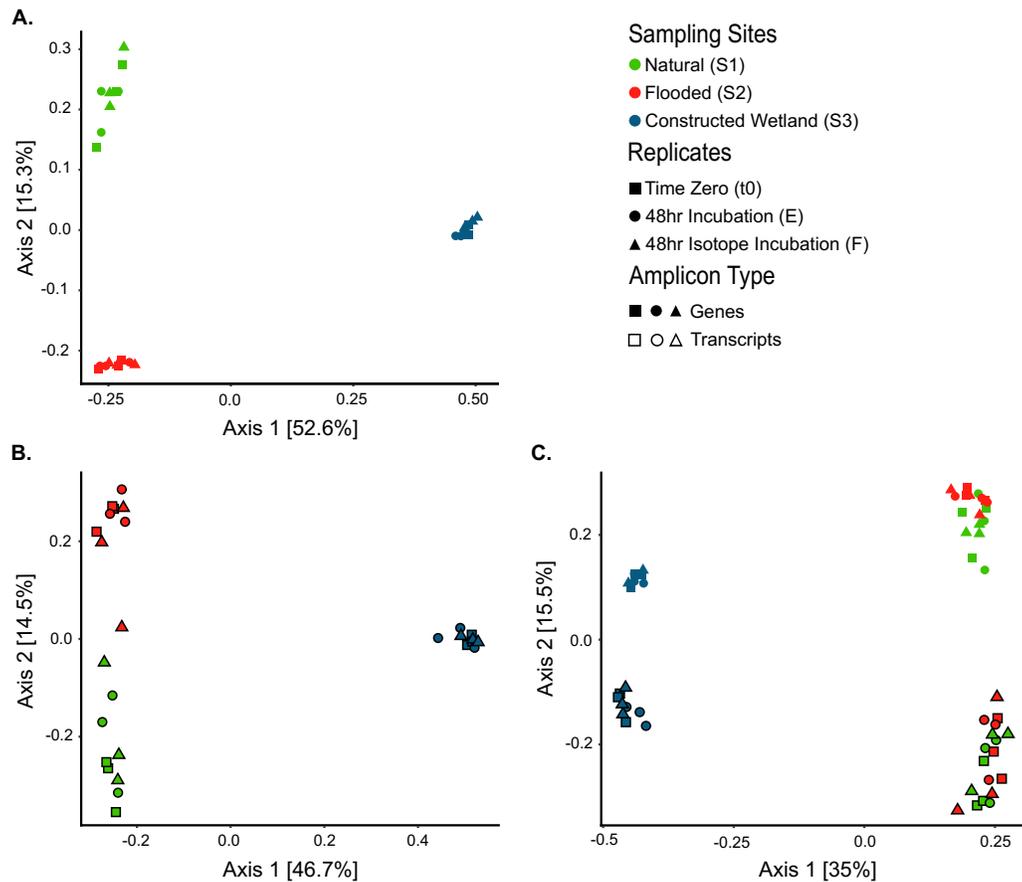


**Figure 4.** Heatmap illustrating the abundance of ASVs of all phyla using the 16S rDNA and rRNA amplicons from all three sampling sites. Higher abundances were observed in 16S rDNA amplicons rather than rRNA, with the greatest diversity of abundant phyla of both rDNA and rRNA observed in the artificial wetland sampling site.

To further explore the diversity and variability of microbial species present, a principal coordinate analysis (PCoA) using the Bray-Curtis dissimilarity matrix was used. The Bray Curtis Principal Coordinate Analysis (PCoA) plot was constructed for both 16S ribosomal gene and transcript data to compare the ASVs at each sampling site and treatment. In both plots the samples group together by site, but not by treatment. The 16S rDNA data shows that the samples in the flooded and wetland site are clustered tightly together within each site while the samples for the natural site are less tightly clustered and more separated along axis 2 (Figure 5a). The natural and flooded sites are at the same position on axis 1, while the wetland site is further away (Figure 5a). However, the sampling sites are evenly spread across axis 2. An Adonis test was run to test the significance of the results from the PCoA plot for the rDNA with 999 permutations. The p-value comparing sites was statistically significant ( $p < 0.01$ ), while neither treatments nor the interaction between treatments and sites were significant (all  $p > 0.05$ ), supporting the cluster patterns of samples grouping together by site and not treatment.

Similar patterns were observed in the 16S rRNA transcript PCoA plot, however the samples were not as tightly clustered together (Figure 5b). Of the three sampling sites, the samples from the wetland site still clustered relatively close, but those from the flooded site do not and are more spread out across axis 2. As seen in the previous plot, the natural and flooded sites are at the same position on axis 1, while the wetland site is further away (Figure 5b). The p-values (999 permutations) for both treatment and treatment and site were not significant. However, the p-value comparing sites was statistically significant. As the treatments are not significant, we will refer to them as replicates for the remainder of the study.

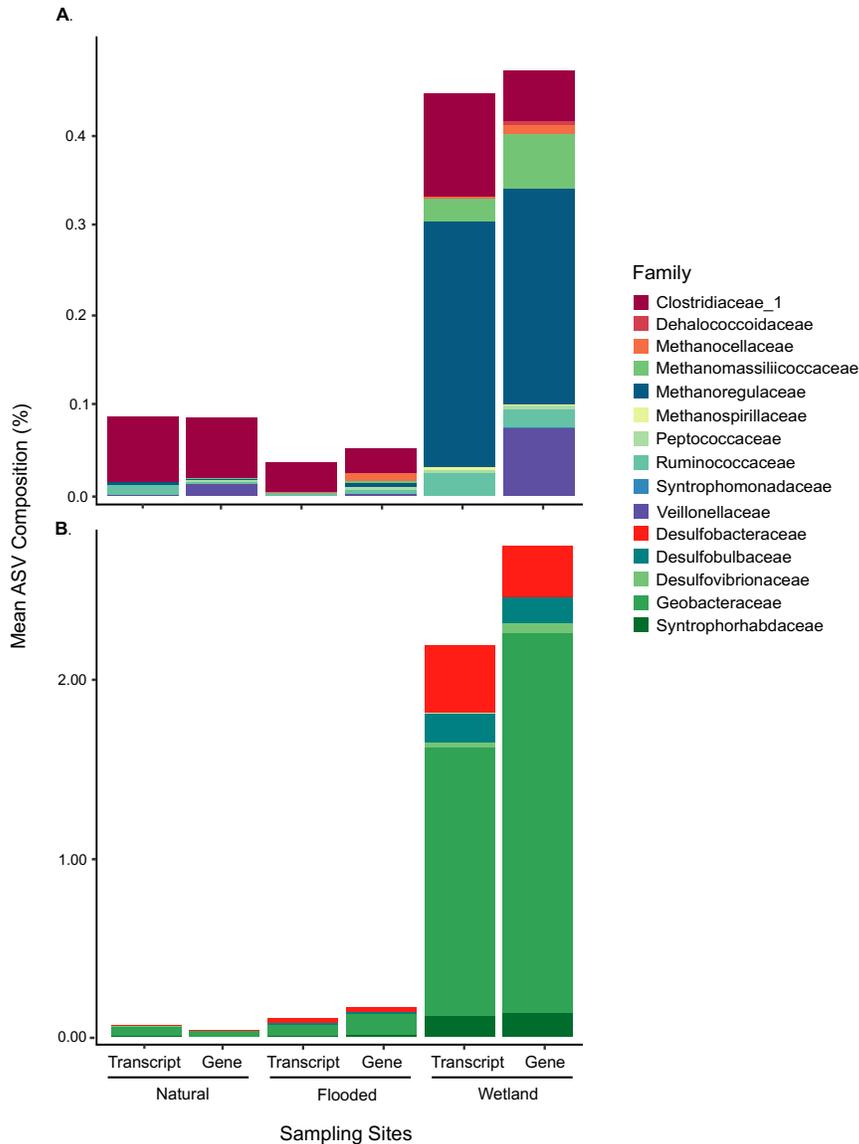
A final PCoA plot was created comparing both 16S ribosomal genes and transcripts on the same plot. Both amplicon types followed similar patterns when plotted individually, with the samples grouping by sampling site, however they are now separated by amplicon type.



**Figure 5A.** Principal coordinate analysis plot using the Bray-Curtis dissimilarity measure of the 16S rDNA. Samples grouped by sampling site are statistically significant from one another but are not statistically significant based on treatments (i.e., replicates). Sampling sites are represented by colours and treatments or replicates by solid shapes. The natural, flooded and wetland sites are green, red and blue respectively, with the replicates; time zero (t0), 48 hr incubation (E), and 48 hr isotope incubation (F) represented by squares, circles and triangles respectively for all plots. **B. 16S rRNA.** Samples grouped by sampling site are statistically significant from one another but are not statistically significant based on treatments (i.e., replicates). Sampling sites are represented by colours and treatments by outlined shapes. **C. Both amplicon types.** Samples grouped by sampling site and amplicon type. Sampling sites and treatments from 16S rRNA are represented by colours and outlined shapes.

### 3.2.3 Taxonomic analysis of 16S rRNA amplicons

Taxonomic analysis of the 16S ribosomal gene and transcript amplicons revealed that *Proteobacteria* and *Cyanobacteria* were the most abundant phyla across all sampling sites averaging 39% and 30% respectively (Figure S2). A more specific taxonomic analysis focused on mercury methylating taxa such as methanogens, sulfate-reducing bacteria and iron-reducing bacteria known to contain the *hgcA* gene. Bar plots containing families of phyla previously implicated in mercury methylating microorganisms were created to observe the abundance and diversity of the periphyton biofilm communities (Figure 6a, 6b). Families of *Delta-Proteobacteria* previously associated with mercury methylation are the most abundant of these groups, accounting for an average of 0.98% rDNA and 0.79% rRNA across all sampling sites (Figure 6b). However, the greatest abundances of *Delta-Proteobacteria* are observed in the wetland site accounting for an average of 2.75% of rDNA and 2.19% of rRNA of the total 16S rRNA ASVs. The *Delta-Proteobacteria* are represented primarily by *Geobacteraceae*, 2.13% of rDNA and 1.50% of rRNA of a total of 16S ASVs in the wetland site, however *Desulfobacteraceae*, *Desulfobulbaceae*, *Desulfovibrionaceae*, and *Syntrophorhabdaceae* are also present (Figure 6b). Other groups present containing potential mercury methylating microorganisms are families within the phyla *Euryarchaeota*, *Firmicutes*, and *Chloroflexi*. *Euryarchaeota*, accounting for an average of 0.11% rDNA and 0.10% rRNA of total 16S rRNA ASVs across all sampling sites, with the greatest abundance seen in the wetland site at 0.31% rDNA and 0.30% rRNA (Figure 6a). The *Euryarchaeota* are primarily represented by *Methanoregulaceae* accounting for 0.23% rDNA and 0.27% rRNA of the total 16S rRNA ASVs in the wetland sampling site with other families including; *Methanocellaceae*, *Methanomassiliicoccaceae*, and *Methanospirillaceae* (Figure 6a). *Firmicutes* account for an average of 0.09% of rDNA and 0.89% transcripts across all sampling sites with the highest abundance at the wetland site accounting for 0.16% rDNA and 0.14% rRNA and are represented by; *Clostridiaceae\_1*, *Peptococcaceae*, *Ruminococcaceae*, *Syntrophomonadaceae*, and *Veillonellaceae* (Figure 6a). *Chloroflexi* is represented by one family, *Dehalococcoidaceae*, accounting for 0.0011% of total 16S rDNA ASVs and 0.0005% of total 16S rRNA ASVs.

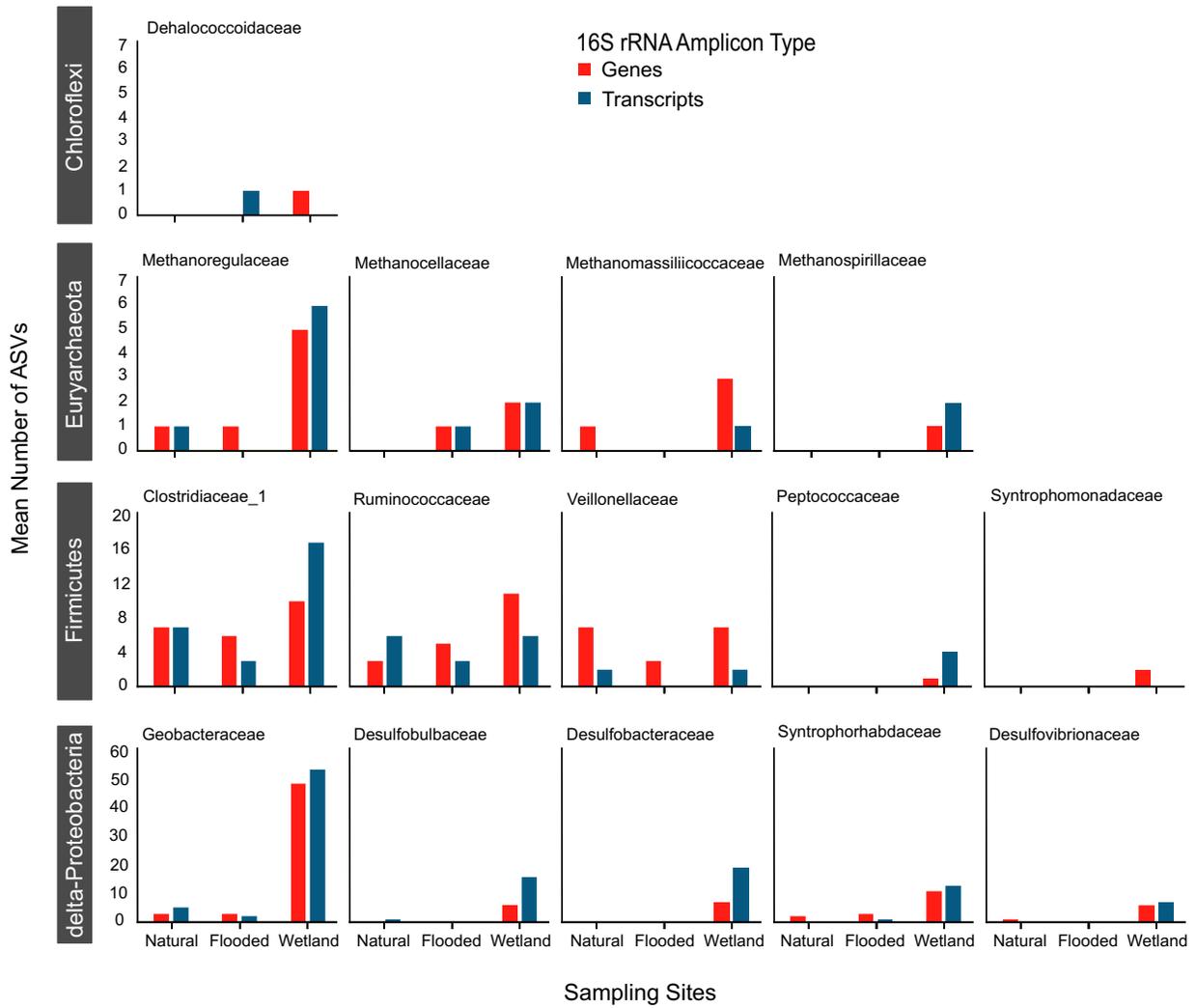


**Figure 6A.** Mean percent read counts from 16S rRNA amplicons of families associated with mercury methylation based on the main groups of mercury methylators (i.e. Methanogens, Firmicutes and Chloroflexi). Bars are broken down into rDNA and rRNA for all three sampling sites. *Methanoregulaceae* are the most abundant family seen across all sampling sites with their greatest abundance observed in the wetland site. **6B.** Mean percent read counts from 16S rRNA amplicons of families associated with mercury methylation based on the main classes of mercury methylators (i.e. *Delta-Proteobacteria*). Bars are broken down into rDNA and rRNA for all three sampling sites. *Geobacteraceae* are the most abundant family seen across all sampling sites with their greatest abundance observed in the wetland site.

### 3.2.4 Indicator species analysis

An indicator species analysis was conducted on both 16S ribosomal gene and transcript amplicons to determine if and how many indicator species were present at each sampling site. The number of indicator species for the primary phyla containing mercury methylating microorganisms (i.e. *Delta-Proteobacteria*, *Firmicutes*, *Euryarchaeota*, and *Chloroflexi*) were plotted at the taxonomic rank of family (Figure 7). On average across all families presented, the greatest abundance of indicator species was seen in the wetland site, with a relatively similar number of indicator species found at the natural and flooded sampling sites in all families (Figure 7). The class *Delta-Proteobacteria* represented by the families; *Geobacteraceae*, *Desulfobulbaceae*, *Desulfobacteraceae*, *Syntrophorhabdaceae*, and *Desulfovibrionaceae*, are the most abundant of all indicator species, all primarily located in the artificial wetland sampling site (Figure 7). *Geobacteraceae* specifically are the most dominant family of indicators with 49 16S rDNA ASVs and 54 16S rRNA identified as indicator species, all found within the wetland site. There are only 3 indicator species from the 16S rDNA amplicons identified as *Geobacteraceae* in both the natural and flooded sites and 5 ASVs in the natural and 2 in the flooded sampling sites from the 16S rRNA transcript amplicons (Figure 7). The remaining families within the class *Delta-Proteobacteria* contain between 6-19 ASVs classified as indicator species for both 16S rDNA and rRNA amplicons in the wetland site, and up to 3 ASVs in both the natural and flooded sites (Figure 7).

The next most abundant group of indicator species are in the phylum *Firmicutes*, represented by the families; *Clostridiaceae\_1*, *Ruminococcaceae*, *Veillonellaceae*, *Peptococcaceae*, and *Syntrophomonadaceae* (Figure 7). The majority of ASVs assigned to the families within *Firmicutes* are present in the wetland sites, however 16S rDNA and rRNA amplicons classified as *Clostridiaceae\_1*, *Ruminococcaceae*, and *Veillonellaceae* are also present in the natural and flooded sites ranging between 2 to 17 ASVs (Figure 7). However, there are no ASVs classified as indicator species from the families *Peptococcaceae* and *Syntrophomonadaceae* in the natural and flooded sites, they are only present at low abundances in the wetland site (Figure 7).



**Figure 7.** Indicator species analysis broken down into important families of mercury methylating microorganisms found in known mercury methylating groups e.g., *Delta-Proteobacteria*, *Firmicutes*, *Methanogens*, and *Chloroflexi*. Each plot contains the data from both 16S rDNA (red) and rRNA (blue) for each sampling site.

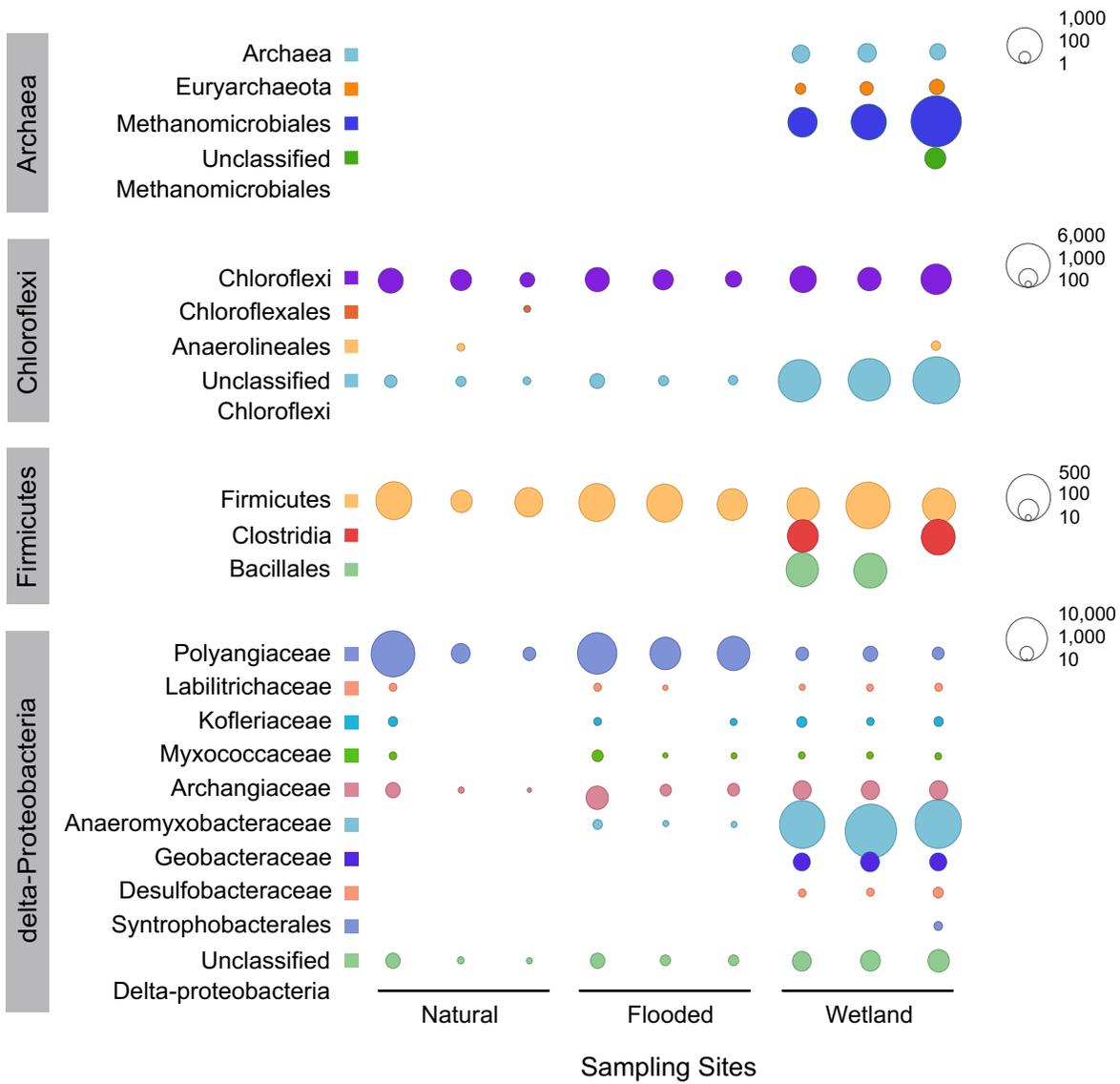
ASVs classified in the phyla *Euryarchaeota* and *Chloroflexi* known to methylate mercury have also been identified as indicator species, however at much lower abundances with a maximum of 6 ASVs. *Methanoregulaceae* are the most abundant indicator species representing *Euryarchaeota*. They are present at all sites with 1 16S rDNA and rRNA ASV found in the natural site, 1 ASV in the flooded site, and 5 16S rDNA and 6 16S rRNA ASVs at the wetland site. The remaining families; *Methanocellaceae*, *Methanomassiliicoccaceae*, and *Methanospirillaceae*, contain a maximum of 3 ASVs for both 16S rDNA and rRNA from the natural and wetland sites, with majority of ASVs located in the wetland site. Only 1 family, *Dehalococcoidaceae*, from the phylum *Chloroflexi*, was identified as an indicator species. 2 16S rRNA ASVs were classified as *Dehalococcoidaceae*, 1 rRNA in the flooded site and 1 rDNA in the wetland site.

### **3.3 Metagenomic analysis**

Following metagenome sequencing of each periphyton biofilm sample, an average of 27.2 million reads in each forward and reverse fasta files were returned. Of the 27.2 million reads produced, approximately 1.3 million reads were assembled from each periphyton biofilm sample using MEGAHIT (Table S5). Indicating that approximately 4.78% of the metagenomes were assembled with the highest proportion of assembly occurring in the natural sampling site at 7.13%, while only 3.81% and 3.4% of the reads in the metagenomes for the flooded and wetland site were assembled. Of the assembled metagenomes, the proportion of prokaryotes and eukaryotes in the proteome are 90.82% and 8.95% respectively, with the lowest percentage of prokaryotes found in the natural sampling site only comprising 80% of the assembled metagenome. The proportion of prokaryotes and eukaryotes in the DNA sequences of the assembled metagenome are 86.5% and 13.2% respectively. The lowest proportion of prokaryotes are seen in the samples collected from the natural sampling site only comprising 68.7% of the annotated proteome.

### 3.3.1 Taxonomic analysis of potential mercury methylating taxa

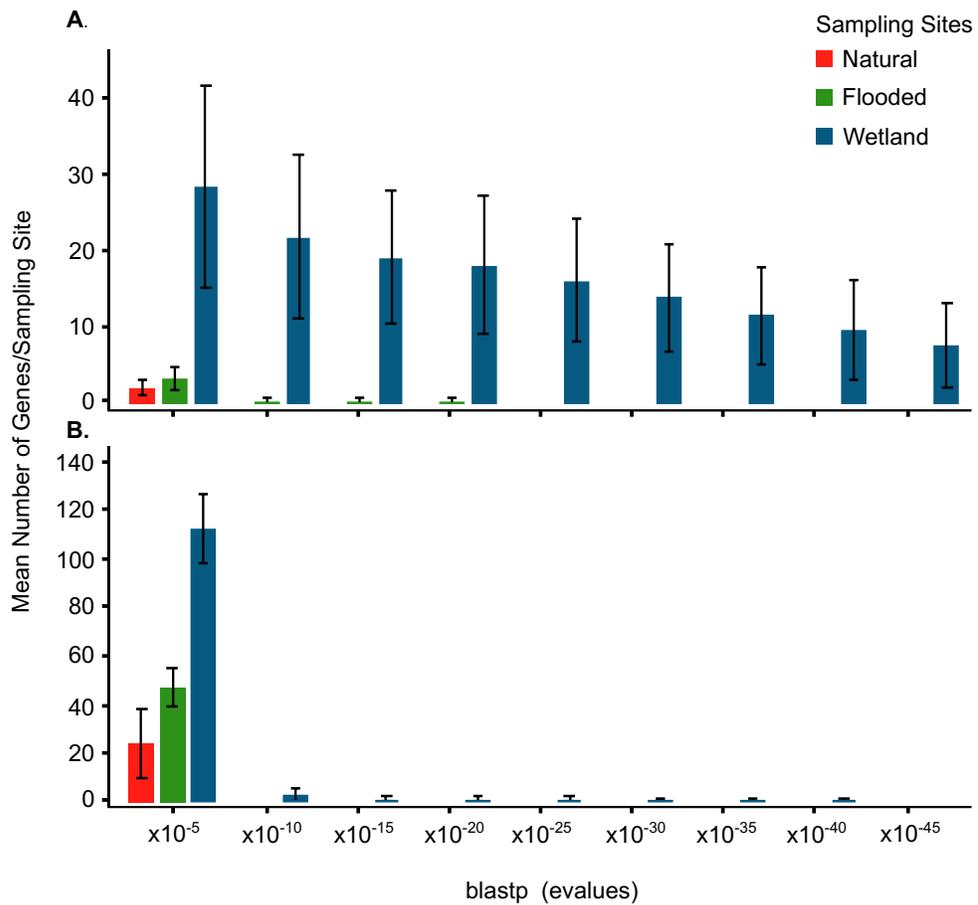
Taxonomic analysis of the proteins of the periphyton biofilm metagenomes using MEGAN6 (Huson et al., 2016) revealed similar patterns seen in the taxonomic analyses of the 16S rRNA amplicons. However, the taxa present are not as diverse as previously observed. Absolute read counts of the assigned taxa of groups previously associated with the methylation of mercury; *Delta-Proteobacteria*, *Chloroflexi*, *Firmicutes*, and *Archaea*, were examined at various taxonomic levels (Figure 8). The most abundant group found within the metagenomes are *Delta-Proteobacteria*, represented by microorganisms at various taxonomic levels such as *Geobacteraceae*, *Desulfobacteraceae*, and *Syntrophobacterales*, all of which are only present in the wetland site with approximately 1,000 amino acid sequences (Figure 8). The class *Methanomicrobiales* from the phylum *Euryarchaeota* are also only present in the wetland site, also with approximately 1,000 amino acid sequences in each replicate. Microbes present in the phyla *Firmicutes* and *Chloroflexi* are not as abundant, containing approximately 100 to 200 proteins, and are present in all sampling sites (Figure 8).



**Figure 8.** Plot illustrating the number of assigned taxa at each taxonomic rank for each replicate at every sampling site using the metagenomic data.

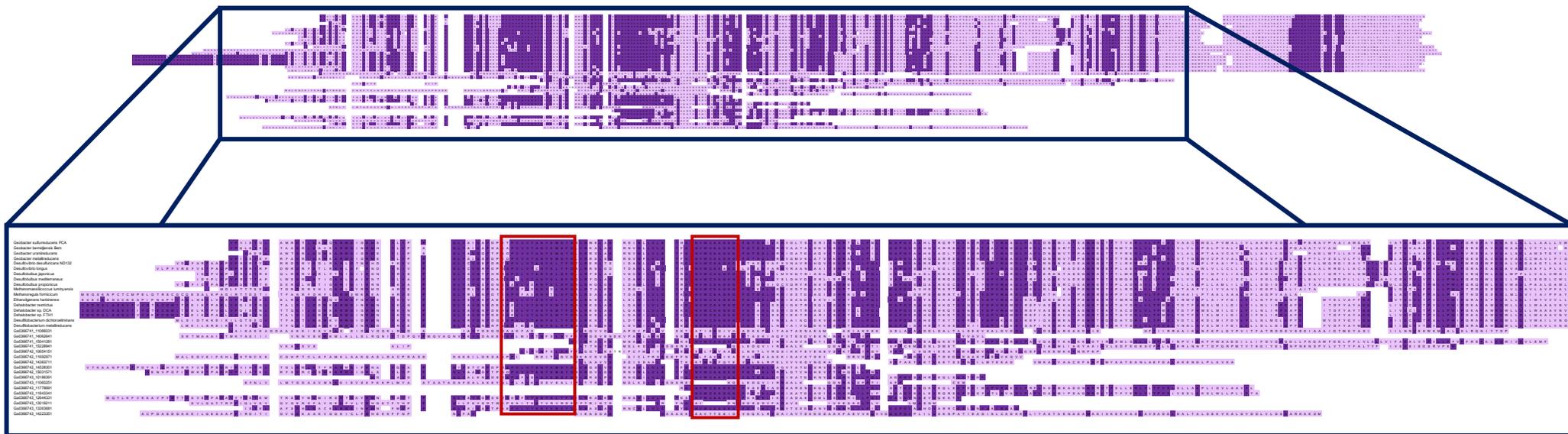
### 3.3.2 Potential *hgcA* and *hgcB* genes in the assembled metagenome

To identify potential *hgcA* and *hgcB* genes from the metagenomic data, BLAST (blastp) searches were conducted using known *hgcA* and *hgcB* protein sequences against the assembled metagenome at various levels of stringency by adjusting the Expect value cut-off (e-value). When searching for the *hgcA* gene, the greatest abundance of potential genes was seen in the artificial wetland site (Figure 9a). As the e-values decreased, the number of potential *hgcA* genes also decreased. At an e-value of  $\times 10^{-5}$ , there are an average of 2 potential *hgcA* genes in the natural site, 3.33 in the flooded site and 28.67 in the wetland site. When the e-value is decreased to  $\times 10^{-10}$ , there are no potential *hgcA* genes in the natural samples, however there are an average 0.33 in the flooded and 22 in the wetland samples (Figure 9a). The number of potential *hgcA* genes further decrease as the e-value is lowered with an average 7.67 potential *hgcA* genes at  $\times 10^{-45}$  in the wetland site (Figure 9a). Similar patterns were observed when searching for *hgcB* genes within the assembled metagenome dataset, however there is a larger decrease of potential *hgcB* genes as the e-value decreases (Figure 9b).

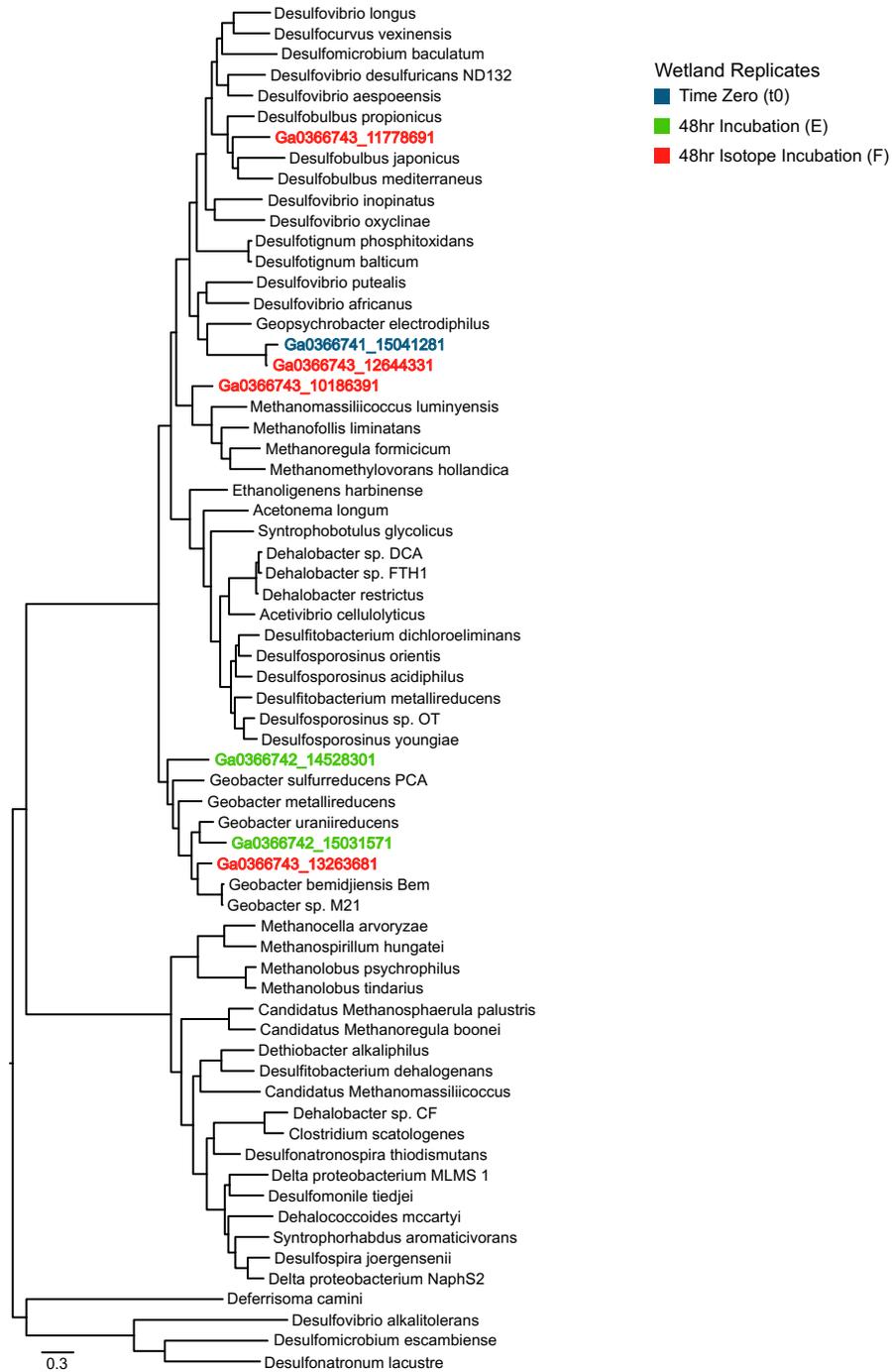


**Figure 9.** The mean and standard error of potential hgcA (A) and hgcB (B) genes per sampling site found in the assembled metagenomic data at various stringency levels. The three sampling sites; natural, flooded and wetland are distinguished by the colours; red, green and blue respectively.

To determine the likelihood of the potential *hgcA* genes being true *hgcA* genes, we compared them to the conserved regions of known *hgcA* protein sequences. To visualize the alignment of potential *hgcA* genes isolated from the assembled metagenomic data (e-value  $\times 10^{-40}$ ), the 17 potential *hgcA* protein sequences of the artificial wetland site from the blastp outputs were aligned to 17 reference *hgcA* protein sequences (Figure 10). Of the 17 sequences aligned, 10 sequences did not align to the reference *hgcA* protein sequences in either of the two conserved regions identified and were removed (Ga0366741\_11086031, Ga0366741\_14062641, Ga0366741\_15228941, Ga0366742\_10654151, Ga0366742\_11692971, Ga0366742\_14383711, Ga0366743\_11060251, Ga0366743\_13019211, Ga0366743\_11843341, and Ga0366743\_14223351) (Figure 10). The remaining 7 sequences from the artificial wetland samples that aligned to the conserved regions of the reference *hgcA* genes were aligned to all reference *hgcA* genes, and used to construct a maximum likelihood phylogenetic tree (Figure 11). 3 potential *hgcA* genes (Ga0366743\_11778691, Ga0366741\_15041281, and Ga0366743\_12644331) grouped with the sulfate-reducing *Delta-Proteobacteria*, while 3 (Ga0366742\_14528301, Ga0366742\_15031571, and Ga0366743\_13263681) were grouped with the iron-reducing *Geobacteraceae* of the *Delta-Proteobacteria* class, and only 1 potential *hgcA* gene identified from the blastp search clustered with the methanogens (Ga0366743\_10186391) (Figure 11).



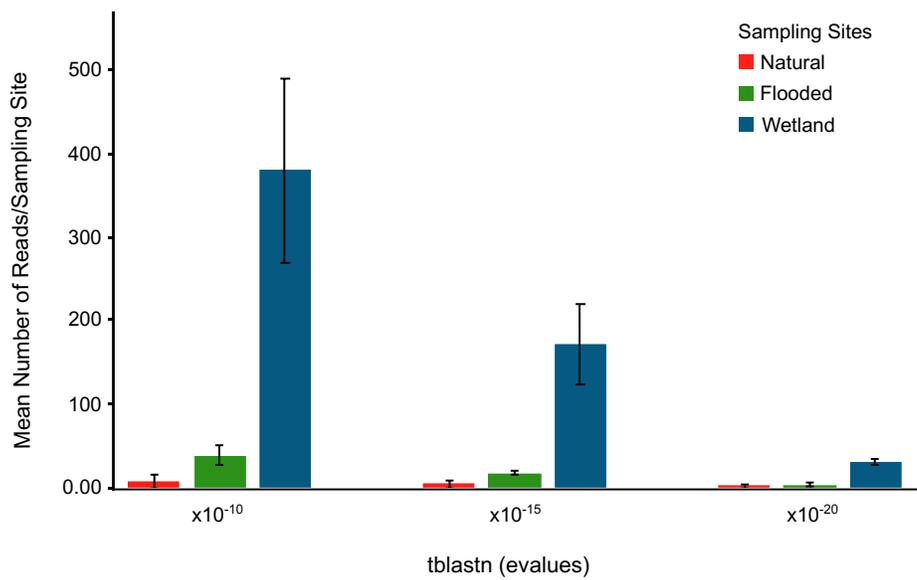
**Figure 10.** Sequence alignments comparing a subset of known *hgca* genes from *Delta-Proteobacteria*, *Firmicutes*, and *Euryarchaeota* to blastp outputs from the assembled metagenomic data. The red boxes indicate highly conserved regions in the putative cobalamin-binding domain of the *hgca* gene. Abbreviations for amino acids; A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.



**Figure 11.** Phylogenetic analysis comparing known *hgcA* sequences to potential *hgcA* genes from the replicates (e-value  $1 \times 10^{-40}$ ) of assembled protein-coding genes from the metagenomes at the artificial wetland sampling site. The replicates are time zero (t0), 48 hr incubation (E), and 48 hr isotope incubation (F) represented by the colours blue, green, and red respectively.

### 3.3.3 Potential *hgcA* genes in the unassembled metagenome

A tblastn search using known *hgcA* protein sequences against the unassembled metagenome reads were used to identify any potential *hgcA* genes that were not assembled into the assembled metagenomes. The tblastn searches were conducted using various levels of stringency by adjusting the e-values ( $1 \times 10^{-10}$ ,  $1 \times 10^{-15}$ , and  $1 \times 10^{-20}$ ) to identify reads with the greatest potential of being true *hgcA* genes. At all levels of stringency, the wetland site had a significantly greater number of potential *hgcA* genes than observed in the natural and flooded sampling sites (Figure 12). As the e-values decreased, so did the number of potential *hgcA* genes at all sampling sites. No potential *hgcA* genes were observed at any sampling site with an e-value greater or equal to  $1 \times 10^{-25}$ .



**Figure 12.** The mean and standard error of potential hgcA genes per sampling site found in the unassembled portion of the metagenomic data at various stringency levels. The three sampling sites; natural, flooded and wetland are distinguished by the colours; red, green and blue respectively.

To visualize the alignment of potential *hgcA* genes isolated from the unassembled metagenomic reads (e-value  $\times 10^{-20}$ ), the 90 potential *hgcA* protein sequences of the artificial wetland site from the tblastn outputs were aligned to 17 reference *hgcA* protein sequences (Figure 13). Of the 90 reads aligned to the reference *hgcA* sequences, only 3 reads aligned to the first conserved region. However, 10 reads aligned to the reference *hgcA* sequences at the second identified conserved region. Only one read aligned to the reference *hgcA* sequences at both conserved regions (Figure 13).

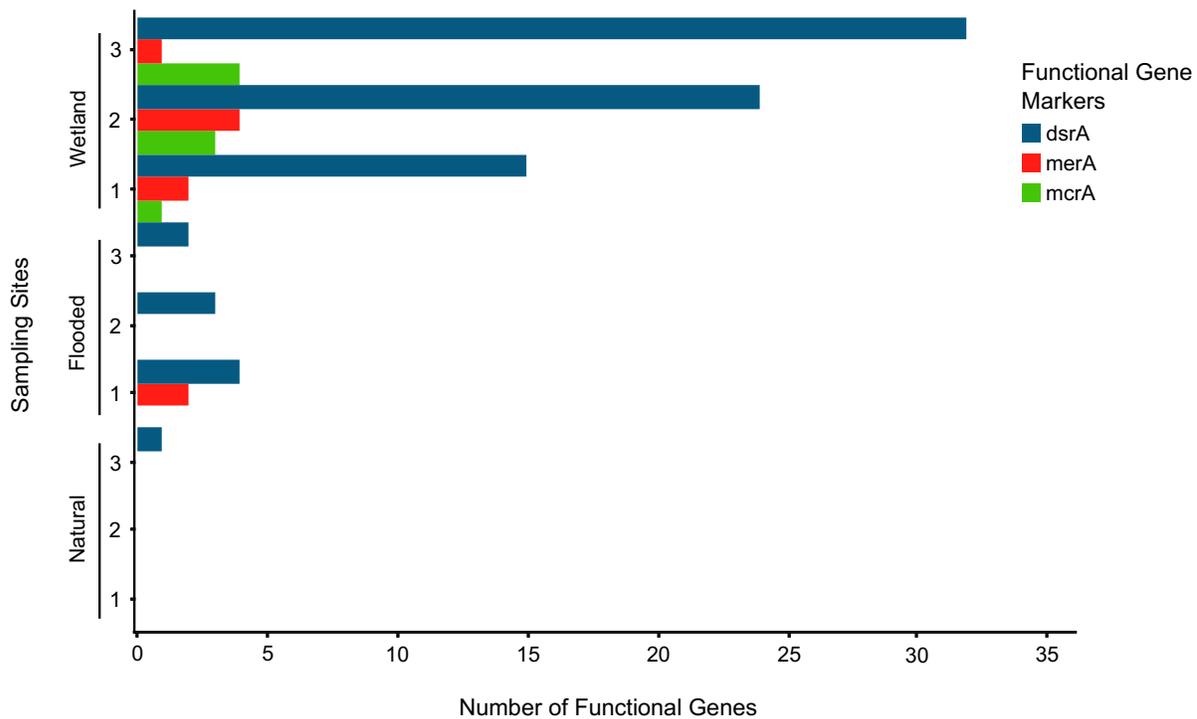


**Figure 13.** Sequence alignments comparing a subset of known *hgcA* genes from *Delta-Proteobacteria*, *Firmicutes*, and *Euryarchaeota* to tblastn outputs from the unassembled metagenomic data. The red boxes indicate highly conserved regions in the putative cobalamin-binding domain of the *hgcA* gene.

Abbreviations for amino acids; A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

### 3.3.4. Functional gene markers associated with mercury methylation

The number of functional gene markers related to mercury methylating microorganisms was calculated at each sampling site from the metagenomic data produced by the JGI IMG/M. The most abundant gene at all sampling sites was *dsrA*, sulfite reductase alpha subunit, with the greatest proportion seen in the wetland site ranging between 15 and 34 genes (Figure 14). *mcrA*, 5-methylcytosine-specific restriction enzyme A, was only present at the wetland with less than 5 genes per replicate, while 2 *merA* genes, mercuric reductase genes, were found in the second replicate at the flooded site and between 1 to 4 genes in the wetland site (Figure 14). No *merB* genes were found in any sample.



**Figure 14.** Bar plot indicating the number of functional gene markers found at each sampling site. dsrA (sulfite reductase alpha subunit), merA (mercuric reductase), and mcrA (methyl-coenzyme M reductase alpha subunit) are represented by the colours blue, red, and green respectively.

## 4. Discussion

Methylmercury (MeHg) is a potent neurotoxin produced by mercury methylating microorganisms. Its neurotoxic effects and ability to bioaccumulate and biomagnify in the food web, specifically within top predator fish populations, are creating public health concerns for the residents of the Wemotaci Reserve. In order to determine the sources of increased MeHg levels, whether from the construction of the Chute-Allard hydroelectric generating station, the wildfires, or local logging practices, all occurring around the same time, we analyzed stable mercury isotopes, for the analysis of Hg methylation and demethylation rates, 16S rRNA amplicons, the *hgcAB* gene pair, and metagenomes extracted from periphyton biofilms and 1) compared the microbial community composition within and between periphyton biofilms, 2) assessed the natural diversity of mercury methylating microorganisms within periphyton biofilms, and 3) quantified the abundance of mercury methylating microorganisms at each sampling site in relation to mercury methylation and demethylation rates. Through the analysis of the periphyton biofilms, we observed the highest mercury methylation rates, as well as a clear pattern of taxa previously associated with mercury methylation, and increased levels of potential *hgcA* and *hgcB* genes were found in samples collected from the artificial wetlands.

### 4.1 Diversity and abundance of microbial communities from 16S rRNA amplicons

16S rRNA gene and transcript amplicon analyses revealed significant variation of microbial communities between sampling sites, but not between the treatments. These findings confirm that the stable mercury isotope addition did not significantly affect the community composition or diversity of microorganisms of samples collected at each site. The samples collected from the natural site in which periphyton biofilms were retrieved from a submerged tree have a greater dissimilarity of community composition within the samples. However, the community composition of periphyton biofilms from the flooded and wetland sites, in which samples were collected from artificial substrates colonized for one year, are less dissimilar from each other, indicating that natural versus artificial substrates may have an effect on the microbial community composition of periphyton biofilms (Khadra et al., 2018; Roulet et al., 2001; Tarkowska-Kukuryk & Mieczan, 2012). One major factor in the microbial community

composition between the periphyton biofilms collected from natural versus artificial substrates could be associated to the age of the periphyton matrix. Newly colonized biofilms rely on the availability of nutrients found in the water column, however, older biofilm communities begin to form more recycling ecological niches and no longer require from the surrounding water and are therefore more independent from external inputs (Hagerthey et al., 2011; Tlili et al., 2011). Another factor possibly impacting the microbial community composition between the natural versus artificial substrates is the type of artificial substrate in which the periphyton biofilms were grown on. Artificial substrates have the ability to promote the growth of certain microbes over others (i.e., glass substrates promote the growth of diatoms). Biofilms collected from non-porous artificial substrates also lack the ability to have biological or chemical exchanges between the microbes and the substrate it is attached to (Personal communication, Dolores Planas).

Although the periphyton biofilms collected from artificial substrates were naturally grown in the river, macrophyte-associated periphyton has also been shown to produce higher levels of MeHg than cultivated periphyton biofilms as Hg methylation is enhanced by the presence of living or decomposing macrophyte mats and roots, and organic matter (Mauro et al., 2002; Roulet et al., 2001). However, results from the methylation experiments indicate that there are higher rates of mercury methylation occurring in the artificial wetlands, from which periphyton biofilms were collected from artificial substrates, rather than the periphyton biofilms retrieved from natural substrates in the natural sampling site, as well as the biofilms grown on artificial substrates in the flooded site. This indicates that the environment may have a greater impact on the structure and microbial community composition of the periphyton biofilms than the type of substrate they were collected from.

Environmental influence on the microbial community composition is supported by Bray-Curtis Principal Coordinate Analyses (PCoA) investigating dissimilarity of the 16S rRNA amplicon sequences from the three sampling sites. The positions of the sample clusters from each sampling site for both the gene and transcript PCoA plots are similar to each other. There is a greater dissimilarity seen on axis 1, where the natural and flooded sites are approximately in the same position, while the wetland site cluster is completely separate. These results are consistent with results seen from the methylation experiment, where we observed both the

natural and flooded sampling sites having low rates of methylation and are very similar to each other while the wetland site has a higher level of methylation, significantly different than from that of the natural and flooded site. The positions of the sample clusters for each sampling site seen on axis 2 follow the same pattern observed from the results of the demethylation rate calculations. The flooded and wetland sampling sites are more closely related, but not significantly different from the natural sampling site.

Further analysis of the 16S rRNA amplicon data revealed greater diversity and abundance of taxa and indicator species in the artificial wetlands, including taxa previously implicated in mercury methylation. Previously, literature has focused on sulfate-reducing bacteria in sediments as being the main source of methylmercury in the environment, but more recently, studies have implicated methanogens in periphyton communities as the primary sources for mercury methylation (Bravo et al., 2018a; Hamelin et al., 2015). However, our results reveal the most abundant microorganisms with potential for mercury methylation found in all sampling sites, with the greatest abundance in the artificial wetlands, as iron-reducing *Delta-Proteobacteria* from the family *Geobacteraceae*. Methanogens are only present in the artificial wetlands and are one of the least abundant potential methylators, contradicting previous findings of methanogens being the principal methylators in periphyton samples (Bravo et al., 2018a; Hamelin et al., 2011). Iron-reducing bacteria are usually considered secondary methylators only after sulfate-reducing bacteria, and are generally investigated in studies focusing on microbial mercury methylation occurring in sediments rather than in periphyton biofilms (Bravo et al., 2018a; Bravo et al., 2018b). However, a study has linked algal biomass, specifically the input of components found in algal organic matter, to sediments as being a potential source of increased MeHg levels in lake sediments (Lei et al., 2019). These findings could imply that iron-reducing bacteria responsible for mercury methylation in lake sediments may have originated from input by periphyton biofilms.

#### **4.2 Diversity and abundance of microbial communities from metagenomic data**

Similar patterns as seen in the 16S rRNA amplicon analyses were observed in the metagenomic data. Once again, there is a greater abundance of taxa previously associated with

mercury methylation in samples collected from the artificial wetlands. However, the families within the main methylating groups being studied are less diverse and not as representative of the known methylating species seen in the 16S rRNA amplicon data. This may be due to the small percentage of the metagenomic data that is assembled. A large proportion of the unassembled metagenome may be eukaryotic DNA from the periphyton biofilms. For this reason, analyses conducted using the metagenomic data were more targeted towards identifying mercury methylating microbes using the *hgcAB* gene pair and functional gene markers that were not possible with the 16S rRNA data analysis.

#### **4.3 Potential *hgcA* and *hgcB* genes in the artificial wetlands**

The *hgcAB* gene pair was first discovered in 2013 when it was determined that the genes were present in methylators, but absent in nonmethylators (Parks et al., 2013). Following the identification of the *hgcAB* gene pair, researchers began compiling lists of organisms containing these genes and used them to identify potential methylators in the environment (Gilmour et al., 2013; Parks et al., 2013; Podar et al., 2015). From these data, numerous studies began designing primer sets specifically targeting both the *hgcA* and *hgcB* genes from environmental samples and have been able to identify methylators in sediments, rice paddy soils and in wetlands (Bae et al., 2014; Christensen et al., 2016; Liu et al., 2014; Schaefer et al., 2014).

We were unable to identify the *hgcAB* gene pair from the periphyton samples using PCR amplification with both broad-range and clade-specific primers. Traditionally, the *hgcAB* gene pair has been difficult to detect (Christensen et al., 2016; Podar et al., 2015). This could be due to multiple causes. The first cause being that the function of the *hgcAB* gene pair is unknown. Without knowing the function, we are limited in the methods that we can use to locate these genes. Another is that there is a patchy distribution of the *hgcAB* gene throughout the tree of life (Podar et al., 2015). Of the taxa that have the potential to methylate mercury, only a few may contain the *hgcAB* gene pair (Podar et al., 2015). For instance, approximately 20% of sequenced genomes from two methanogenic orders that are known to contain the *hgcAB* gene pair actually have the genes, and only about 10% of the species from the orders of the class *Delta-Proteobacteria* associated with mercury methylating families contain the *hgcAB* gene pair

(Gilmour et al., 2018). A third limitation of working with the *hgcAB* gene pair is that when they are found in environmental samples, they are found at extremely low copy numbers. A metagenomic study found that of the 3,500 metagenomes studied, only about one third of the samples contained the *hgcA* gene, of those samples, there was only approximately 22.3 genes found per 1 billion base pairs (Podar et al., 2015). Another reason why we may not have been able to amplify the *hgcA* genes from periphyton biofilms could be that the primers used may not target the correct microorganisms. Recent studies have identified other methanogens containing the *hgcAB* gene pair with the ability to methylate mercury that were not originally discussed in the papers from which our primers were designed (Gilmour et al., 2013, 2018; Parks et al., 2013; Podar et al., 2015).

Further analysis of the metagenomic data through BLAST searches revealed potential *hgcA* and *hgcB* genes in samples collected from three sites that were not detected in the PCR amplifications. The greatest abundance of potential *hgcA* genes were discovered in samples collected from the artificial wetlands, with very few potential *hgcA* genes identified in both the natural or flooded sampling sites. The majority of potential genes identified aligned next to species from the class *Delta-proteobacteria* and are evenly distributed among sulfate-reducing and iron-reducing bacteria. Our findings are supported by numerous studies that have identified sulfate-reducing bacteria as primary methylators, however the alignment of potential *hgcA* genes to iron-reducing bacteria, specifically from the family *Geobacteraceae*, is supported by more recent studies highlighting the importance of iron-reducing bacteria in mercury cycling (Acha et al., 2005; Bravo et al., 2018a; Bravo et al., 2018b; Compeau & Bartha, 1987). Even though our data depicts a greater abundance of iron-reducing and sulfate-reducing bacteria, taxa from the phylum *Euryarchaeota* known to methylate mercury were present in the highest abundance in the artificial wetlands. Functional gene markers related to methanogenesis were also found in the greatest abundance at the wetland site. Studies focusing on periphyton have shown that inhibiting sulfate reduction only partially inhibits methylation, however, the addition of methanogenesis inhibitors fully inhibit mercury methylation (Hamelin et al., 2011). Even though only one potential *hgcA* gene was identified, the majority of methanogens contain *hgcA* (different from the believed patchy distribution) (Gilmour et al., 2018). Further analysis of the metagenomic data focused on the unassembled reads using BLAST searches to identify potential *hgcA* and *hgcB*

genes that may not have been assembled into scaffolds. A greater abundance of potential *hgcA* genes were identified from the unassembled data, however we are less confident that they are actual *hgcA* genes based on the amino acid sequence length and alignment to the reference *hgcA* sequences. The potential *hgcAB* genes produced through BLAST searches of the assembled metagenomes have a greater probability of being actual *hgcA* and *hgcB* genes due to the length of the sequences and their alignment to the conserved regions of the reference sequences.

#### **4.4 Conclusions and future work**

Our study aimed to better understand how the construction of the Chute-Allard run-of-river hydro facility could have altered the pathways responsible for the mobilization, transformation and transfer of mercury in the St. Maurice river, or whether other factors altering the landscape around the dammed area, such as forest fires and logging, are responsible for the increased levels of MeHg in the environment. We hypothesized that environments within the St. Maurice river conducive to the formation and growth of periphyton biofilms will have the greatest diversity of potential mercury methylating microorganisms, and therefore the highest potential for increased MeHg production. Our results revealed the greatest abundance of taxa previously associated with mercury methylation, as well as the *hgcAB* gene pair and functional gene markers are present in samples collected from the artificial wetlands, and not from the natural or flooded sampling sites in which if they were present it was in very low abundances. Natural wetlands have been implicated in numerous studies for being environments that support high levels of mercury methylation (e.g., Hamelin et al., 2011) and are have been shown to be a major source of increased MeHg levels in local fish populations (Driscoll et al., 2007; Evers et al., 2011).

Our results suggest that the construction of the dam and subsequent flooding of the artificial wetland channels are responsible for the increased levels of MeHg in local fish populations. Since the forest fires and local logging operations occurred in large areas surrounding the river, we would expect to see less localized microbial communities with the ability to methylate mercury. However, the patterns we have observed could be caused by the artificial wetlands providing the ideal habitat for mercury methylating microorganisms to grow

due to a greater abundance of aquatic plant biomass not present in the natural and flooded sites (Hsu-Kim et al., 2018). Although our research points to the flooding of the artificial wetlands and increased amount of biomass providing a conducive environment for mercury methylating microorganisms, more information is needed to further investigate the cause of increased MeHg levels in the St. Maurice river and to determine where the Hg originates from.

## References

- Achá, D., Hintelmann, H., & Yee, J. (2011). Importance of sulfate reducing bacteria in mercury methylation and demethylation in periphyton from Bolivian Amazon region. *Chemosphere*, 82(6), 911–916. <https://doi.org/10.1016/j.chemosphere.2010.10.050>
- Achá, D., Iñiguez, V., Roulet, M., Dave, J. R., Alanoca, L., Sanchez, S., Guimarães, G. R. D., & Luna, R. (2005). Sulfate-Reducing Bacteria in Floating Macrophyte Rhizospheres from an Amazonian Floodplain Lake in Bolivia and Their Association with Hg Methylation. *Applied and Environmental Microbiology*, 71(11), 7531–7535. <https://doi.org/10.1128/AEM.71.11.7531>
- Ackerman, J. T., & Eagles-Smith, C. A. (2010). Agricultural wetlands as potential hotspots for mercury bioaccumulation: Experimental evidence using caged fish. *Environmental Science and Technology*, 44(4), 1451–1457. <https://doi.org/10.1021/es9028364>
- Aiken, G. R., Gilmour, C. C., Krabbenhoft, D. P., & Orem, W. (2011). Dissolved organic matter in the Florida everglades: Implications for ecosystem restoration. *Critical Reviews in Environmental Science and Technology*, 41(SUPPL. 1), 217–248. <https://doi.org/10.1080/10643389.2010.530934>
- Aménagements hydroélectriques de la Chute-Allard et des Rapides-des-Cœurs*. (2014).
- Apprill, A., McNally, S., Parsons, R., & Weber, L. (2015). Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquatic Microbial Ecology*, 75(2), 129–137. <https://doi.org/10.3354/ame01753>
- 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), 6517–6526. <https://doi.org/10.1128/aem.01666-14>
- Barkay, T., Liebert, C., & Gillman, M. (1989). Environmental significance of the potential for mer(Tn21)-mediated reduction of Hg<sup>2+</sup> to Hg<sup>0</sup> in natural waters. *Applied and Environmental Microbiology*, 55(5), 1196–1202. [https://doi.org/10.1016/S0065-2164\(05\)57001-1](https://doi.org/10.1016/S0065-2164(05)57001-1)
- Barre, J. P. G., Amouroux, D., Perrot, V., Rodríguez Martín-Doimeadios, R. C., Jiménez-Moreno, M., & Bérail, S. (2016). Sources and fate of mercury pollution in Almadén mining

- district (Spain): Evidences from mercury isotopic compositions in sediments and lichens. *Chemosphere*, 147, 430–438. <https://doi.org/10.1016/j.chemosphere.2015.12.094>
- Béliveau, A., Lucotte, M., Davidson, R., do Canto Lopes, L. O., & Paquet, S. (2009). Early Hg mobility in cultivated tropical soils one year after slash-and-burn of the primary forest, in the Brazilian Amazon. *Science of the Total Environment*, 407(15), 4480–4489. <https://doi.org/10.1016/j.scitotenv.2009.04.012>
- Benoit, J. M., Gilmour, C. C., Heyes, A., Mason, R. P., & Miller, C. L. (2003). Geochemical and Biological Controls over Methylmercury Production and Degradation in Aquatic Ecosystems. *American Chemical Society*, 262–297. <https://doi.org/10.1021/bk-2003-0835.ch019>
- Blum, J. D., Popp, B. N., Drazen, J. C., Anela Choy, C., & Johnson, M. W. (2013). Methylmercury production below the mixed layer in the North Pacific Ocean. *Nature Geoscience*, 6(10), 879–884. <https://doi.org/10.1038/ngeo1918>
- Bodaly, R. A., Jansen, W. A., Majewski, A. R., Fudge, R. J. P., Strange, N. E., Derksen, A. J., & Green, D. J. (2007). Postimpoundment time course of increased mercury concentrations in fish in hydroelectric reservoirs of northern Manitoba, Canada. *Archives of Environmental Contamination and Toxicology*, 53(3), 379–389. <https://doi.org/10.1007/s00244-006-0113-4>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bravo, A. G., Bouchet, S., Tolu, J., Björn, E., Mateos-Rivera, A., & Bertilsson, S. (2017). Molecular composition of organic matter controls methylmercury formation in boreal lakes. *Nature Communications*, 8, 1–9. <https://doi.org/10.1038/ncomms14255>
- Bravo, A. G., Peura, S., Buck, M., Ahmed, O., Mateos-Rivera, A., Ortega, S. H., Schaefer, Jeffra, K., Bouchet, S., Tolu, J., Björn, E., & Bertilsson, S. (2018a). Methanogens and iron-reducing bacteria: The overlooked members of mercury-methylating microbial communities in boreal lakes. *Applied and Environmental Microbiology*, 84(23). <https://doi.org/10.1128/AEM.01774-18>
- Bravo, A. G., Zopfi, J., Buck, M., Xu, J., Bertilsson, S., Schaefer, J. K., Poté, J., & Cosio, C. (2018b). Geobacteraceae are important members of mercury-methylating microbial communities of sediments impacted by waste water releases. *ISME Journal*, 12(3), 802–

812. <https://doi.org/10.1038/s41396-017-0007-7>
- Broussard, L. A., Hammett-stabler, C. A., & Winecker, R. E. (2002). Synonym : Formula : Valence : Chemical State : Physical State : Toxicity : *Laboratory Medicine*, 33(8).
- Buchfink, B., Xie, C., & Huson, D. H. (2014). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59–60. <https://doi.org/10.1038/nmeth.3176>
- Callahan, B., McMurdie, P., Rosen, M., Han, A., Johnson, A., & Holmes, S. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, 13, 581–583. <https://doi.org/10.1038/nmeth.3869>
- Celo, V., Lean, D. R. S., & Scott, S. L. (2006). Abiotic methylation of mercury in the aquatic environment. *Science of the Total Environment*, 368(1), 126–137. <https://doi.org/10.1016/j.scitotenv.2005.09.043>
- Christensen, G. A., Wymore, A. M., King, A. J., Podar, M., Hurt, R. A., Santillan, E. U., Soren, A., Brandt, C. C., Brown, S. D., Palumbo, A. V., Wall, J. D., Gilmour, C. C., & Elias, D. A. (2016). Development and Validation of Broad-Range Qualitative and Clade-Specific Quantitative Molecular Probes for Assessing Mercury Methylation in the Environment. *Applied and Environmental Microbiology*, 82(19), 6068–6078. <https://doi.org/10.1128/aem.01271-16>
- Cleckner, L. B., Gilmour, C. C., & Hurley, J. P. (1999). Mercury Methylation in Periphyton of the Florida Everglades. *Krabbenhof Source: Limnology and Oceanography*, 44(7), 1815–1825. Retrieved from <http://www.jstor.org/stable/2670419>
- Compeau, G. C., & Bartha, R. (1987). Effect of salinity on mercury-methylating activity of sulfate-reducing bacteria in estuarine sediments. *Applied and Environmental Microbiology*, 53(2), 261–265.
- Comte, I., Lucotte, M., Davidson, R., Reis De Carvalho, C. J., De Assis Oliveira, F., & Rousseau, G. X. (2013). Impacts of land uses on mercury retention in long-time cultivated soils, brazilian amazon. *Water, Air, and Soil Pollution*, 224(4). <https://doi.org/10.1007/s11270-013-1515-3>
- Correia, R. R. S., Miranda, M. R., & Guimarães, J. R. D. (2012). Mercury methylation and the microbial consortium in periphyton of tropical macrophytes: Effect of different inhibitors. *Environmental Research*, 112, 86–91. <https://doi.org/10.1016/j.envres.2011.11.002>
- De Caceres, M., & Legendre, P. (2009). Associations between species and groups of sites:

- indices and statistical inference. *Ecology*. Retrieved from <http://sites.google.com/site/miqueldecaceres/>
- Desrosiers, M., Planas, D., & Mucci, A. (2006). Mercury methylation in the epilithon of boreal shield aquatic ecosystems. *Environmental Science and Technology*, *40*(5), 1540–1546. <https://doi.org/10.1021/es0508828>
- Driscoll, C. T., Blette, V., Yan, C., Schofield, C. L., Munson, R., & Holsapple, J. (1995). The role of dissolved organic carbon in the chemistry and bioavailability of mercury in remote Adirondack lakes. *Water, Air, & Soil Pollution*, *80*(1–4), 499–508. <https://doi.org/10.1007/BF01189700>
- Driscoll, C. T., Mason, R. P., Chan, H. M., Jacob, D. J., & Pirrone, N. (2013). Mercury as a global pollutant: Sources, pathways, and effects. *Environmental Science and Technology*, *47*(10), 4967–4983. <https://doi.org/10.1021/es305071v>
- Driscoll, Han, Y., Chen, C. Y., Evers, D. C., Lambert, K. F., Holsen, T. M., ... Munson, R. K. (2007). Mercury Contamination in Forest and Freshwater Ecosystems in the Northeastern United States. *BioScience*, *57*(1), 17–28.
- Durnford, D., Dastoor, A., Figueras-Nieto, D., & Ryjkov, A. (2010). Long range transport of mercury to the Arctic and across Canada. *Atmospheric Chemistry and Physics*, *10*(13), 6063–6086. <https://doi.org/10.5194/acp-10-6063-2010>
- Eckley, C. S., & Hintelmann, H. (2006). Determination of mercury methylation potentials in the water column of lakes across Canada. *Science of the Total Environment*, *368*(1), 111–125. <https://doi.org/10.1016/j.scitotenv.2005.09.042>
- Eckley, C. S., Luxton, T. P., Goetz, J., & McKernan, J. (2017). Water-level fluctuations influence sediment porewater chemistry and methylmercury production in a flood-control reservoir. *Environmental Pollution*, *222*, 32–41. <https://doi.org/10.1016/j.envpol.2017.01.010>
- Evers, D. C., Wiener, J. G., Driscoll, C. T., Gay, D. A., Basu, N., Monson, B. A., Lambert, K. F., Morrison, H. A., Morgan, J. T., Williams, K. A., & Soehl, A. G. (2011). *Great Lakes Mercury Connections: The Extent and Effects of Mercury Pollution in the Great Lakes Region*. Gorham, Maine.
- Figueiredo, N., Serralheiro, M. L., Canário, J., Duarte, A., Hintelmann, H., & Carvalho, C. (2018). Evidence of mercury methylation and demethylation by the estuarine microbial

- communities obtained in stable Hg isotope studies. *International Journal of Environmental Research and Public Health*, 15(10). <https://doi.org/10.3390/ijerph15102141>
- Fleming, E. J., Erin Mack, E., Green, P. G., & Nelson, D. C. (2006). Mercury Methylation from Unexpected Sources: Molybdate-Inhibited Freshwater Sediments and an Iron-Reducing Bacterium. *Applied and Environmental Microbiology*, 72(1), 457–464. <https://doi.org/10.1128/AEM.72.1.457>
- Foucher, D., Ogrinc, N., & Hintelmann, H. (2009). Tracing mercury contamination from the idrija mining region (slovenia) to the gulf of trieste using Hg isotope ratio measurements. *Environmental Science and Technology*, 43(1), 33–39. <https://doi.org/10.1021/es801772b>
- Fredrik, H., Braaten, V., Wit, H. A. De, Fjeld, E., Rognerud, S., Lydersen, E., & Larssen, T. (2014). Science of the Total Environment Environmental factors in influencing mercury speciation in Subarctic and Boreal lakes. *Science of the Total Environment*, 476–477, 336–345. <https://doi.org/10.1016/j.scitotenv.2014.01.030>
- Gammons, C. H., Pape, B. L., Parker, S. R., Poulson, S. R., & Blank, C. E. (2013). Geochemistry, water balance, and stable isotopes of a “clean” pit lake at an abandoned tungsten mine, Montana, USA. *Applied Geochemistry*, 36, 57–69. <https://doi.org/10.1016/j.apgeochem.2013.06.011>
- Garcia, E., & Carignan, R. (2005). Mercury concentrations in fish from forest harvesting and fire-impacted Canadian Boreal lakes compared using stable isotopes of nitrogen. *Environmental Toxicology and Chemistry*, 24(3), 685–693. <https://doi.org/10.1897/04-065R.1>
- Garcia, E., Carignan, R., & Lean, D. R. S. (2007). Seasonal and inter-annual variations in methyl mercury concentrations in zooplankton from boreal lakes impacted by deforestation or natural forest fires. *Environmental Monitoring and Assessment*, 131(1–3), 1–11. <https://doi.org/10.1007/s10661-006-9442-z>
- Gilmour, C. C., Bullock, A. L., McBurney, A., Podar, M., & Elias, D. A. (2018). Robust Mercury Methylation across Diverse Methanogenic Archaea. *mBio*, 9(2), 1–13. <https://doi.org/10.1128/mbio.02403-17>
- Gilmour, C. C., Henry, E. A., & Ralph, M. (1992). Sulfate Stimulation of Mercury Methylation in Freshwater Sediments. *Environmental Science and Technology*, 26(11), 2281–2287. <https://doi.org/10.1021/es00035a029>

- Gilmour, C. C., Podar, M., Bullock, A. L., Graham, A. M., Brown, S. D., Somenahally, A. C., ... 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>
- Grégoire, D. S., & Poulain, A. J. (2014). A little bit of light goes a long way: The role of phototrophs on mercury cycling. *Metallomics*, *6*(3), 396–407. <https://doi.org/10.1039/c3mt00312d>
- Hagan, N., Robins, N., Hsu-Kim, H., Halabi, S., Espinoza Gonzales, R. D., Ecos, E., Richter, D., & Vandenberg, J. (2015). Mercury hair levels and factors that influence exposure for residents of Huancavelica, Peru. *Environmental Geochemistry and Health*, *37*(3), 507–514. <https://doi.org/10.1007/s10653-014-9665-9>
- Hagerthey, S. E., Bellinger, B. J., Wheeler, K., Gantar, M., & Gaiser, E. (2011). Everglades periphyton: A biogeochemical perspective. *Critical Reviews in Environmental Science and Technology*, *41*(SUPPL. 1), 309–343. <https://doi.org/10.1080/10643389.2010.531218>
- Hall, B. D., Cherewyk, K. A., Paterson, M. J., & Bodaly, R. (Drew) A. (2009). Changes in methyl mercury concentrations in zooplankton from four experimental reservoirs with differing amounts of carbon in the flooded catchments This paper is part of the series “Forty Years of Aquatic Research at the Experimental Lakes Area”. *Canadian Journal of Fisheries and Aquatic Sciences*, *66*(11), 1910–1919. <https://doi.org/10.1139/f09-123>
- Hamelin, S., Amyot, M., Barkay, T., Wang, Y., & Planas, D. (2011). Methanogens: Principal methylators of mercury in lake periphyton. *Environmental Science and Technology*, *45*(18), 7693–7700. <https://doi.org/10.1021/es2010072>
- Hamelin, S., Planas, D., & Amyot, M. (2015). Spatio-temporal variations in biomass and mercury concentrations of epiphytic biofilms and their host in a large river wetland (Lake St. Pierre, Qc, Canada). *Environmental Pollution*, *197*, 221–230. <https://doi.org/10.1016/j.envpol.2014.11.007>
- Holmes, C. D., Jacob, D. J., Mason, R. P., & Jaffe, D. A. (2009). Sources and deposition of reactive gaseous mercury in the marine atmosphere. *Atmospheric Environment*, *43*(14), 2278–2285. <https://doi.org/10.1016/j.atmosenv.2009.01.051>
- Hsu-Kim, H., Eckley, C. S., Achá, D., Feng, X., Gilmour, C. C., Jonsson, S., & Mitchell, C. P. J. (2018). Challenges and opportunities for managing aquatic mercury pollution in altered

- landscapes. *Ambio*, 47(2), 141–169. <https://doi.org/10.1007/s13280-017-1006-7>
- Huson, D. H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., Ruscheweyh, H. J., & Tappu, R. (2016). MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLoS Computational Biology*, 12(6), 1–12. <https://doi.org/10.1371/journal.pcbi.1004957>
- Johnson, N. W., Mitchell, C. P. J., Engstrom, D. R., Bailey, L. T., Coleman Wasik, J. K., & Berndt, M. E. (2016). Methylmercury production in a chronically sulfate-impacted sub-boreal wetland. *Environmental Science: Processes and Impacts*, 18(6), 725–734. <https://doi.org/10.1039/c6em00138f>
- Kamman, N. C., Burgess, N. M., Driscoll, C. T., Simonin, H. A., Goodale, W., Linehan, J., Estabrook, R., Hutcheson, M., Major, A., Scheuhammer, A. M., & Scruton, D. A. (2005). Mercury in freshwater fish of northeast North America - A geographic perspective based on fish tissue monitoring databases. *Ecotoxicology*, 14(1–2), 163–180. <https://doi.org/10.1007/s10646-004-6267-9>
- Kelly, C. A., Rudd, J. W. M., Bodaly, R. A., Roulet, N. P., St.Louis, V. L., Heyes, A., Moore, T. R., Schiff, S., Aravena, R., Scott, K. J., Dyck, B., Harris, R., Warner, B., & Edwards, G. (1997). Increases in fluxes of greenhouse gases and methyl mercury following flooding of an experimental reservoir. *Environmental Science and Technology*, 31(5), 1334–1344. <https://doi.org/10.1021/es9604931>
- Kerry, A., Welbourn, P. M., Prucha, B., & Miede, G. (1991). Mercury Methylation by Sulfate-Reducing Bacteria from Sediments of an Acid Stressed Lake. *Water, Air, & Soil Pollution*, 56, 565–575.
- Khadra, M., Planas, D., Girard, C., & Amyot, M. (2018). Age matters: Submersion period shapes community composition of lake biofilms under glyphosate stress. *Facets*, 3(1), 934–951. <https://doi.org/10.1139/facets-2018-0019>
- Lei, P., Nunes, L. M., Liu, Y. R., Zhong, H., & Pan, K. (2019). Mechanisms of algal biomass input enhanced microbial Hg methylation in lake sediments. *Environment International*, 126(January), 279–288. <https://doi.org/10.1016/j.envint.2019.02.043>
- Li, D., Liu, C. M., Luo, R., Sadakane, K., & Lam, T. W. (2015). MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*, 31(10), 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>

- Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, 26(5), 589–595. <https://doi.org/10.1093/bioinformatics/btp698>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Li, Y., Mao, Y., Liu, G., Tachiev, G., Roelant, D., Feng, X., & Cai, Y. (2010). Degradation of methylmercury and its effects on mercury distribution and cycling in the Florida everglades. *Environmental Science and Technology*, 44(17), 6661–6666. <https://doi.org/10.1021/es1010434>
- Liu, Y. R., Yu, R. Q., Zheng, Y. M., & He, J. Z. (2014). Analysis of the microbial community structure by monitoring an Hg methylation gene (*hgcA*) in paddy soils along an Hg gradient. *Applied and Environmental Microbiology*, 80(9), 2874–2879. <https://doi.org/10.1128/AEM.04225-13>
- Lozano, V. L., Vinocur, A., Sabio y García, C. A., Allende, L., Cristos, D. S., Rojas, D., Wolansky, M., & Pizarro, H. (2018). Effects of glyphosate and 2,4-D mixture on freshwater phytoplankton and periphyton communities: a microcosms approach. *Ecotoxicology and Environmental Safety*. <https://doi.org/10.1016/j.ecoenv.2017.12.006>
- Magurran, A. E. (2005). *Measuring Biological Diversity*. Blackwell Publishing Company. <https://doi.org/10.1021/es00038a601>
- Matilainen, T., Verta, M., Niemi, M., & Uusi-Rauva, A. (1991). Specific rates of net methylmercury production in lake sediments. *Water, Air, & Soil Pollution*, 56, 595–605.
- Mauro, J. B. N., Guimaraes, J. R. D., Hintelmann, H., Watras, C., Haack, E., & Coelho-Souza, S. (2002). Mercury methylation in macrophytes, periphyton, and water - comparative studies with stable and radio-mercury additions. *Analytical and Bioanalytical Chemistry*, 374(6), 983–989. <https://doi.org/10.1007/s00216-002-1534-1>
- McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE*, 8(4), e61217. Retrieved from <http://dx.plos.org/10.1371/journal.pone.0061217>
- Meng, B., Feng, X., Qiu, G., Liang, P., Li, P., Chen, C., & Shang, L. (2011). The process of methylmercury accumulation in rice (*Oryza sativa* L.). *Environmental Science and Technology*, 45(7), 2711–2717. <https://doi.org/10.1021/es103384v>

- Morel, F. M. M., Kraepiel, A. M. L., & Amyot, M. (2002). The Chemical Cycle and Bioaccumulation of Mercury. *Annual Review of Ecology and Systematics*, 29(1), 543–566. <https://doi.org/10.1146/annurev.ecolsys.29.1.543>
- Munthe, J., Bodaly, R. A. D., Branfireun, B. A., Driscoll, C. T., Gilmour, C. C., & Harris, R. (2007). Recovery of Mercury-Contaminated Fisheries, 33–44.
- Ni Chadhain, S., Schaefer, J. K., Crane, S., Zylstra, G. J., & Barkay, T. (2006). Analysis of mercuric reductase, 8(10), 1746–1752. <https://doi.org/10.1111/j.1462-2920.2006.01114.x>
- Obrist, D., Kirk, J. L., Zhang, L., Sunderland, E. M., Jiskra, M., & Selin, N. E. (2018). A review of global environmental mercury processes in response to human and natural perturbations: Changes of emissions, climate, and land use. *Ambio*, 47(2), 116–140. <https://doi.org/10.1007/s13280-017-1004-9>
- Oksanen, J., Guillaume, F. B., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., & Wagner, H. (2013). R package Vegan : ecological diversity, 1–12. Retrieved from <http://www2.uaem.mx/r-mirror/web/packages/vegan/vignettes/diversity-vegan.pdf>
- Olsen, T. A., Brandt, C. C., & Brooks, S. C. (2016). Periphyton Biofilms Influence Net Methylmercury Production in an Industrially Contaminated System. *Environmental Science and Technology*, 50(20), 10843–10850. <https://doi.org/10.1021/acs.est.6b01538>
- Oremland, R. S., Miller, L. G., Dowdle, P., Connell, T., & Barkay, T. (1995). Methylmercury oxidative degradation potentials in contaminated and pristine sediments of the Carson River, Nevada. *Applied and Environmental Microbiology*, 61(7), 2745–2753.
- Parada, A. E., Needham, D. M., & Fuhrman, J. A. (2016). Every base matters: Assessing small subunit rRNA primers for marine microbiomes with mock communities, time series and global field samples. *Environmental Microbiology*, 18(5), 1403–1414. <https://doi.org/10.1111/1462-2920.13023>
- Parks, J. M., Guo, H., Momany, C., Liang, L., Miller, S. M., Summers, A. O., & Smith, J. C. (2009). Mechanism of Hg-C protonolysis in the organomercurial lyase MerB. *Journal of the American Chemical Society*, 131(37), 13278–13285. <https://doi.org/10.1021/ja9016123>
- Parks, J. M., Johs, A., Podar, M., Bridou, R., Hurt, R. A., Smith, S. D., Tomanicek, Stephen J., Qian, Y., Brown, S. D., Brandt, C. C., Palumbo, A. V., Smith, J. C., Wall, J. D., Elias, D.

- A., & Liang, L. (2013). The genetic basis for bacterial mercury methylation. *Science*, 339(6125), 1332–1335. <https://doi.org/10.1126/science.1230667>
- Pirrone, N., Cinnirella, S., Feng, X., Finkelman, R. B., Friedli, H. R., Leaner, J., Mason, R., Mukherjee, A. B., Stracher, G. B., Streets, D. G., & Telmer, K. (2010). Global mercury emissions to the atmosphere from anthropogenic and natural sources. *Atmospheric Chemistry and Physics*, 10(13), 5951–5964. <https://doi.org/10.5194/acp-10-5951-2010>
- 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), e1500675. <https://doi.org/10.1126/sciadv.1500675>
- Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., & Glöckner, F. O. (2007). SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, 35(21), 7188–7196. <https://doi.org/10.1093/nar/gkm864>
- Robinson, J. B., & Tuovinen, O. H. (1984). Mechanisms of microbial resistance and detoxification of mercury and organomercury compounds: physiological, biochemical, and genetic analyses. *Microbiological Reviews*, 48(2), 95–124. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/6377034> <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC373215>
- Roulet, M., Guimarães, J.-R. D., & Lucotte, M. (2001). Methylmercury Production and Accumulation in Sediments and Soils of an Amazonian Floodplain - Effect of Seasonal Inundation. *Water, Air, and Soil Pollution*, 128(3), 41–60.
- Rudd, J. W. M., & Furutani, A. (1980). Measurement of mercury methylation in lake water and sediment samples. *Applied and Environmental Microbiology*, 40(4), 770. Retrieved from <http://aem.asm.org/cgi/content/abstract/40/4/770>
- Rytuba, J. J. (2000). Mercury mine drainage and processes that control its environmental impact. *Science of the Total Environment*, 260(1–3), 57–71. [https://doi.org/10.1016/S0048-9697\(00\)00541-6](https://doi.org/10.1016/S0048-9697(00)00541-6)
- 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., Yagi, J., Reinfelder, J. R., Cardona, T., Ellickson, K. M., Tel-Or, S., & Barkay, T. (2004). Role of the bacterial organomercury lyase (MerB) in controlling methylmercury accumulation in mercury-contaminated natural waters. *Environmental Science and Technology*, *38*(16), 4304–4311. <https://doi.org/10.1021/es049895w>
- Selin, N. E., Wu, S., Nam, K. M., Reilly, J. M., Paltsev, S., Prinn, R. G., & Webster, M. D. (2009). Global health and economic impacts of future ozone pollution. *Environmental Research Letters*, *4*(4). <https://doi.org/10.1088/1748-9326/4/4/044014>
- Simpson, E. (1949). Measurement of diversity. *Nature*, *163*(1943), 688.
- Sprovieri, F., Pirrone, N., Hedgecock, I. M., Landis, M. S., & Stevens, R. K. (2002). Intensive atmospheric mercury measurements at Terra Nova Bay in Antarctica during November and December 2000. *Journal of Geophysical Research Atmospheres*, *107*(23), 1–9. <https://doi.org/10.1029/2002JD002057>
- Sunderland, E. M., Krabbenhoft, D. P., Moreau, J. W., Strode, S. A., & Landing, W. M. (2009). Mercury sources, distribution, and bioavailability in the North Pacific Ocean: Insights from data and models. *Global Biogeochemical Cycles*, *23*(2), 1–14. <https://doi.org/10.1029/2008GB003425>
- Tamashiro, H., Akagi, H., Arakaki, M., Futatsuka, M., & Roht, L. H. (1984). Causes of death in Minamata disease: analysis of death certificates. *International Archives of Occupational and Environmental Health*, *54*(2), 135–146. <https://doi.org/10.1007/BF00378516>
- Tarkowska-Kukuryk, M., & Mieczan, T. (2012). Effect of substrate on periphyton communities and relationships among food web components in shallow hypertrophic lake. *Journal of Limnology*, *71*(2), 279–290. <https://doi.org/10.4081/jlimnol.2012.e30>
- Tlili, A., Corcoll, N., Bonet, B., Morin, S., Montuelle, B., Bérard, A., & Guasch, H. (2011). In situ spatio-temporal changes in pollution-induced community tolerance to zinc in autotrophic and heterotrophic biofilm communities. *Ecotoxicology*, *20*(8), 1823–1839. <https://doi.org/10.1007/s10646-011-0721-2>
- Todorova, S. G., Driscoll, C. T., Matthews, D. A., Effler, S. W., Hines, M. E., & Henry, E. A. (2009). Evidence for regulation of monomethyl mercury by nitrate in a seasonally stratified, eutrophic lake. *Environmental Science and Technology*, *43*(17), 6572–6578. <https://doi.org/10.1021/es900887b>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Yu, R. Q., Flanders, J. R., MacK, E. E., Turner, R., Mirza, M. B., & Barkay, T. (2012). Contribution of coexisting sulfate and iron reducing bacteria to methylmercury production in freshwater river sediments. *Environmental Science and Technology*, 46(5), 2684–2691. <https://doi.org/10.1021/es2033718>

## 5. Supplementary Material and Methods

### 5.1 *hgcAB* gene pair analysis

PCR amplification was completed on all periphyton samples using broad-range and clade-specific primers that target the *hgcA* and *hgcB* genes to determine their presence or absence using primers found in the literature. All primer sets were optimized using temperature gradients prior to amplifying environmental samples. The *hgcAB* genes were targeted using the ORNL-HgcA-uni-F (5' AAYGTCTGGTGYGCNGCVGG 3') and ORNL-hgcB-uni-R (5' CABGCNCCRCAYTCCATRCA 3') spanning across both the *hgcA* and *hgcB* genes (Christensen *et al.*, 2016). The 25  $\mu$ L PCR reaction consisted of 15.5  $\mu$ L autoclaved milliQ water, 5  $\mu$ L of 5x Phusion Reaction HF Buffer, 0.5  $\mu$ L dNTPs (10 mM), 0.5  $\mu$ L Phusion HF DNA Polymerase (2000 U/mL), 1.25  $\mu$ L Forward Primer, 1.25  $\mu$ L Reverse Primer, and 1  $\mu$ L of extracted periphyton DNA. 1  $\mu$ L of autoclaved milliQ dH<sub>2</sub>O was used as a template for the negative reaction and 1  $\mu$ L of isolated *Geobacter sulfurreducens* PCA was used as a template for the positive control reaction. The reactions were incubated at 98 °C for 1 minute, and cycled 35 times at 98 °C for 10 seconds, 63.2 °C for 30 seconds, and 72 °C for 20 seconds, then incubated at 72 °C for 5 minutes. Following the PCR, 1  $\mu$ L of loading dye was added to each sample and 5  $\mu$ L was loaded and run on a 1 % agarose gel containing 0.03  $\mu$ L/mL ethidium bromide for 45 minutes at 90 volts. A GeneRuler 1 kb DNA Plus Ladder (Thermo Scientific) was used to compare the size of the PCR products against the known sizes of the bands within the ladder.

Conducted PCR to amplify clade-specific *hgcA* genes from the periphyton samples using the primers, conditions and reaction mixture are outlined by Christensen and colleagues (2016). Following the PCR, 1  $\mu$ L of loading dye was added to 5  $\mu$ L of each sample and 5  $\mu$ L was loaded and run on a 1 % agarose gel containing 0.03  $\mu$ L/mL ethidium bromide for 45 minutes at 90 volts. A GeneRuler 1kb DNA Plus Ladder (Thermo Scientific) was used to compare the size of the PCR products against the known sizes of the bands within the ladder.

## 5.2 Dilution series

A dilution series was constructed using known *hgcAB* gene copy numbers per sample of *Geobacter sulfurreducens PCA* extracted from pure cultures to determine detection limit of primers and to rule out environmental DNA inhibiting *hgcAB* amplification. *Geobacter sulfurreducens PCA* was diluted into  $1 \times 10^9$  copy number/sample,  $1 \times 10^7$  copy number/sample,  $1 \times 10^5$  copy number/sample,  $1 \times 10^3$  copy number/sample, and  $1 \times 10$  copy number/sample. A PCR was completed using the ORNL-HgcAB-uni forward and reverse primers on each previously created dilution, as well as on each dilution with  $1 \mu\text{L}$  of environmental DNA following the PCR conditions above and then visualized through gel electrophoresis.

## 5.3 Clade-Specific primer development

Using the primers developed by others to amplify the *hgcAB* genes in the literature, I developed new, shorter, and less degenerate clade-specific primers (Table S7b). The primers were developed by using the same conserved region of the *hgcA* gene, but by either shifting the primers slightly or shortening them to reduce the level of degeneracy. Multiple primers were also created for each clade to be used in conjunction with each other, creating fewer degenerate primers with a higher level of coverage from previously identified mercury methylating species with the *hgcA* genes. The primers were ordered from Integrated DNA Technologies (IDT) and optimized using the same reaction mixture and thermocycler conditions outlined above. The primers were tested on the environmental periphyton DNA and visualized by gel electrophoresis following the same protocols outlined above.

## 6. Supplementary Results

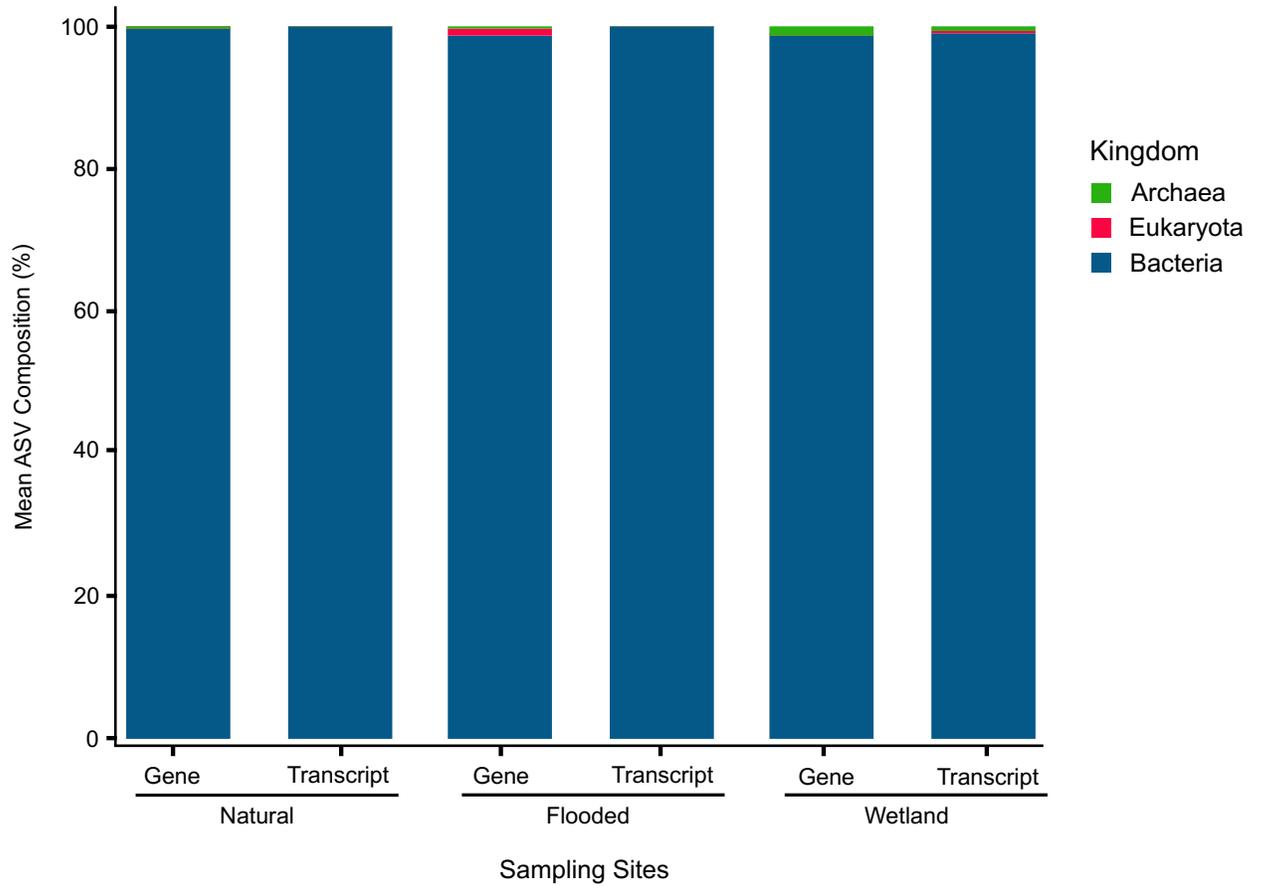
### 6.1 *hgcAB* gene pair analysis

PCR amplification was completed using various clade-specific primers, primers that target a specific class of bacteria, and broad-range primers, primers that target the *hgcAB* gene pair across all classes. The primers used were a mixture of those found in the literature (Table

S7a) as well as newly designed primers (Table S7b). No amplification was seen from any primer pair used following PCR amplification from the extracted DNA of periphyton samples. The dilution series revealed that the detection limits of the primers created by Christensen and colleagues (2016) was around  $1 \times 10^6$  gene copies/sample. The addition of environmental DNA into the dilution series confirmed that the periphyton DNA did not inhibit the PCR reaction preventing amplification. Further PCR amplification was completed using newly designed, less degenerate and more inclusive primers to try and amplify the *hgcA* gene from periphyton samples. No amplification was seen from any primer pair following PCR amplification.

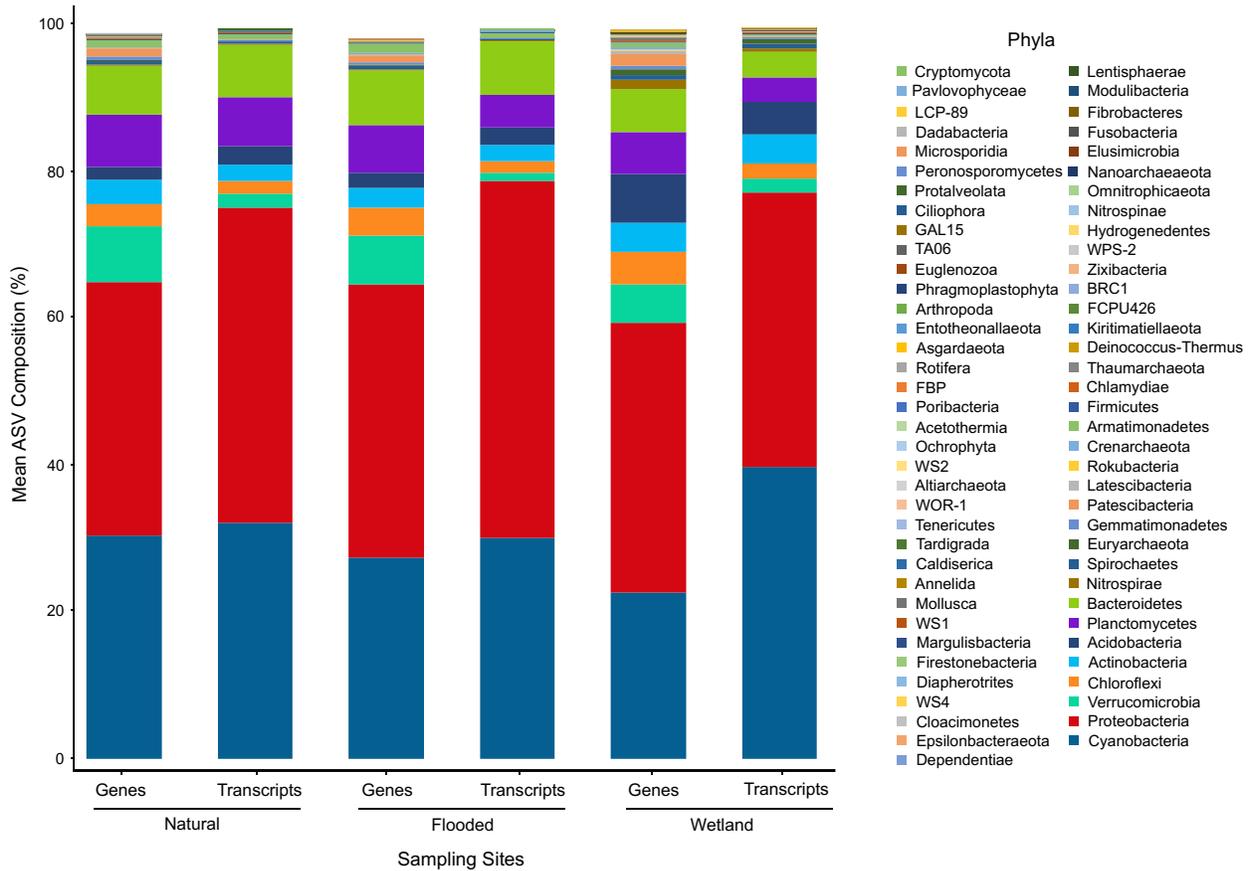
## Supplementary Figures and Tables

**Figure S1**



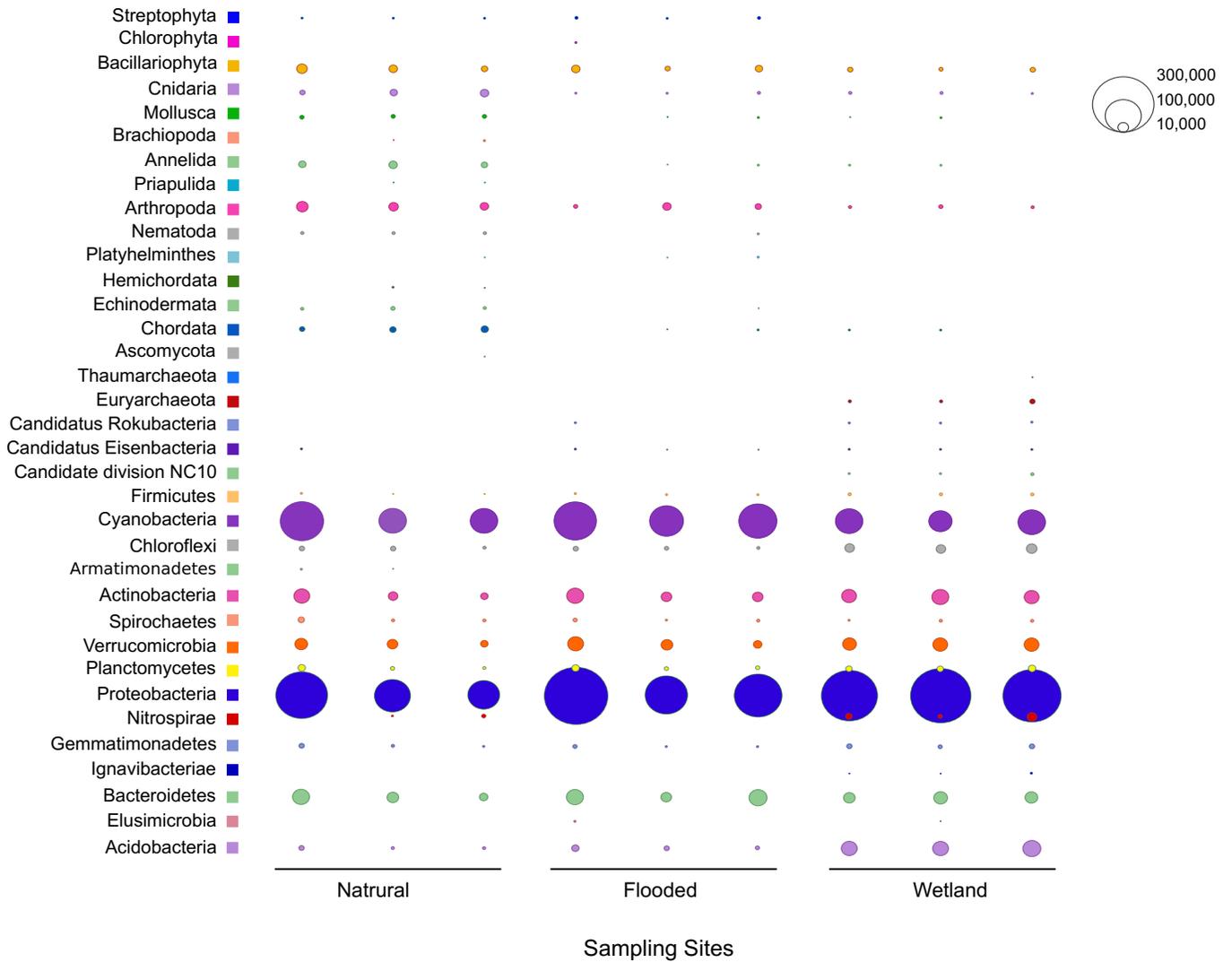
**Figure S1.** Bar plot representing the mean relative percent abundance from 16S rDNA and rRNA for each sampling site at the kingdom level. Over 90% of all samples are Bacteria, with less than 5% being Eukaryotes and Archaea. Archaea were only detected at the wetland site.

**Figure S2**



**Figure S2.** Bar plot representing the mean relative percent abundance from 16S rDNA and rRNA for each sampling site at the phylum level from the bacterial kingdom. The most abundant phyla across all sampling sites for both 16S rDNA and rRNA are cyanobacteria, proteobacteria, verrucomicrobia, chloroflexi, actinobacteria, acidobacteria and planctomycetes.

**Figure S3**



**Figure S3.** Plot illustrating the number of assigned taxa at the phylum rank for each replicate at every sampling site of the metagenomic data. Proteobacteria and cyanobacteria are the most abundant phyla across all sampling sites.

**Table S1**

**A.**

Amplicons	Unique ASVs	Read Counts
S1-A-t0-C	634	15767
S1-A-t0-G	2065	88652
S1-B-t0-C	637	27798
S1-B-t0-G	1715	108171
S1-C-t0-C	831	21725
S1-C-t0-G	1784	85270
S1-E1-C	1236	39411
S1-E1-G	2166	90217
S1-E2-C	944	31430
S1-E2-G	1515	52962
S1-E3-C	921	27202
S1-E3-G	2071	87246
S1-F1-C	1291	47216
S1-F1-G	1965	76693
S1-F2-C	685	20445
S1-F2-G	2106	79903
S1-F3-C	1561	54405
S1-F3-G	1912	76424
S2-A-t0-C	1315	54841
S2-A-t0-G	2065	87106
S2-B-t0-C	932	32355
S2-B-t0-G	1910	93199
S2-C-t0-C	1255	48928
S2-C-t0-G	2004	84334
S2-E1-C	1211	46075
S2-E1-G	2294	103213
S2-E2-C	1248	61973
S2-E2-G	1874	79028
S2-E3-C	945	40082
S2-E3-G	1787	78512
S2-F1-C	1399	68476
S2-F1-G	2048	87705
S2-F2-C	882	31278
S2-F2-G	2051	86844
S2-F3-C	573	13225
S2-F3-G	2060	76124
S3-A-t0-C	1423	89393
S3-A-t0-G	1500	64177
S3-B-t0-C	1304	87940
S3-B-t0-G	1876	79395
S3-C-t0-C	1587	92394
S3-C-t0-G	2161	81785
S3-E1-C	1300	97868
S3-E1-G	1934	87424
S3-E2-C	1398	102319
S3-E2-G	1610	70436
S3-E3-C	1022	48427
S3-E3-G	1756	80729
S3-F1-C	1941	108306
S3-F1-G	2085	84692
S3-F2-C	1630	93276
S3-F2-G	2131	83894
S3-F3-C	1677	109341
S3-F3-G	1900	73110

**B.**

Genes	Unique ASVs	Read Counts
S1-t0	1854.67	94031.00
S1-E	1917.33	76808.33
S1-F	1994.33	77673.33
S2-t0	1993.00	88213.00
S2-E	1985.00	86917.67
S2-F	2053.00	83557.67
S3-t0	1845.67	75119.00
S3-E	1766.67	79529.67
S3-F	2038.67	80565.33
S1	1922.11	82837.56
S2	2010.33	86229.44
S3	1883.67	78404.67
Overall	1938.70	82490.56

**C.**

Transcripts	Unique ASVs	Read Counts
S1-t0	700.67	21763.33
S1-E	1033.67	32681.00
S1-F	1179.00	40688.67
S2-t0	1167.33	45374.67
S2-E	1134.67	49376.67
S2-F	951.33	37659.67
S3-t0	1438.00	89909.00
S3-E	1240.00	82871.33
S3-F	1749.33	103641.00
S1	971.11	31711.00
S2	1084.44	44137.00
S3	1475.78	92140.44
Overall	1177.11	55996.15

**Table S1A.** 16S rRNA amplicon sequencing data showing the number of unique ASVs and read counts for each sampling site. Samples ending in -C are 16S rRNA amplicons and those ending in -G are 16S rDNA amplicons. **B.** Average number of unique ASVs and read counts based on 16S rDNA sequences for each treatment at all three sampling sites. **C.** Average number of unique ASVs and read counts based on 16S rRNA sequences for each treatment at all three sampling sites.

**Table S2**

Kingdoms	Natural Site		Flooded Site		Wetland Site	
	Genes	Transcripts	Genes	Transcripts	Genes	Transcripts
Bacteria	0.9971953	0.9992502	0.9877897	0.9987388	0.9861049	0.9917939
Eukaryota	0.001145482	0.00038893	0.009978546	0.00090879	0.000734083	0.000639121
Archaea	0.000539208	0.00015067	0.001224124	0.00025174	0.012599874	0.007394509

**Table S2.** Mean relative percent abundance of each kingdom for the 16S rDNA and rRNA at each sampling site. Values used to create Figure S1.

Table S3

Phyla	Natural Site		Flooded Site		Wetland Site	
	Genes	Transcripts	Genes	Transcripts	Genes	Transcripts
Cyanobacteria	0.3039751	0.3221315	0.273258	0.3005692	0.2263187	0.397532
Proteobacteria	0.3458791	0.4299104	0.3741594	0.4882575	0.3679118	0.374672
Verrucomicrobia	0.07692029	0.01816404	0.06561564	0.01102879	0.05211424	0.0193931
Chloroflexi	0.03001457	0.01860553	0.03837179	0.01532602	0.04548907	0.02046514
Actinobacteria	0.03372866	0.02152075	0.02818578	0.02208276	0.0392111	0.04024653
Acidobacteria	0.01697432	0.02510871	0.02038747	0.02367376	0.06748465	0.04413311
Planctomycetes	0.07192122	0.06685027	0.06476648	0.04502899	0.05655276	0.03352491
Bacteroidetes	0.06552181	0.07236185	0.0750633	0.07367716	0.0592836	0.03392647
Nitrospirae	0.00146874	0.00080589	0.00107723	0.00043551	0.0116107	0.00566888
Spirochaetes	0.00621699	0.00278557	0.00506787	0.0017496	0.00736776	0.00607647
Euryarchaeota	2.08E-04	9.11E-05	7.86E-04	1.51E-04	7.82E-03	6.61E-03
Gemmatimonadetes	0.00550475	0.00252629	0.00412079	0.00189058	0.00468651	0.00277475
Patescibacteria	0.0108861	0.00043098	0.00855598	0.00016615	0.01621077	0.00021947
Latescibacteria	0.00185772	0.00034338	0.00239542	0.00037761	0.00387165	0.00026168
Rokubacteria	5.23E-05	2.10E-05	1.61E-04	1.26E-05	2.42E-03	7.25E-04
Crenarchaeota	1.74E-05	0.00E+00	2.94E-04	7.55E-06	3.10E-03	1.91E-04
Armatimonadetes	0.00943077	0.00690262	0.01384034	0.0062734	0.00557507	0.00173407
Firmicutes	0.00116024	0.00119482	0.00061206	0.00043048	0.00229436	0.00222969
Chlamydiae	0.00231242	0.00249125	0.00070226	0.00089872	0.0023879	0.00094662
Thaumarchaeota	0.00025619	0	0	0	0.00064055	6.3912E-05
Deinococcus-Thermus	2.98E-04	1.05E-05	2.65E-04	0.00E+00	1.84E-05	0.00E+00
Kiritimatiellaeota	5.63E-05	7.01E-05	3.89E-04	2.95E-04	1.51E-03	8.53E-04
Fer1	4.00E-04	4.24E-04	5.15E-05	1.51E-05	4.46E-04	7.60E-05
BRC1	6.21E-04	2.98E-04	7.40E-04	2.27E-04	3.00E-04	2.05E-05
Zixibacteria	2.86E-04	2.45E-05	4.12E-05	2.52E-05	6.11E-04	2.07E-04
WPS-2	4.48E-04	8.06E-05	6.49E-04	1.36E-04	7.00E-04	1.71E-04
Hydrogenedentes	1.89E-04	2.10E-05	1.64E-04	8.81E-05	6.24E-04	8.32E-05
Nitrospinae	2.01E-05	0.00E+00	6.44E-06	0.00E+00	3.50E-04	9.65E-06
Omnitrophicaeota	4.56E-05	2.42E-04	1.24E-04	2.39E-04	9.68E-04	5.88E-04
Nanoarchaeaeota	5.50E-05	5.96E-05	1.44E-04	9.31E-05	8.90E-04	4.90E-04
Elusimicrobia	1.37E-04	5.61E-05	3.40E-04	6.80E-05	1.19E-03	2.85E-04
Fusobacteria	5.37E-06	0.00E+00	4.77E-05	0.00E+00	1.19E-04	9.65E-05
Fibrobacteres	1.22E-04	8.41E-05	1.22E-04	7.55E-05	4.70E-04	6.97E-04
Modulibacteria	5.37E-06	1.75E-05	0.00E+00	0.00E+00	1.39E-04	4.46E-05
Lentisphaerae	1.07E-05	0.00E+00	6.44E-06	5.03E-06	2.71E-04	1.76E-04
Dependentiae	0.00069212	0.00065172	0.00054506	0.00041286	0.00059945	0.00035935
Epsilonbacteraeota	4.83E-05	1.47E-04	0.00E+00	2.01E-05	8.50E-06	9.65E-06
Cloacimonetes	0.00E+00	0.00E+00	3.87E-06	0.00E+00	1.83E-04	2.29E-05
WS4	0.00E+00	0.00E+00	1.29E-05	0.00E+00	1.60E-04	0.00E+00
Diapherotrites	0.00E+00	0.00E+00	0.00E+00	0.00E+00	8.36E-05	0.00E+00
Firestonebacteria	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.70E-05	3.50E-05
Margulisbacteria	5.77E-05	2.45E-05	7.47E-05	5.54E-05	1.98E-05	9.65E-06
WS1	2.68E-06	0.00E+00	0.00E+00	0.00E+00	1.49E-04	1.21E-05
Mollusca	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.65E-05
Annelida	0.00E+00	1.02E-04	0.00E+00	0.00E+00	0.00E+00	2.41E-05
Caldisericia	0.00E+00	0.00E+00	0.00E+00	0.00E+00	3.54E-05	1.21E-05
Tardigrada	0.00E+00	1.05E-05	0.00E+00	0.00E+00	0.00E+00	3.62E-05
Tenericutes	1.61E-05	0.00E+00	2.45E-05	7.55E-06	1.84E-05	3.62E-06
WOR-1	0.00E+00	0.00E+00	0.00E+00	0.00E+00	5.10E-05	0.00E+00
Altiaarchaeota	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.13E-05	2.29E-05
WS2	0.00E+00	0.00E+00	3.87E-06	0.00E+00	5.67E-05	0.00E+00
Ochrophyta	0.00E+00	3.50E-05	0.00E+00	1.13E-04	0.00E+00	6.51E-05
Acetothermia	5.37E-06	0.00E+00	0.00E+00	0.00E+00	2.41E-05	0.00E+00
Poribacteria	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.28E-05	0.00E+00
FBP	9.39E-06	7.01E-06	3.87E-06	5.03E-06	8.50E-06	0.00E+00
Rotifera	0.00E+00	0.00E+00	0.00E+00	1.51E-05	0.00E+00	0.00E+00
Asgardaeota	0.00E+00	0.00E+00	0.00E+00	0.00E+00	8.50E-06	1.09E-05
Entotheonellaeota	2.68E-06	1.75E-05	0.00E+00	0.00E+00	0.00E+00	1.57E-05
Arthropoda	0.00E+00	1.75E-05	0.00E+00	1.26E-05	0.00E+00	0.00E+00
Phragmoplastophyta	0.00E+00	0.00E+00	0.00E+00	2.01E-05	0.00E+00	3.62E-06
Euglenozoa	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.65E-05
TA06	0.00E+00	0.00E+00	0.00E+00	0.00E+00	9.92E-06	0.00E+00
GAL15	9.39E-06	0.00E+00	0.00E+00	0.00E+00	4.25E-06	0.00E+00
Ciliophora	0.00E+00	1.40E-05	0.00E+00	1.51E-05	0.00E+00	1.09E-05
Protalveolata	0.00E+00	1.40E-05	0.00E+00	0.00E+00	0.00E+00	0.00E+00
Peronosporomycetes	0.00E+00	0.00E+00	0.00E+00	1.01E-05	0.00E+00	0.00E+00
Microsporidia	0.00E+00	0.00E+00	0.00E+00	1.01E-05	0.00E+00	0.00E+00
Dadabacteria	2.68E-06	0.00E+00	0.00E+00	0.00E+00	5.67E-06	0.00E+00
LCP-89	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.83E-06	3.62E-06
Pavlovophyceae	0.00E+00	0.00E+00	0.00E+00	5.03E-06	0.00E+00	0.00E+00
Cryptomycota	0.00E+00	0.00E+00	0.00E+00	5.03E-06	0.00E+00	0.00E+00

**Table S3.** Mean relative percent abundance of each phyla for the 16S rDNA and rRNA at each sampling site. Values used to create Figure S2.

**Table S4**

Families	Natural Site		Flooded Site		Wetland Site	
	Genes	Transcripts	Genes	Transcripts	Genes	Transcripts
Geobacteraceae	0.0003098	0.0005186	0.0012022	0.0006142	0.0213394	0.0150447
Desulfobacteraceae	5.37E-06	4.20E-05	2.36E-04	2.54E-04	2.89E-03	3.78E-03
Desulfobulbaceae	0.00E+00	4.20E-05	1.22E-04	8.06E-05	1.44E-03	1.65E-03
Syntrophorhabdaceae	9.39E-06	5.26E-05	9.02E-05	6.55E-05	1.33E-03	1.15E-03
Desulfovibrionaceae	2.28E-05	1.40E-05	6.44E-06	0.00E+00	5.17E-04	3.20E-04
Methanoregulaceae	5.37E-06	3.50E-05	5.15E-05	5.03E-06	2.39E-03	2.73E-03
Clostridiaceae_1	0.000672	0.000722	0.000269	0.000325	0.000554	0.001143
Methanomassiliococcaceae	9.39E-06	0.00E+00	1.68E-05	0.00E+00	6.05E-04	2.56E-04
Veillonellaceae	1.35E-04	1.75E-05	2.06E-05	0.00E+00	7.55E-04	6.03E-06
Ruminococcaceae	1.48E-05	1.05E-04	5.15E-05	3.52E-05	2.00E-04	2.48E-04
Methanocellaceae	6.71E-06	0.00E+00	8.89E-05	7.55E-06	9.78E-05	1.09E-05
Peptococcaceae	3.22E-05	0.00E+00	2.58E-05	0.00E+00	3.54E-05	4.10E-05
Dehalococcoidaceae	0.00E+00	0.00E+00	0.00E+00	5.03E-06	4.82E-05	0.00E+00
Methanospirillaceae	0.00E+00	0.00E+00	0.00E+00	0.00E+00	2.13E-05	2.29E-05
Syntrophomonadaceae	0.00E+00	0.00E+00	0.00E+00	0.00E+00	1.13E-05	0.00E+00

**Table S4.** Mean relative percent abundance of each family previously implicated in mercury methylation found in the 16S rDNA and rRNA dataset at each sampling site. Values used to create Figure 6.

**Table S5**

	Nanuq			Fasta File	JGI Gold		Calculated	
	Reads from Nanuq	Bases from Nanuq	/302 for reads	Initial Reads in Unassembled	Number of Bases	Read Length	Number of assembled reads	
S1-A-t0	37066516	11194087832	37066516	35496284	665719195	302	2204368.195	6.210137927
S1-E1	27358328	8262215056	27358328	26159976	603680438	302	1998941.848	7.641222024
S1-F1	26833355	8103673210	26833355	25674618	574155719	302	1901177.877	7.404892558
S2-A-t0	29568800	8929777600	29568800	28233194	361691402	302	1197653.649	4.242005524
S2-E1	22427525	6773112550	22427525	21229379	206395554	302	683428.9868	3.219260379
S2-F1	24024723	7255466346	24024723	23054374	270967798	302	897244.3642	3.891861754
S3-A-t0	29072668	8779945736	29072668	27827250	279842668	302	926631.351	3.329942236
S3-E1	31083212	9387130024	31083212	29742390	309748739	302	1025658.076	3.448472285
S3-F1	29466714	8898947628	29466714	28192456	285549434	302	945527.9272	3.353833122
							Average	4.749069756
							S1 Average	7.085417503
							S2 Average	3.784375885
							S3 Average	3.377415881

**Table S5.** Breakdown of metagenomic data after each step from sequencing to assigning taxonomy.**Table S6**

<b>A.</b>				<b>B.</b>			
Samples	evaluate 10	evaluate 15	evaluate 20	Site	evaluate 10	evaluate 15	evaluate 20
S1.A.t0	15	8	3	S1	7.00	3.67	1.00
S1.E1	5	2	0	S2	38.00	17.00	3.33
S1.F1	1	1	0	S3	380.00	171.67	30.00
S2.A.t0	51	19	3	<b>C.</b>			
S2.E1	35	17	5	Site	evaluate 10	evaluate 15	evaluate 20
S2.F1	28	15	2	S1	7.21	3.79	1.73
S3.A.t0	299	137	28	S2	11.79	2.00	1.53
S3.E1	334	150	28	S3	111.37	49.22	3.46
S3.F1	507	228	34				

**Table S6A.** Number of potential *hgcA* genes found in unassembled metagenome data at various levels of stringency for each sampling site and treatment. **B.** Average number of potential *hgcA* genes found in the unassembled metagenome data at various levels of stringency (e-values) at each sampling site. **C.** Standard variance of potential *hgcA* genes found in the unassembled metagenome data at various levels of stringency at each sampling site.

**Table S7**

**A.**

Primer Type	Primer Name	Sequence	Length (bp)	Degeneracy (fold)	Target Gene	Percent Coverage	Fragment length (bp)	Average Detection Limit (copies)	Reference
Broad-Range	ORNL-HgcAB-uni-F	5'-AAYGTCTGGTGYGCNCGVGG-3'	20	48	hgcA/hgcB	64	880-1020	1 X 10 <sup>6</sup>	(Christensen et al. 2016) (Christensen et al. 2016)
	ORNL-HgcAB-uni-R	5'-CABGCNCCRCAYTCCATRCA-3'	20	96		76			
	hgcA_F	5'-GGNRTYAAAYRTNTGGTGYGC-3'	20	128	hgcA/hgcB	85	888-945	N/A	(Bae, Dierberg, and Ogram 2014) (Bae, Dierberg, and Ogram 2014)
	hgcB_R	5'-CADGCNCCRCAYTCVATRCA-3'	20	288					
	hgcA4F	5'-GGNRTYAAAYRTCTGGTGYGC-3'	20	128	hgcA	56	680	N/A	(Liu et al. 2014)
	hgcA4R	5'-CGCATYTCCTTYTYBACNCC-3'	20	96					
Clade-Specific	Delta-hgcA-F	5'-GCCAACTACAAGMTGASCTWC-3'	21	8	hgcA	21	107	2 X 10 <sup>5</sup>	(Christensen et al. 2016) (Christensen et al. 2016)
	Delta-hgcA-R	5'-CCSGCNGCRCACCAGACR-3'	20	32		57			
	Archaea-hgcA-F	5'-AAYTAYWCNCTSAGYTTYGAYGC-3'	23	512	hgcA	64	125	2 X 10 <sup>4</sup>	(Christensen et al. 2016) (Christensen et al. 2016)
	Archaea-hgcA-R	5'-TCDGTCCCRABGTSCCYTT-3'	20	72		45			
	Firm-hgcA-F	5'-TGGDCCGGTDARAGCWAARGATA-3'	23	72	hgcA	6	167	2 X 10 <sup>5</sup>	(Christensen et al. 2016) (Christensen et al. 2016)
	Firm-hgcA-R	5'-AAAAGAGHAYBCCAAAATCA-3'	21	18		25			

**B.**

Primer Type	Primer Name	Sequence	Length (bp)	Degeneracy (fold)	Target Gene	Percent Coverage	Fragment length (bp)	Average Detection Limit (copies)
Clade-Specific	Deltaproteo F1	5'-AAYRTVTGGTGYGCNG-3'	16	96	hgcA/hgcB	100	880-1020	N/A
	Deltaproteo F2	5'-AAYRTVTGGTGYGCWG-3'	16	48		25		
	DeltaProteo F3	5'-AAYRTVTGGTGYGCSSG-3'	16	48		75		
	Deltaproteo R1	5'-GCNCCRCAYTCVATRCA-3'	17	64		100		
	Firmicutes F1	5'-AAYGTHTGGTGYGCYGC-3'	17	24	hgcA	44	479-503	N/A
	Firmicutes F2	5'-AAYGTHTGGTGYGCRGC-3'	17	24		56		
	Firmicutes R1	5'-TTYGCNTGGAAARGGHTGG-3'	18	48		88		
	Archaea F1	5'-AAYTAYWCNCTSAGYTTYGA-3'	20	256	hgcA	64	94-100	N/A
	Archaea F2/R1	5'-AAYGTVTGGTGYGCYGC-3'	17	24		82		
	Archaea F3/R2	5'-AAYGTVTGGTGYGCRGC-3'	17	24	hgcA/hgcB	18	700-908	N/A
	Archaea R3	5'-CNCCRCAYTCCTYCA-3'	16	32		100		

**Table S7A.** Table of primer sets used to amplify *hgcA* and *hgcB* found in the literature. Primers are sorted into broad-range or clade-specific. Red represents degenerate bases; Y = C/T, R = A/G, M = A/C, S = C/G, W = A/T, V = A/C/G, B = C/G/T, D = A/G/T, H = A/C/T, N = A/T/C/G. **B.** Table of less degenerate and more specific primer sets used to amplify *hgcA* and *hgcB* modified from primers found in literature. Primers are sorted into clades. Red represents degenerate bases; Y = C/T, R = A/G, M = A/C, S = C/G, W = A/T, V = A/C/G, B = C/G/T, D = A/G/T, H = A/C/T, N = A/T/C/G.