

The Role of Arctic Permafrost Microbiomes in Mercury Mobilization

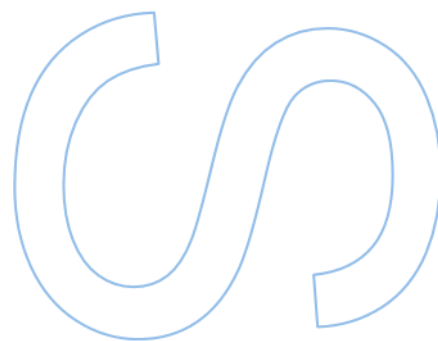
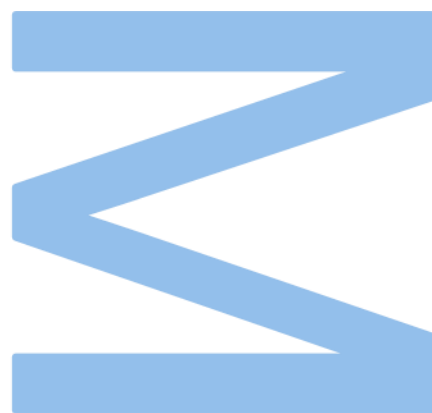
Ana Patrícia de Sousa Ribeiro

Master in Ecology and Environment

Department of Biology

Faculty of Sciences of the University of Porto

2024



The Role of Arctic Permafrost Microbiomes in Mercury Mobilization

Ana Patrícia de Sousa Ribeiro

Dissertation carried out as part of the
Master in Ecology and Environment
Department of Biology
2024

Supervisor

Joana Faria da Costa
Researcher at the Interdisciplinary Center for Marine and
Environmental Research (CIIMAR)

Co-supervisor

Catarina Maria Pinto Moura Pinto de Magalhães
Assistant Professor at the Faculty of Sciences of the University of
Porto and Researcher at the Interdisciplinary Center for Marine
and Environmental Research (CIIMAR)



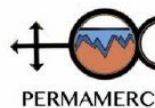
“All we have to decide is what to do with the time that is given us.”

— J.R.R. Tolkien, The Fellowship of the Ring

This project was presented at IJUP 2024 as an oral communication, with the following reference:

Ribeiro Ana; Costa, Joana; Gambardella, Nicola; Magalhães, Catarina, The Role of Arctic Permafrost Microbiomes in Mercury Mobilization, 17th edition of the Encontro de Investigação Jovem da Universidade do Porto (IJUP) 2024, 8-10 May 2024, Porto, Portugal.

This research is being supported by FCT, Fundação para a Ciência e Tecnologia) which supports the development of the PERMAMERC (PTDC/CTA-AMB/4744/2020) and N-MICROARCTIC (2022.02983.PTDC) projects through national funds.



Acknowledgements

Agradeço à minha orientadora, Doutora Joana Costa, em primeiro lugar pela sua simpatia e dedicação, em segundo lugar por toda a sua ajuda, disponibilidade, orientação e confiança durante este ano. Agradeço ao Nicola Gambardella por toda a sua orientação, pela sua paciência, disponibilidade e por me incentivar a aprender bioinformática. Agradeço à minha coorientadora, Doutora Catarina Magalhães pela oportunidade de realizar esta tese de mestrado e pelo seu apoio e incentivo. Agradeço ainda ao Professor João Canário e à sua equipa.

Um agradecimento especial...

Às da faculdade, Bia, Catarina e Mariana, pela vossa amizade, por estarem sempre lá para mim e fazerem todas as minhas longas viagens valerem a pena. Por todos os bons momentos e todas as aventuras ao longo destes 6 longos anos.

À minha amiga Bárbara, obrigada pela tua amizade, por todo o apoio e por me incentivares sempre a superar-me.

À minha prima Alícia, por estar sempre lá para mim, por me fazer sorrir nos piores momentos e aturar todas as minhas crises existenciais.

À maior surpresa que me trouxe este percurso, Rui, obrigada por me lembrares sempre que sou capaz, por me fazeres sempre querer aprender mais e ser melhor, por seres o meu melhor amigo nos piores e melhores momentos e sobretudo por todo o amor.

A maior gratidão aos meus pais, por tudo o que fazem por mim, por nunca duvidarem de quem sou, e por todo o carinho e amor que me dão. À minha avó-madrinha que está sempre lá para mim, obrigada por todo o amor e suporte ao longo da minha vida. Ao meu avô, que está sempre a olhar por mim lá de cima. À minha avó poeta, que sempre me incentivou a seguir os meus sonhos e ser mais do que posso imaginar, por todo o seu amor e carinho. Ao meu avô, por acreditar em mim e me ajudar em tudo o que precisar.

A todas as pessoas que direta ou indiretamente me apoiaram, um obrigada!

Sou eternamente grata pelas pessoas que tenho na minha vida, e por tudo o que fazem por mim e me ajudarem a chegar onde cheguei. *Não foi por acaso.*

Resumo

Os microrganismos desempenham um papel essencial nos ciclos biogeoquímicos da Terra, sendo por isso fundamental compreender a sua resposta às alterações climáticas. O aquecimento global ameaça a libertação de mercúrio (Hg) armazenado em ambientes de permafrost ao acelerar o seu degelo. Certos microrganismos que habitam nesses ambientes têm a capacidade de metilar o Hg em metilmercúrio (MMHg), um composto neurotóxico que se acumula e biomagnifica ao longo da cadeia alimentar, representando riscos significativos para a saúde de predadores de topo e humanos. Ao longo do tempo, vários estudos foram realizados com o intuito de compreender os mecanismos relacionados com o ciclo do Hg em ambientes naturais. Diversas descobertas importantes foram feitas, sobretudo a identificação de dois genes chave para a metilação ocorrer, *hgcA* e *hgcB*. Além disso, os mecanismos da desmetilação foram aprofundados com a identificação dos genes *merA* e *merB*, que facilitam a desmetilação microbiana. O nosso estudo teve como objetivo identificar as comunidades microbianas nos lagos de termocarso do Ártico Canadano e o seu papel no ciclo do Hg, de modo a obter informações sobre a relação entre o microbioma e o ambiente, bem como o impacto das estações nestes processos. Amostras de sedimentos foram recolhidas em dois lagos de termocarso durante o inverno e o verão de 2022, a várias profundidades. Após a extração de eDNA, foi realizada a análise metagenómica utilizando sequenciação com tecnologia *Illumina* e uma *pipeline* específica (*marky-coco* & *marky-coco-visualizer*) para identificar os genes de interesse. A análise posterior foi realizada no software R, integrando dados biológicos e ambientais, o que permitiu estimar os valores relativos à *cobertura* dos genes associados ao ciclo do Hg e avaliar o perfil taxonómico das comunidades procariotas locais. Os nossos resultados revelaram uma elevada diversidade de microrganismos capazes de metilação e desmetilação de Hg em ambos os lagos. Relativamente à metilação, observámos uma maior *coverage* do gene *hgcA* e uma maior diversidade taxonómica no inverno, em comparação com o verão, e postulámos um possível aumento da atividade de metilação durante o inverno. No verão, ocorre uma mudança na comunidade microbiana, resultando numa menor *coverage* do gene *hgcA*, simultaneamente com uma maior *coverage* dos genes *merAB* e menores concentrações de MMHg, sugerindo uma predominância da atividade de desmetilação nesta estação. Por fim, estes resultados destacam a complexidade das dinâmicas do ciclo do Hg. Num mundo onde as temperaturas globais continuam a aumentar, é essencial investigar como as alterações climáticas afetarão estes ecossistemas, antecipando e mitigando os riscos ambientais associados à libertação de Hg.

Palavras-chave: metilação, desmetilação, lagos de termocarso, metagenómica, alterações climáticas, ecologia microbiana, genes *hgcAB*, genes *merAB*, Canadá, bioinformática

Abstract

Microorganisms play an elementary role in the earth's biogeochemical cycles, therefore understanding their response to climate change is essential. Rising temperatures threaten the release of mercury (Hg) stored in permafrost environments by accelerating its thawing. Specific microorganisms inhabiting these habitats are known to have the ability to methylate Hg into monomethylmercury (MMHg), a neurotoxic compound that bioaccumulates and biomagnifies within the food web, posing significant health risks to top predators and humans. Over the years many studies were conducted in order to comprehend the mechanisms behind Hg pathways in natural environments. Many achievements were made, in particular, the identification of two key genes, *hgcA* and *hgcB*, that enable methylation. Besides, a better understanding of demethylation was aided by the identification of *merA* and *merB* genes within the *mer* operon allowing microbial demethylation. Our study aimed to identify microbial communities in Canadian Arctic thermokarst lakes and their involvement in Hg cycling, providing insights into the relation between microbiome and environment and how seasonal variations may influence these processes. Sediment samples were collected from two thermokarst lakes during winter and summer 2022, at various depths. Following eDNA extraction, metagenomics analysis was performed utilizing next-generation sequencing with Illumina technology and a specific pipeline (*marky-coco* & *marky-coco-visualizer*) to identify the genes of interest. Downstream analysis was performed in R Statistical Software by integrating biological and environmental datasets. The use of these tools allowed for the estimation of Hg-cycle associated gene coverage values and assessment of the local prokaryotic communities' taxonomic profile. Our findings revealed a high diversity of microorganisms capable of Hg methylation and demethylation in both lakes. Regarding methylation, we observed higher *hgcA* gene coverage and higher taxonomic diversity in winter, when compared to summer and postulated a possible increase in methylation activity during winter. In summer, a microbial community shift occurs, resulting in lower *hgcA* gene coverage alongside with higher *merAB* gene coverage and lower MMHg concentrations, suggesting a predominance of demethylation activity in this season. Overall, these results highlight the complexity of Hg-cycle dynamics and confirm the necessity to direct more studies into its understanding. In a world where global temperatures continue to increase it is essential to forecast how climate change will affect these ecosystems to anticipate and mitigate environmental risks related to Hg release.

Keywords: methylation, demethylation, thermokarst lakes, metagenomics, climate change, microbial ecology, *hgcAB* genes, *merAB* genes, Canada, bioinformatics

Table of contents

Acknowledgements.....	1
Resumo	2
Abstract	4
List of tables	8
List of figures	9
List of abbreviations	11
1. Introduction.....	1
1.1. Arctic ecosystems and global warming.....	1
1.2. Permafrost.....	1
1.3. Mercury	3
1.4. Microbial mediated processes in mercury cycle.....	4
1.5. Metagenomics	7
1.6. Objectives	8
2. Methods.....	9
2.1. Samples collection	9
2.2. eDNA processing	10
2.2.1. eDNA extraction and quantification.....	10
2.2.2. Sequencing	10
2.3. Bioinformatic pipelines	11
2.3.1. Kbase metagenomic analysis	11
2.3.2. Marky-coco.....	11
2.3.3. R libraries	14
3. Results	15
3.1. eDNA extraction and sequencing technology	15
3.1.1. Quality control of the sequenced reads.....	16
3.2. Microbial composition of SAS lakes.....	17
3.2.1. Taxonomic classification	17
3.3. Mercury cycle.....	19

3.3.1.	<i>hgcA</i> coverage	19
3.3.2.	<i>hgcB</i> coverage	21
3.3.3.	<i>merA</i> coverage	22
3.3.4.	<i>merB</i> coverage	23
3.4.	Environmental mercury data.....	25
3.4.1.	Total mercury	25
3.4.2.	Monomethylmercury	26
4.	Discussion	27
4.1.	Taxonomic classification.....	27
4.2.	Mercury Cycle	29
4.2.1.	<i>hgcA</i> and <i>hgcB</i> genes.....	29
4.2.2.	<i>merA</i> and <i>merB</i> genes.....	32
4.2.3.	Seasonal dynamics of mercury concentrations	34
4.2.4.	Contrasts between the thermokarst lakes.....	36
5.	Conclusion.....	37
	References	38
	Supplementary files	48

List of tables

Table 1. Identification of each sample according to depth, location, and extraction season.....	10
Table 2. eDNA quantification results, obtained with the Qubit method.....	15
Table 3. Distribution of <i>merA</i> and <i>merB</i> genes across microbial taxa, grouped by phylum.	32

List of figures

- Figure 1. *Permafrost layers.*** The image depicts the vertical profile of permafrost, illustrating the relationship between surface and subsurface layers. Permafrost is soil at or below 0 °C for at least two consecutive years and the active layer is the surface layer of soil above permafrost that thaws in summer and refreezes in winter. Depth of Zero Annual Amplitude (DZAA) is the point below the surface where the temperature remains constant throughout the year, regardless of seasonal changes. Beneath permafrost base, ground is permanently unfrozen. Adapted from Palangi & Abyaneh, (2017). 2
- Figure 2. *Mercury cycle schematic representation.*** The image illustrates the movement and transformation of Hg through the atmosphere, ocean, soil, and deep mineral reservoirs. Hg emissions from natural sources (volcanoes and biomass burning) and human activities (industrial emissions) enter the atmosphere, where Hg exists in elemental (Hg(0)) and oxidized (Hg(II)) forms. It deposits on land and oceans, where it can undergo further transformations, including methylation to methylmercury (MMHg). This accumulates in food webs, posing health risks through fish consumption. Additional sources include river flux and permafrost thawing, and eventually being buried in sediments as mercury sulphide (HgS). Adapted from: Physical processes in the mercury biogeochemical cycle, MacKenzieEJewell, Wikipedia, CC BY-SA 4.0. 3
- Figure 3. *Study area in the Canadian arctic, Quebec.*** Sasapimakwananisikw (SAS) valley, Quebec (left). Aerial view of the two study lakes, SAS1 and SAS2 (right). Source: Google maps **Error! Bookmark not defined.**
- Figure 4. *Per base sequence quality for winter sample PE22_1 (a) and summer sample PE22_27 (b).*** Winter sample reads quality is high showing scores above 30 percent from 1-150 bases with a slight decline near the final region. Summer sample shows a high quality of reads in the initial region (1-75) with an evident decline from the intermediate to the final region (75-150 bases) 16
- Figure 5. *Taxonomic classification of reads with Kaiju.*** Winter (a) and Summer (b) taxonomic classification graphics display the percentage of each phylum, assigned with Kaiju, according to different lakes (SAS1 and SAS2) and depths (0–5 cm, 5–10 cm, 10–15 cm). Phyla are presented in order from the highest to the lowest percentage. The analysed data excludes unclassified reads, which account for 56.45% to 60.63% of reads in winter samples and 67.09% to 72.28% in summer samples. Microorganisms for which taxonomy could not be assigned to the reads are labelled as “could_not_be_assigned.” “Others” represents microorganisms that were assigned but were below the 0.5% abundance threshold. 17

Figure 6. <i>hgcA</i> gene coverage across lakes, depths and seasons. Stacked bar plot (a) Winter and bar plot (b) Summer, each color represents a different identified taxonomy. Both plots display the taxonomic composition based on the normalized coverage of <i>hgcA</i> gene, across different soil depths (0-5cm, 5-10cm, and 10-15cm) at two sampling sites, SAS1 and SAS2.	19
Figure 7. <i>hgcB</i> gene coverage across lakes, depths and seasons. Stacked bar plot (a) Winter and bar plot (b) Summer, each color represents a different identified taxonomy. Both plots display the taxonomic composition based on the normalized coverage of <i>hgcB</i> gene, across different soil depths (0-5cm, 5-10cm, and 10-15cm) at two sampling sites, SAS1 and SAS2.	21
Figure 8. <i>merA</i> gene coverage across lakes, depths and seasons. The figure displays two stacked bar plots, bar plot (a) Winter and bar plot (b) Summer where each color represents a different microbial taxon. Both plots display the taxonomic composition based on the normalized coverage of <i>merA</i> gene, across different soil depths (0-5cm, 5-10cm, and 10-15cm) at two sampling sites, SAS1 and SAS2.	22
Figure 9. <i>merB</i> gene coverage across lakes, depths and seasons. Bar plot (a) Winter and bar plot (b) Summer, each color represents a different microbial taxon. Both plots display the taxonomic composition based on the normalized coverage of <i>merB</i> gene, across different soil depths (0–5 cm, 5–10 cm, and 10–15 cm) at two sampling sites, SAS1 and SAS2. Taxa found exclusively in a specific lake are labelled with the number of the lake where they were discovered (SAS1- 1 and SAS2- 2).....	23
Figure 10. <i>merB</i> gene coverage comparative analysis between seasons. Stacked bar plot showing winter (leftt) and summer (right) normalized coverage of <i>merB</i> gene across different sediment depths. The y-axis represents normalized coverage, indicating the relative abundance of each taxon represented by different colors. Taxa found exclusively in a season are labelled accordingly (Winter- W; Summer- S)	24
Figure 11. <i>Total mercury (THg) concentrations.</i> Winter (left) and summer (right) THg concentrations measured in (ng/g) at lake SAS1 (dark blue) and lake SAS2 (light blue). The bars represent the mean THg concentrations (ng/g), and the errors correspond to the standard deviation (SD) of the means.	25
Figure 12. <i>Monomethylmercury (MMHg) concentrations.</i> Winter (left) and summer (right) MMHg concentrations measured in (ng/g) at lake SAS1 (dark blue) and lake SAS2 (light blue). The bars represent the mean MMHg concentrations (ng/g), and the errors correspond to the standard deviation (SD) of the means.	26

List of abbreviations

°C - Degrees Celsius

IPCC – Intergovernmental Panel on Climate Change

NOAA – National Oceanic and Atmospheric Administration

OM – Organic Matter

O₂ – Oxygen

C – Carbon

CH₄ – Methane

CO₂ – Carbon dioxide

DZAA – Depth of Zero Annual Amplitude

Hg – Mercury

Hg(0) – Elemental Mercury

Hg(II) – Divalent mercury

MMHg – Monomethylmercury

CH₃ – Methyl group

Gg – Giga grams

SRB – Sulphate reducing bacteria

IRB – Iron reducing bacteria

HgcA – Corrinoid protein

HgcB – Ferredoxin-like protein

CFeSP – Corrinoid iron-sulfur proteins

DOM – Dissolved Organic Matter

mer – Mercuric

MerA – Mercuric reductase

MerB – Organomercury lysase

MerR – Mercuric resistance operon regulatory protein

DOC – Dissolved Organic Carbon

ROS – Reactive Oxygen Species

H₂S – Hydrogen sulphide

(CH₃)₂HgS – Dimethylmercury sulphide

HgS – Mercury sulphide

OD – Oxidative demethylation

DNA – Deoxyribonucleic acid

eDNA – Environmental Deoxyribonucleic acid

QC – Quality Control

SAS – Sasapimakwananisikw

IST – Instituto Superior Técnico

IGOT – Instituto de Geografia e Ordenamento do Território da Universidade de Lisboa

CIIMAR – Centro Interdisciplinar de Investigação Marinha e Ambiental

WS – Working solution

DNBs – Deoxyribonucleic acid nanoballs

Gbp – Giga base-pair

HG-MATE – Hg-cycling Microorganisms in Aquatic and Terrestrial Ecosystems

Prodigal – Prokaryotic Dynamic Programming Gene finding Algorithm

SAM – Sequence Alignment/Map

BAM – Binary Alignment/Map

AAs – Amino acids

HMM – Hidden Markov Model

BlastP – Protein Basic Local Alignment Search Tool

MUSCLE – Multiple Sequence Comparison by Log-Expectation

UniProt – Universal Protein Resource

RNA – Ribonucleic acid

rpoB – RNA polymerase beta subunit-encoding gene

THg – Total mercury

bp – base pairs

PVC Superphylum – Planctomycetota, Verrucomicrobiota, Chlamydiota Superphylum

NOM – Natural Organic Matter

(CH₃)₂Hg – Dimethylmercury

1. Introduction

1.1. Arctic ecosystems and global warming

In the past decades, global surface temperature has risen 1.1°C. This happened at an increased rate after 1970 due to the higher usage of greenhouse gases worldwide for energy and resources ([IPCC - Climate change report, 2023](#)). It is noticeable that climate is changing with the aggravation of extreme phenomena across the globe which are rapidly affecting earth's ecosystems and biodiversity. Arctic ecosystems are particularly vulnerable to the increasing temperatures causing glacier retreat and the thawing of permafrost, with studies showing Arctic warming rates four times the global average ([Rantanen *et al.*, 2022](#); [Blunden *et al.*, 2023](#)). According to the National Oceanic and Atmospheric Administration (NOAA) the ice cover in the Arctic is shrinking every year, and the ice that remains yearly it's becoming thinner and more frail ([Lindsey & Scott, 2022](#); [Meng *et al.*, 2023](#)).

Seasonal contrasts are evident worldwide, but the Northern hemisphere has a unique seasonal pattern, due to earth's tilt and orbit around the sun, characterised by long winters that last nearly 8 months denoted by long nights and nearly no daylight, alongside temperatures below 0°C. In contrast, summer season lasts about 4 months, characterized by mild temperatures, shorter nights and progressively longer days ([Hartmuth *et al.*, 2022](#)). Several studies have noted that during summer, microbial communities in open water systems can be highly influenced by environmental conditions such as sunlight exposure, higher input of organic matter (OM), nutrients, and microbial activity ([MacMillan *et al.*, 2015](#); [Vigneron *et al.*, 2019](#)). However, during winter, ice cover acts as a barrier to atmospheric inputs, leading to the depletion of oxygen (O₂) under the ice due to the lack of exchange with the atmosphere ([Vigneron *et al.*, 2019](#)). With warmer winters this balance may be at risk, threatening overall life forms as well as the biogeochemical cycles that take place in these ecosystems.

1.2. Permafrost

Permafrost is soil that has been at or below 0 °C for at least two years in a row, and the active layer is the top layer of soil above permafrost that thaws in summer and refreezes in winter (Figure 1) ([Schaefer *et al.*, 2020](#)). Permafrost covers approximately 25% of the earth's terrestrial surface being more widespread in higher latitudes ([Burn, 2023](#)). Its formation occurs through the sedimentation of OM and inorganic compounds in the active layer, year after year, when the soil remains frozen due to cold temperatures ceasing or significantly reducing microbial activity, consequently trapping these

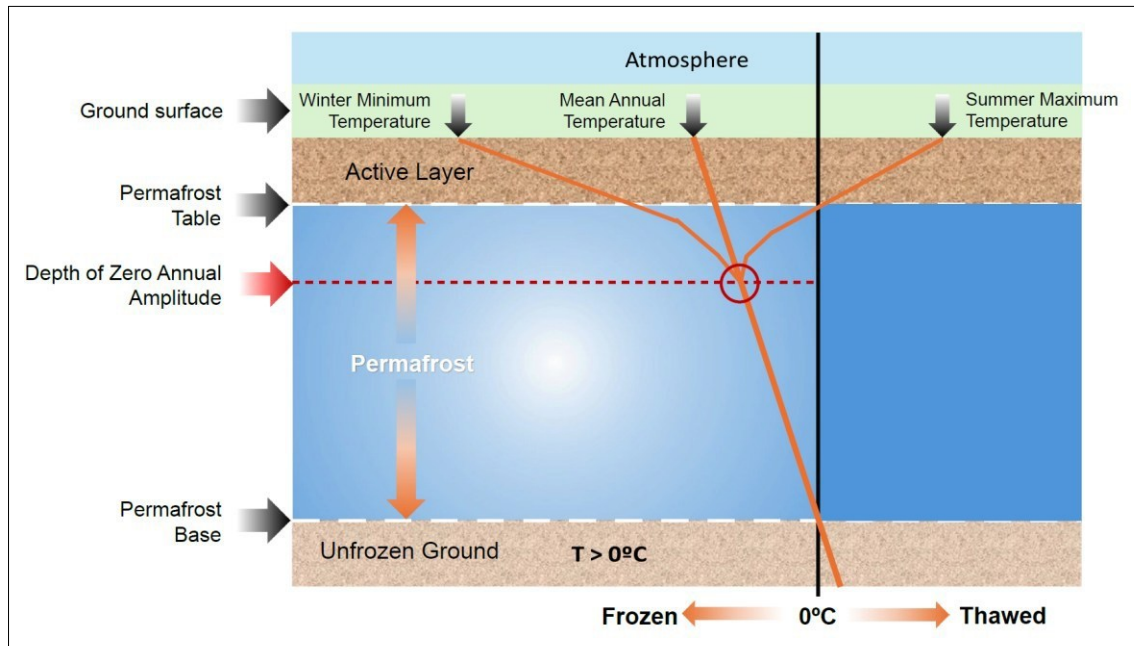


Figure 1. Permafrost layers. The image depicts the vertical profile of permafrost, illustrating the relationship between surface and subsurface layers. Permafrost is soil at or below 0 °C for at least two consecutive years and the active layer is the surface layer of soil above permafrost that thaws in summer and refreezes in winter. Depth of Zero Annual Amplitude (DZAA) is the point below the surface where the temperature remains constant throughout the year, regardless of seasonal changes. Beneath permafrost base, ground is permanently unfrozen. Adapted from [Palangi & Abyaneh, \(2017\)](#).

compounds at the bottom ([Schuster *et al.*, 2018](#); [Schaefer *et al.*, 2020](#); [Ottoni *et al.*, 2022](#)). There are various formation types of permafrost. In the particular case of this study, it is known that the thermokarst lakes were formed by the thawing of palsas ([Matveev *et al.*, 2016](#); [Folhas *et al.*, 2020](#)). Palsas are ice-rich low mounds with rounded shapes found in peatlands from the northern hemisphere and formed over long periods of time. For instance, during phases of cold temperatures, as a result of accumulation of water that freezes underground, ice expands pushing the ground upwards and creating a mound rich in OM, since their main composition is peat ([Folhas *et al.*, 2020](#); [Tarbier *et al.*, 2021](#)). When temperatures rise, thawing causes the collapse of these structures and forms thermokarst basins that fill with water from the melting ice ([Matveev *et al.*, 2016](#)). Moreover, several studies have shown that permafrost stores large pools of greenhouse gases such as methane (CH₄) and carbon dioxide (CO₂) and heavy metals ([Schaefer *et al.*, 2014](#); [Schuster *et al.*, 2018](#)). Therefore, the thawing of this frozen ground not only results in extended bodies of water, named thermokarst lakes, but also might result in the release of a myriad of unknown microbes and pollutants into the environment, posing a concerning threat to ecosystems and human health.

1.3. Mercury

Mercury (Hg) is naturally present in the environment in various forms, the most common of which are elemental Hg (Hg(0)), divalent Hg (Hg(II)), and monomethylmercury (MMHg). Its emission has been occurring naturally via volcanic eruptions, rock and soil erosion, geothermal activity, plant respiration, and biomass burning for millennia (Driscoll *et al.*, 2013). Hg(0) can be transformed into Hg(II) through oxidation processes causing Hg(II) to be available in large quantities in water bodies (Figure 2). Hg emissions are not only natural but also a result of anthropogenic activities. Coal-fired power stations, artisanal and small-scale mining, and industries, among others, have been significantly increasing Hg concentrations in the atmosphere surpassing natural emissions (Driscoll *et al.*, 2013, Pavithra *et al.*, 2023). A primary source of Hg(0) is the Hg present in lithospheric reservoirs that can be transferred to the atmosphere through volatilization and deposit on the surface through dry deposition. Hg(0) can reside in the atmosphere for up to a year and travel great distances. After its deposition, Hg(0) may be oxidised to Hg(II). This oxidised form can be re-emitted to the atmosphere after being reduced back to Hg(0) (Driscoll *et al.*, 2013). Hg(II) can also be converted into MMHg by the addition of a methyl-group (CH₃) in a microbial mediated process (Schaefer *et al.*, 2011). In this form, Hg presents a high toxicity due to its increased bioavailability and its long half-life. When incorporated by organisms, it remains in the tissues for a long time before being excreted or degraded.

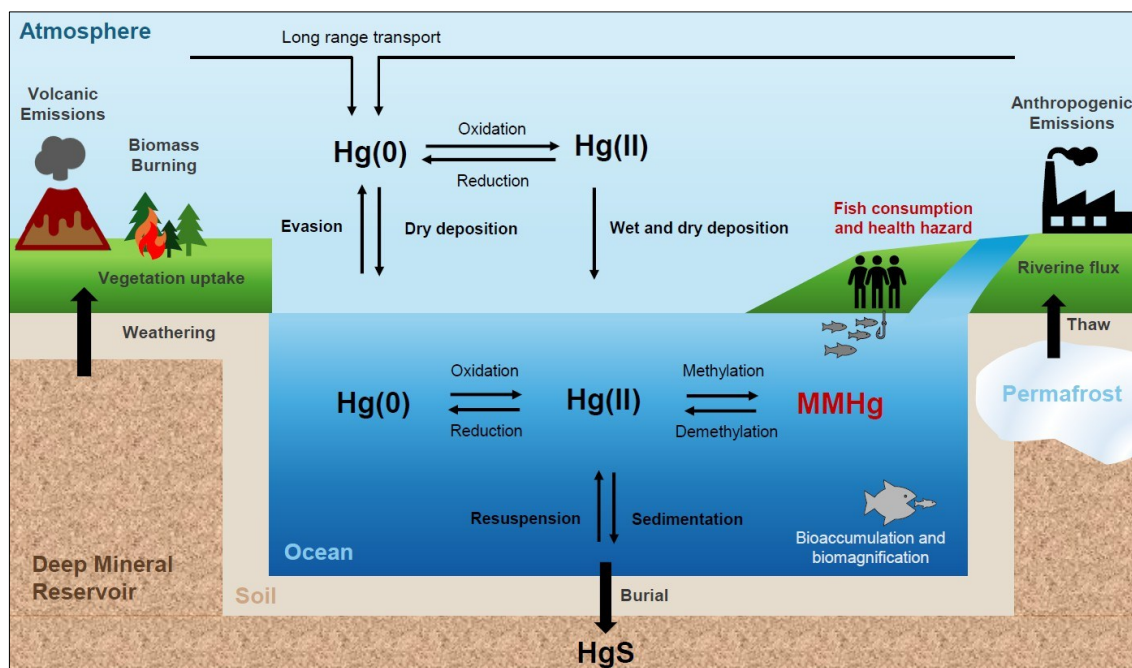


Figure 2. Mercury cycle schematic representation. The image illustrates the movement and transformation of Hg through the atmosphere, ocean, soil, and deep mineral reservoirs. Hg emissions from natural sources (volcanoes and biomass burning) and human activities (industrial emissions) enter the atmosphere, where Hg exists in elemental (Hg(0)) and oxidized (Hg(II)) forms. It deposits on land and oceans, where it can undergo further transformations, including methylation to methylmercury (MMHg). This accumulates in food webs, posing health risks through fish consumption. Additional sources include riverine flux and permafrost thawing, and eventually being buried in sediments as mercury sulphide (HgS). Adapted from: Physical processes in the mercury biogeochemical cycle, MacKenzieEJewell, Wikipedia, [CC BY-SA 4.0](#).

This leads to its bioaccumulation, and biomagnification along the trophic chains, reaching high concentrations in the top organisms, with harmful effects on humans and wildlife (Ha *et al.*, 2017). Hg was recognized as a chemical of global concern by the Minamata Convention on Mercury (MINAMATA CONVENTION ON MERCURY TEXT AND ANNEXES, 2019). It is therefore of great importance to assess what impact climate change will have on the release of Hg into the environment. Northern permafrost peatlands store, in the top 3 meters, an estimated ~560Gg to ~1700Gg of Hg of which ~800Gg are frozen into permafrost (Schaefer *et al.*, 2020; Tarbier *et al.*, 2021). Consequently, once thawed, permafrost conditions change, enabling microorganisms' activity as well as the input of external OM and nutrients, and exchange of contaminants, including Hg, to the surrounding environments (Schuster *et al.*, 2018). With warmer temperatures worldwide it is thought that permafrost will thaw creating the ideal conditions for Hg methylation (Schuster *et al.*, 2018; Vigneron *et al.*, 2019; Tarbier *et al.*, 2021).

Microorganisms play a crucial role in biogeochemical cycles by making elements available for other living beings (Meng *et al.*, 2022). They can be found ubiquitously on earth, even in very harsh environments. Their large environmental distribution can be explained by a wide genetic diversity which results in metabolic versatility allowing them to adapt to different environments (Barberán *et al.*, 2014). This metabolic versatility allows them to adjust to the ecosystem where they live due to the expression of distinct sets of genes activated in response to specific environmental conditions, allowing them to thrive in diverse circumstances ranging from hydrothermal vents to freezing sea water and from extreme acid to extreme basic conditions (VanInsberghe *et al.*, 2020). Microorganisms form communities whose diversity and abundance are controlled by resources and environmental conditions; furthermore, the complex interactions within these communities are crucial to the functioning and resilience of ecosystems. Microbial communities are, therefore, the main target of microbial ecology studies, being these essential to predict ecosystem responses to environmental changes.

1.4. Microbial mediated processes in mercury cycle

The classification of prokaryotic phyla has recently changed due to updates in the International Code of Nomenclature of Prokaryotes (ICNP) (Oren, 2024). In this thesis the previous classification system was used for identification of the local microbiome. Therefore, in order to facilitate results interpretation, and to align our study with the existing body of research on these microorganisms, all phyla will be addressed by the former classification. The first microorganisms described as Hg methylators were

sulphate-reducing bacteria (SRB), iron reducing bacteria (IRB), and methanogens. At present, excluding syntrophic and fermentative Firmicutes, all Hg methylators tested are strictly anaerobic bacteria that use sulphate, iron and CO₂ as terminal electron acceptors (Gilmour *et al.*, 2018; Capo *et al.*, 2023). A critical study to understand the microbial role in Hg-cycles was developed by Parks *et al.* (2013) where they found that the presence of the two genes *hgcA* and *hgcB*, in cultured *Desulfovibrio desulfuricans* ND132 and *Geobacter sulfurreducens* PCA, was essential for Hg methylation activity, confirming their potential as Hg methylators. They also found no significant impact on the bacteria's growth rate or extent of growth after deletion of *hgcA* and *hgcB* genes implying that, while the genes are necessary for Hg methylation, they are not essential for survival, seemingly not operating as a Hg detoxification system (Goñi-Urriza *et al.*, 2015; Capo *et al.*, 2022). The two model bacteria actively absorbed Hg(II) to the cell, added the CH₃ into the cytosol, and released MMHg to the environment (Parks *et al.*, 2013). In these organisms, the two genes found to be related to Hg methylation, encoded for a corrinoid protein (HgcA), thought to facilitate CH₃ transfer, and a downstream ferredoxin-like protein (HgcB) that allows for corrinoid reduction. Based on their function, they have been stated to be co-expressed (Parks *et al.*, 2013). This is reinforced by the fact that *hgcB* is almost always located downstream of *hgcA*, and the absence of one of the two genes blocks methylation, suggesting their co-evolution as a conserved operon (Lin *et al.*, 2023). These genes and their orthologs were later confirmed to be present in all sequenced known methylators and absent in non-methylators (Parks *et al.*, 2013; Christakis *et al.*, 2021; Capo *et al.*, 2023b). Furthermore, they suggest that thiolate ligands, which contain sulphur, strongly bind to Hg(II). Thiols affinity for Hg implies that HgcA may use this complex to stabilize Hg(II) during methylation which likely aids the transfer of a CH₃ to Hg. In the same study, a highly conserved motif (N(V/I)WCA(A/G)GK) was identified after analysing 52 HgcA orthologs. The presence of this sequence across many species suggests that it poses a crucial role in Hg methylation. Moreover, Parks *et al.* (2013) propose that this conserved region belongs to the cap helix of a corrinoid iron- sulphur protein (CFeSP), which is involved in transferring CH₃. Recent studies suggest that the methylation process evolved as an antimicrobial strategy that would favour microorganisms with the ability to metabolise Hg in a potentially resource limited environment, such as the early earth (Lin *et al.*, 2023). They suggested this process as an accidental enzymatic reaction by explaining that early microorganisms with the ability to transfer CH₃ during processes like methionine or acetyl coenzyme-A synthesis, over time, may have started transferring CH₃ to Hg(II) further exporting MMHg to the environment where it posed as toxin to surrounding microorganisms (Lin *et al.*, 2023).

Hg exposure alone does not induce the expression of *hgc* genes and does not confer Hg resistance (Capo *et al.*, 2022). Also, levels of *hgc* transcripts and levels of dissolved Hg(II) sulphide species are strong indicators of increased Hg(II) methylation and subsequently MMHg levels (Capo *et al.*, 2022). Hg methylation in thermokarst lakes is also influenced by an array of environmental factors like pH, temperature, reductive environment conditions, concentration of Hg and the presence of dissolved organic matter (DOM) (Bravo *et al.*, 2017). Hg is most effectively methylated under conditions with low sulphide concentrations and the presence of cysteine, a thiol-containing compound (Bravo *et al.*, 2017; Gilmour *et al.*, 2018).

Demethylation (Figure 2) is a chemical process where CH₃ is removed from an organic molecule (Hu *et al.*, 2020). This is done by utilising the mercuric (*mer*) resistance operon, which encodes for MerA (mercuric reductase) and MerB (organomercury lysase) proteins. However, the two are individually expressed; while MerA reduces Hg(II) to Hg(0), MerB breaks the MMHg bond, transforming MMHg to Hg(II) (Christakis *et al.*, 2021). The expression of these genes is controlled by Hg(II) exposure through a responsive regulatory protein, MerR. Hg(II) binds to MerR and induces transcription of the genes encoding MerA and MerB (Schaefer *et al.*, 2004). Therefore, transcription of *mer* genes is proportional to Hg concentrations in the environment (Nascimento *et al.*, 2003). Extreme concentrations of organic and inorganic Hg are thought to be the reason why microorganisms developed a demethylation reductive pathway that allows for detoxification, granting them an environmental advantage (Lin *et al.*, 2023). The origin of the *mer* operon is related to a thermophilic archaea origin and was later transferred to bacteria as an advantageous mechanism against Hg toxicity (Christakis *et al.*, 2021; Lin *et al.*, 2023). In particular, *merA* is thought to have appeared early in microbial evolution as a response to oxidative release of Hg(II) from sulphide minerals in aerobic environments allowing the reduction of Hg(II) to Hg(0). Further, *merB* is linked to a likely new addition to the *mer* operon, related to mesophilic environments where MMHg was toxic as a result of mesophilic Hg methylators activity, this is also supported by the fact that many *merB*-only microorganisms have been described and may possess other mechanisms for Hg(II) reduction such as thiols and orthologs that bind to Hg(II) and other redox reactions (Christakis *et al.*, 2021). Degradation of MMHg can also be mediated through abiotic photocatalysis, interaction with dissolved organic carbon (DOC), and DOM (Klapstein *et al.*, 2018; Christakis *et al.*, 2021). Barkay & Gu (2022) highlight different biotic and abiotic demethylation pathways. Biomediated pathways include *mer*-mediated reductive demethylation and *mer*-independent oxidative demethylation (OD). The former widely occurs in aerobic sedimentary conditions and at high levels of Hg. The

latter often involves methanogens and anaerobic SRB in anaerobic environments (Barkay & Gu., 2022). Abiotic pathways include DOM facilitated MMHg breakdown through direct and indirect photochemical reactions (Klapstein *et al.*, 2018). DOM absorbs ultraviolet radiation and transfers energy to MMHg breaking the C-Hg bond or generates reactive oxygen species (ROS) that interact with MMHg (Klapstein *et al.*, 2018; Bravo *et al.*, 2018). Further, dark demethylation enables the breakdown of MMHg in low light environments like deep waters or sediments. This process occurs through abiotic reactions, often involving DOM, sulphides, or minerals. In particular, hydrogen sulphide (H₂S), produced in anoxic conditions, reacts with MMHg to form compounds like dimethylmercury sulphide ((CH₃)₂HgS), which may eventually degrade into mercury sulphide (HgS) and CH₄ (Barkay & Gu., 2022).

1.5. Metagenomics

Historically, the study of microorganisms evolved around pure laboratory cultures which limited the vision of microbiologists on what they can actually do when in a structured community. When in the environment, these organisms live in very intricate, vast and changing habitats, thus for thousands of years they have been creating a distinct adaptability to environmental change (Committee on Metagenomics, 2007). However, this represents a challenge when replicating these complex environments in laboratory settings. Only recently, facilitated by technological progress, new solutions have emerged. Metagenomics focuses on genomic analysis applied to microbe communities, therefore diminishing the reliance on pure cultures. Shotgun-metagenomic sequencing, is an innovative specialised technology that focuses on sequencing the entire environmental DNA (eDNA) of a sample. First, samples are collected from the chosen habitats; subsequently, DNA is extracted and fragmented into smaller pieces, which will then be sequenced using high-throughput sequencing technologies. This will generate a vast volume of raw sequence data. Hence, at the bioinformatic analysis phase specific pipelines are employed which consist of several steps. Quality control (QC) is essential to remove errors and artefacts that may have been added during the sequencing. After this, trimming is done in order to remove the base pairs (bp) of less quality that may interfere later on the analysis. Furthermore, functional annotation assigns possible functions to genes, providing insights into the metabolic potential and ecological roles of the microbial community. At last, all the results are analysed for a better understanding of the metabolic potential of the community, its diversity, and its composition (Thomas *et al.*, 2012; Navgire *et al.*, 2022).

Particularly in this work, a specialised pipeline was used to recover the genes associated with the methylation and demethylation pathways of Hg and verify their presence and potential functions in permafrost soils. This pipeline was created by [Capo et al \(2023a\)](#) and is based on a "de novo" single assembly, which assembles each metagenome individually. This allowed for easy and accurate detection and characterization of *hgc* genes from environmental samples. This pipeline relies on a database where *hgc* and *mer* genes from isolated, single-cell and metagenome-reconstructed genomes are compiled ([Christakis et al., 2021](#); [Capo et al., 2023a](#)).

1.6. Objectives

This master thesis, part of the PERMAMERC project, aims to contribute to the understanding of the microbial involvement in Hg-cycle in Canadian Arctic permafrost thaw systems. The investigation specifically involves the characterization of Arctic permafrost microbiomes, the assessment of their involvement in Hg methylation and demethylation pathways, the relationship between microbiome and environment, and the evaluation of seasonal variation in the communities.

2. Methods

2.1. Samples collection

The study sites were located in the Sasapimakwananisikw (SAS) valley, in Quebec, at the subarctic Canada, near Whapmagoostui-Kuujuarapik village, a discontinuous permafrost area (Figure 3). Samples were collected by Dr. João Canário's team, affiliated with the Centro de Química Estrutural, Instituto Superior Técnico (IST), (Lisbon). Soil samples were collected from two lakes, identified as SAS1 and SAS2, in winter (March) and summer (August) of 2022. SAS1A (55°13'07.7 N, 77°42'28.4 W) is located southern of the SAS river while SAS2A (55°13'35.6 N, 77°41'49.1 W) is located north. During winter these lakes are covered by a moderate snow layer that melts during summer months and both the lakes are near palsas. Lake SAS1 has a large surface area (1661m²) and is shallow (1.3m max. depth), while lake SAS2 has a smaller surface area (104m²) but is of greater depth (2.5m max. depth) (IGOT, ULisboa). Both lakes present a surrounding dense population of macrophytes, specifically *Carex aquatilis* during warmer months (Folhas *et al.*, 2020).



Figure 3. Study area in the Canadian arctic, Quebec. Aerial view of the study area showing the locations of the study lakes, SAS1 and SAS2, in the Sasapimakwananisikw (SAS) Valley, Quebec, Canadian Arctic. The inset (top left) provides the regional context of the SAS Valley. Scale bar: 500 meters. Source: Google Maps.

Sediment samples were collected at the bottom of the lakes at three different depths (0-15 cm range) as described in Table 1. The samples were stored in individual sterile bags, accurately labelled, and preserved at -80°C until analysis. These methods were performed equally in both seasons.

Table 1. Identification of each sample according to depth, location, and extraction season.

Season	Sample ID	Depth	Lake	Latitude	Longitude
Winter	PE22_1	0-5 cm	SAS1	55°13'07.7 N	77°42'28.4 W
	PE22_2	5-10 cm			
	PE22_3	10-15 cm			
	PE22_9	0-5 cm	SAS2	55°13'35.6 N	77°41'49.1 W
	PE22_10	5-10 cm			
	PE22_11	10-15 cm			
Summer	PE22_27	0-5 cm	SAS1	55°13'07.7 N	77°42'28.4 W
	PE22_29	5-10 cm			
	PE22_31	10-15 cm			
	PE22_39	0-5 cm	SAS2	55°13'35.6 N	77°41'49.1 W
	PE22_41	5-10 cm			
	PE22_43	10-15 cm			

2.2. eDNA processing

2.2.1. eDNA extraction and quantification

eDNA extraction and quantification were performed at the Interdisciplinary Center of Marine Research (CIIMAR). DNeasy® PowerSoil® Pro Kit, was used for the sample's DNA extraction under sterile conditions following manufacturer's instructions.

To determine if the eDNA concentrations were sufficient for sequencing, we used a Qubit® fluorometer to measure the eDNA concentration of each sample, always following the producer's instructions.

2.2.2. Sequencing

Samples were sequenced by BGI company (Shenzhen, China) using DBNseq™ technology. This technology uses peculiar DNA nanoballs (DNBs) generated by rolling circle amplification of the DNA fragments that, placed onto a patterned array chip, are sequenced. Laser excitation and imaging capture the incorporation of fluorescently labelled nucleotides whose signal is converted into DNA sequence data. This technology is distinguished by its low amplification errors, high accuracy, and minimal duplication rates, making it highly effective for applications in whole-genome and whole-exome

sequencing (Kim *et al.*, 2021). The sequencing outputs, obtained from our samples, had an average of 71,966,315 sequencing reads, each comprising 150 nucleotides. This resulted in an average coverage of 10,7Gbp per sample.

2.3. Bioinformatic pipelines

2.3.1. Kbase metagenomic analysis

Initial metagenomics analysis was conducted using Kbase (Arkin *et al.*, 2018), an open software and data platform. The first step consisted of assessing the quality of the sequences using FastQC 0.12.1 (Andrews, 2010), a tool that performs QC tests on raw sequences produced by high-throughput sequencing pipelines. Next, Trimmomatic 0.36 (Bolger *et al.*, 2014), a tool for removing possible adapters and the lowest quality parts of the reads, was used with the default parameters:

Adapter clipping options: Enabled (Seed mismatches = 2; Palindrome clip threshold = 30; Simple clip threshold = 10); *Sliding window options:* (Sliding window size = 4; Sliding window minimum quality = 15); *Advanced parameters:* (Post tail crop length = null; Head crop length = 0; Leading minimum quality = 3; Trailing minimum quality = 3; Minimum read length = 36).

Afterwards, another FastQC was performed to evaluate if the trimming resulted in satisfactory quality trimmed reads.

Taxonomic classification was performed using Kaiju 1.9.0. (Menzel & Krog, 2016) in the trimmed paired end reads. Kaiju reports were subsequently analysed and filtered to obtain the necessary data to generate taxonomic profile plots displaying the taxonomic structure of microbial communities in the analysed samples, in different seasons, using R Statistical Software (v4.1.2; R Core Team 2021).

2.3.2. Marky-coco

Identification of Hg methylation genes requires *ad-hoc* reference databases, such as the one implemented in the pipeline developed by Capo *et al* (2023a) which detects and identifies *hgcA* and *hgcB* genes from raw FastQ files. This pipeline was created into a collaborative project which is paired to the *hgcAB* gene catalogue Hg-MATE (Hg-cycling Microorganisms in Aquatic and Terrestrial Ecosystems) (Gionfriddo *et al.*, 2021) created by the same team. This pipeline was used to investigate the presence of Hg methylation and demethylation genes in the local microbial communities.

This process initiated with the use of Fastp, an ultra-fast tool for quality control and data filtering, to trim and clean the raw data (Chen *et al.*, 2018). Posteriorly a *de*

novo assembly was performed using megahit 1.1.2, a tool used to assemble large and complex metagenomic data (Li *et al.*, 2015), followed by annotation of the contigs using Prodigal 2.6.3 (Prokaryotic Dynamic Programming Gene finding Algorithm) which predicts the location of protein-coding genes within the genome (Hyatt *et al.*, 2010).

Next, Bowtie 2 was used to align the reads to the assembled DNA sequences (contigs), providing insights into how these reads match the assembled sequence's structure (Langmead & Salzberg, 2012). Samtools 1.9 was used to transform the Bowtie2 output SAM (Sequence Alignment/Map) format into BAM (Binary Alignment/Map) since the latter are more efficient for storage and processing (Danecek *et al.*, 2021). FeatureCounts was used to estimate read counts, in combination with the BAM files and the Prodigal output, thereby determining the number of reads that were aligned with each predicted gene (Liao *et al.*, 2014).

Afterwards, the amino acids (AAs) sequences obtained from each assembly were retrieved and stored in a FASTA file format. The Hidden Markov Model (HMM) profiles obtained from Hg-MATE.db.v1 were then applied to these AAs sequences using the *hmmsearch* software from Hmmer 3.2.1, a bioinformatics software package that employs HMM for sequence analysis (Potter *et al.*, 2018). AAs sequences identified as *hgcA* gene homologs were then aligned to an existing alignment of *hgcA* sequences using *hmmalign* software from Hmmer 3.2.1. These aligned sequences were posteriorly mapped to a phylogenetic tree specifically constructed for *hgcA*. Finally, the sequences were classified using specialized functions (*rppr* and *guppy_classify*) from the program *pplacer*, aiding in categorising sequences based on their relationships within the phylogenetic tree (Matsen *et al.*, 2010; Gionfriddo *et al.*, 2020).

Ultimately, custom bash scripts, designed to analyse sequencing data and extract relevant information from the identified homologous genes, were used to obtain coverage values, taxonomy information, and AAs sequences.

The output data was evaluated by applying a high stringency cutoff which involved selecting the true *hgcA* genes by the motifs NVWCAAGK, NVWCASGK, NVWCAGGK, NIWCAAGK, NIWCAGGK or NVWCSAGK. True *hgcB* genes were identified by filtering the motifs CMECGA and CIECGA and by being co-located with true *hgcA*. These motifs serve as markers for functional *hgcA* genes involved in Hg methylation. This approach allows us to distinguish between functional *hgcA* and *hgcA*-like genes with unknown Hg methylation capability. Similarly, true *hgcB* must have motifs CMECGA or CIECGA and be collocated with true *hgcA*, while *hgcB*-like genes do not necessarily have this collocation (Capo *et al.*, 2023a).

Furthermore, fused *hgcAB* genes, although they contain motifs from both *hgcA* and *hgcB*, lack functional capability for Hg methylation. Fused HgcAB proteins combine motifs from both HgcA and HgcB but lack functional capability for Hg methylation, while true collocated *hgcA* and *hgcB* possess separate functionalities and are capable of methylating Hg (Capo *et al.*, 2023a; Lin *et al.*, 2023).

Functional *merA* and *merB* genes are known to contain specific AAs at particular positions, which are essential for their function allowing for the identification of true *merA* and *merB* (Boyd & Barkay, 2012). For instance, *merA* homologs were screened in a study by Christakis *et al* (2021), across different species such as *Bacillus cereus* RC607 and found to have conserved cysteine pairs and tyrosine/phenylalanine at specific positions which are fundamental for protein's redox activity and function. Similarly, *merB* homologs were identified in microorganisms such as *Escherichia coli* (plasmid R831b), *Halobacterium* sp. DL1, and *Candidatus Nitrososphaera evergladensis* SR1 which exhibited conserved AAs including cysteine and aspartic acid at given positions that are essential for protein's role in Hg resistance (Christakis *et al.*, 2021).

True *merA* and *merB* were identified by the presence of these conserved AAs according to a script created by ECOBIOTEC colleague Nicola Gambardella. This script is based on the available literature on the mer operon composition. In order to perform this identification, the putative *merA* and *merB* sequences were compared against reference sequences from a reference database created by Christakis *et al* (2021). Separate databases were created for each putative *merA* and *merB* sequence in order to compare with the reference database. BLASTp was then used to align the query sequences with the reference database, and sequences with similarity greater than 30% were retained, as a higher protein similarity suggests higher functional equivalence (Abo-Elkhier *et al.*, 2019). The next step involved precise alignment using MULTiple Sequence Comparison by Log-Expectation (MUSCLE), which compared the query sequences to the reference sequences containing critical AAs (Edgar, 2004). A score table was created based on the alignment results for each query sequence.

After identifying true *merA* and *merB* genes, taxonomic assignment of the sequences was performed using the BLASTp tool against the UniProt database. Specifically, *merA* sequences were compared to 5111 sequences, and *merB* sequences were compared to 442 sequences (Consortium, T. U., 2022). Given that sequence similarities observed were typically less than 40%, we may consider that our sequences may be shortened, or that they may represent new sequences not yet categorized in databases.

Nonetheless, the taxonomic assignment was sufficient to evaluate the community structure for this study.

Coverage is the number of reads that uniquely align to a specific part of the genome. Ideally, these reads are evenly distributed throughout the genome, resulting in a uniform coverage, which indicates that each region of the genome was sequenced at approximately the same level (Waston *et al.*, 2017). A higher coverage usually indicates greater confidence in the data, while a lower coverage indicates that there are fewer reads aligning to the genome, leading to uncertainty about the accuracy of the sequence. To account for variations in sequencing depth, it is important to normalize coverage across different samples to ensure that comparisons between samples are valid. The coverage of Hg-cycle genes was determined by dividing the number of reads mapping to the contig by the length of the gene. This value was normalized by dividing the coverage value by the average coverage of RNA polymerase beta subunit-encoding gene from bacteria (*rpoBb*) and archaea (*rpoBa*), genes that are expected to be present in single copy in each organism (Capo *et al.*, 2023a). *rpoB* is a highly conserved gene within the bacterial genome (Khamis *et al.*, 2004). *hgcAB* and *merAB* normalised coverage values will be addressed as *hgcAB* and *merAB* coverage.

2.3.3. R libraries

R Statistical Software was used for data analysis of Kaiju and marky-coco pipeline outputs. Several R packages were used in RStudio to facilitate data visualization and manipulation. Ggplot2 is a widely used package in the R environment, and it was used to visually represent our data by allowing the construction of plots (Wickham, 2016). Complementing ggplot2, dplyr package was used for efficient data processing using functions such as filtering rows, selecting columns, mutating (creating new columns), summarizing data, and arranging data (Wickham, 2023). Furthermore, other auxiliary packages were used in our workflow, namely: purrr, readxl, forcats, RColorBrewer, GGally, patchwork, and cowplot. These packages served a variety of purposes such as functional programming, Excel data importing, factor manipulation, colour palettes, pairwise plots, and arranging multiple plots, respectively (Wickham, 2023 ; Neuwirth, 2022, Schloerke *et al.*, 2024; Wilke, 2024).

3. Results

3.1. eDNA extraction and sequencing technology

Table 2. eDNA quantification results, obtained with the Qubit method.

Season	Lake	Sample	Depth	Soil (mg)	DNA (ng/ μ L)	DNA/Soil ratio
Winter	SAS1	PE22_1	0-5cm	275	42,8	0,16
		PE22_2	5-10cm	330	59,0	0,18
		PE22_3	10-15cm	326	47,7	0,15
	SAS2	PE22_9	0-5cm	360	107	0,30
		PE22_10	5-10cm	362	51	0,14
		PE22_11	10-15cm	312	58	0,19
Summer	SAS1	PE22_27	0-5cm	393	48,5	0,12
		PE22_29	5-10cm	367	10,9	0,03
		PE22_31	10-15cm	385	8,6	0,02
	SAS2	PE22_39	0-5cm	363	15	0,04
		PE22_41	5-10cm	453	9,32	0,02
		PE22_43	10-15cm	372	9,07	0,02

High concentrations of eDNA (ng/ μ L) were obtained from samples collected during winter and summer campaigns, as seen in Table 2. In total 12 samples were analysed. Assuming similar extraction efficiencies between samples, and in order to compare eDNA concentration between samples, a normalization was made by dividing the obtained eDNA concentrations by the amount of soil (mg) used for the extraction process (DNA/Soil ratio). Winter samples highest DNA/soil ratio (0.30) was observed in SAS2 at the 0-5cm depth. Additionally, the highest DNA/soil ratio in summer (0.12) was in SAS1 at the 0-5cm depth. DNA/soil ratio were consistently lower in the summer samples compared to winter. Furthermore, surface layers (0-5cm) consistently hold higher DNA/soil ratios relatively to deeper layers (10-15cm) across both seasons.

3.1.1. Quality control of the sequenced reads

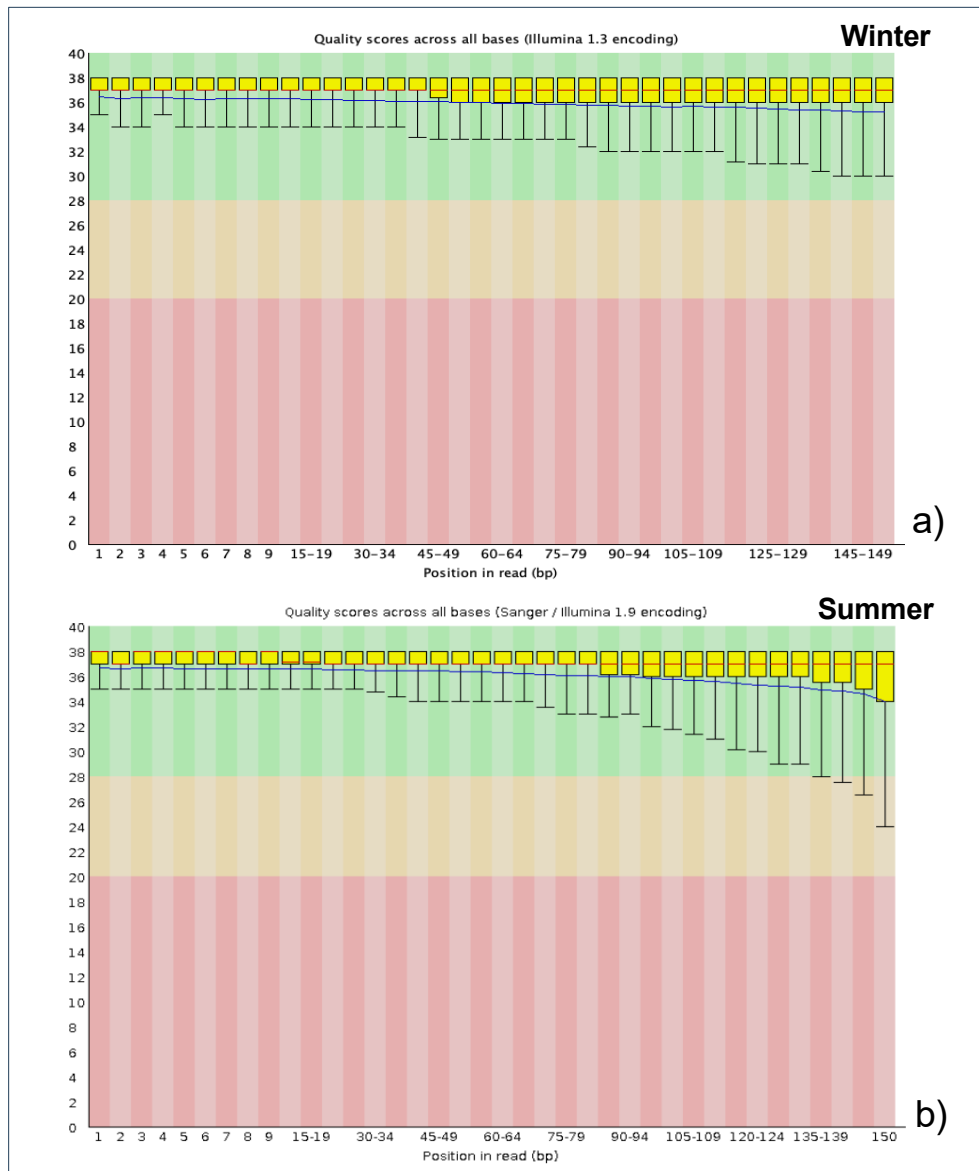


Figure 4. Per base sequence quality for winter sample PE22_1 (a) and summer sample PE22_27 (b). Winter sample reads quality is high showing scores above 30 percent from 1-150 bases with a slight decline near the final region. Summer sample shows a high quality of reads in the initial region (1-75) with an evident decline from the intermediate to the final region (75-150 bases).

The Phred scale is a scoring system that accesses the quality of DNA sequences by assigning a quality score (Q) to each base, reflecting the likelihood of an error in reading that base. This scale is logarithmic therefore each increase in the score represents a tenfold decrease in the probability of an error. High scores (above 30) indicate high-quality sequencing, confirming the reliability of the results. The sequencing reads obtained in our analysis, in both winter (Figure 4a) and summer (Figure 4b), demonstrated good results with all other samples following this trend (see supplementary files, S1-S12). Overall, we obtained a mean quality score of 35 for all samples, as assessed on the Phred scale confirming the effectiveness of the performed sequencing.

3.2. Microbial composition of SAS lakes

3.2.1. Taxonomic classification

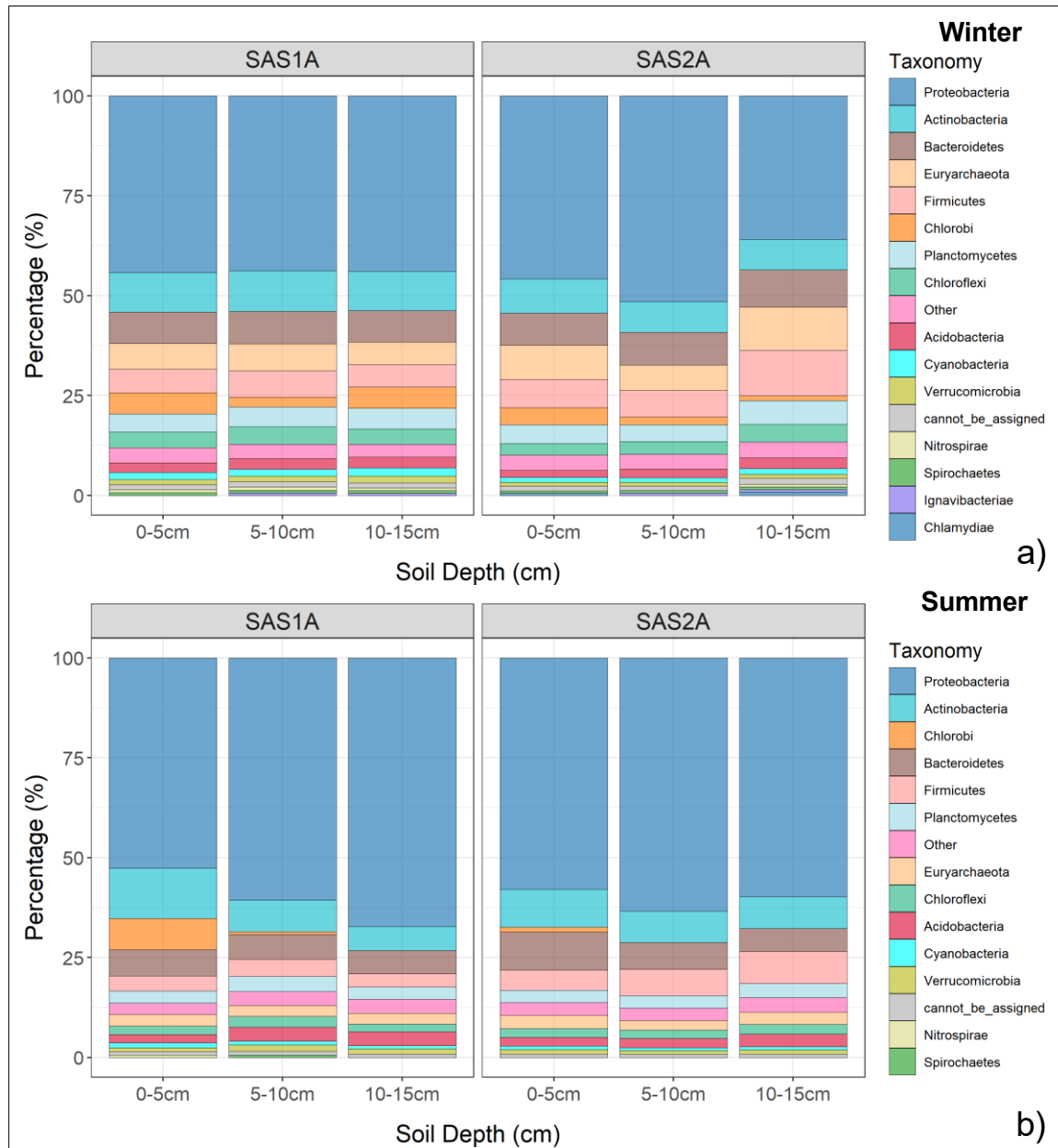


Figure 5. Taxonomic classification of reads with Kaiju. Winter (a) and Summer (b) taxonomic classification graphics display the percentage of each phylum, assigned with Kaiju, according to different lakes (SAS1 and SAS2) and depths (0–5 cm, 5–10 cm, 10–15 cm). Phyla are presented in order from the highest to the lowest percentage. The analysed data excludes unclassified reads, which account for 56.45% to 60.63% of reads in winter samples and 67.09% to 72.28% in summer samples. Microorganisms for which taxonomy could not be assigned to the reads are labelled as “could_not_be_assigned.” “Others” represents microorganisms that were assigned but were below the 0.5% abundance threshold.

Figure 5a and 5b displays Kaiju taxonomic classification for both winter and summer samples, respectively, at the phylum level. The analysed data excludes the unclassified reads which represent 56.45% to 60.63% of reads for winter samples and 67.09% to 72.28% of reads for summer samples.

Taxonomic results (Figure 5) showed that both lakes had a similar microbial composition, although different proportions of some of the phyla were observed. During winter (Figure 5a), Proteobacteria was the most abundant phylum, representing $44.23 \pm 4.98\%$ of the total identified microorganisms. Other abundant phyla included Actinobacteria ($8.96 \pm 1.17\%$), Bacteroidetes ($8.21 \pm 0.55\%$), Euryarchaeota ($7.40 \pm 1.90\%$), and Firmicutes ($7.22 \pm 2.14\%$). Phyla with lower relative abundance included Planctomycetes ($4.93 \pm 0.60\%$), Chloroflexi ($3.76 \pm 0.61\%$), Chlorobi ($3.43 \pm 1.76\%$), Acidobacteria ($2.43 \pm 0.41\%$), Cyanobacteria ($1.53 \pm 0.31\%$), and Verrucomicrobia ($1.23 \pm 0.33\%$).

In summer (Figure 5b), there was an increase in the relative abundance of Proteobacteria (+16.03%), reaching $60.26 \pm 4.93\%$, compared to winter. In contrast, Euryarchaeota showed a decreasing trend, falling (-4.57%) to $2.83 \pm 0.32\%$, and similar decreasing trends were observed for Firmicutes ($5.19 \pm 1.82\%$), Chlorobi ($1.75 \pm 3.01\%$), Planctomycetes ($3.25 \pm 0.32\%$), Chloroflexi ($2.25 \pm 0.29\%$) and Bacteroidetes ($6.75 \pm 1.43\%$). The relative abundance of Actinobacteria ($8.64 \pm 2.22\%$), Cyanobacteria ($0.88 \pm 0.20\%$) and Verrucomicrobia ($1.16 \pm 0.25\%$) slightly decreased in summer. However, Acidobacteria exhibited a slight increase to $2.77 \pm 0.67\%$ in summer.

3.3. Mercury cycle

3.3.1. *hgcA* coverage

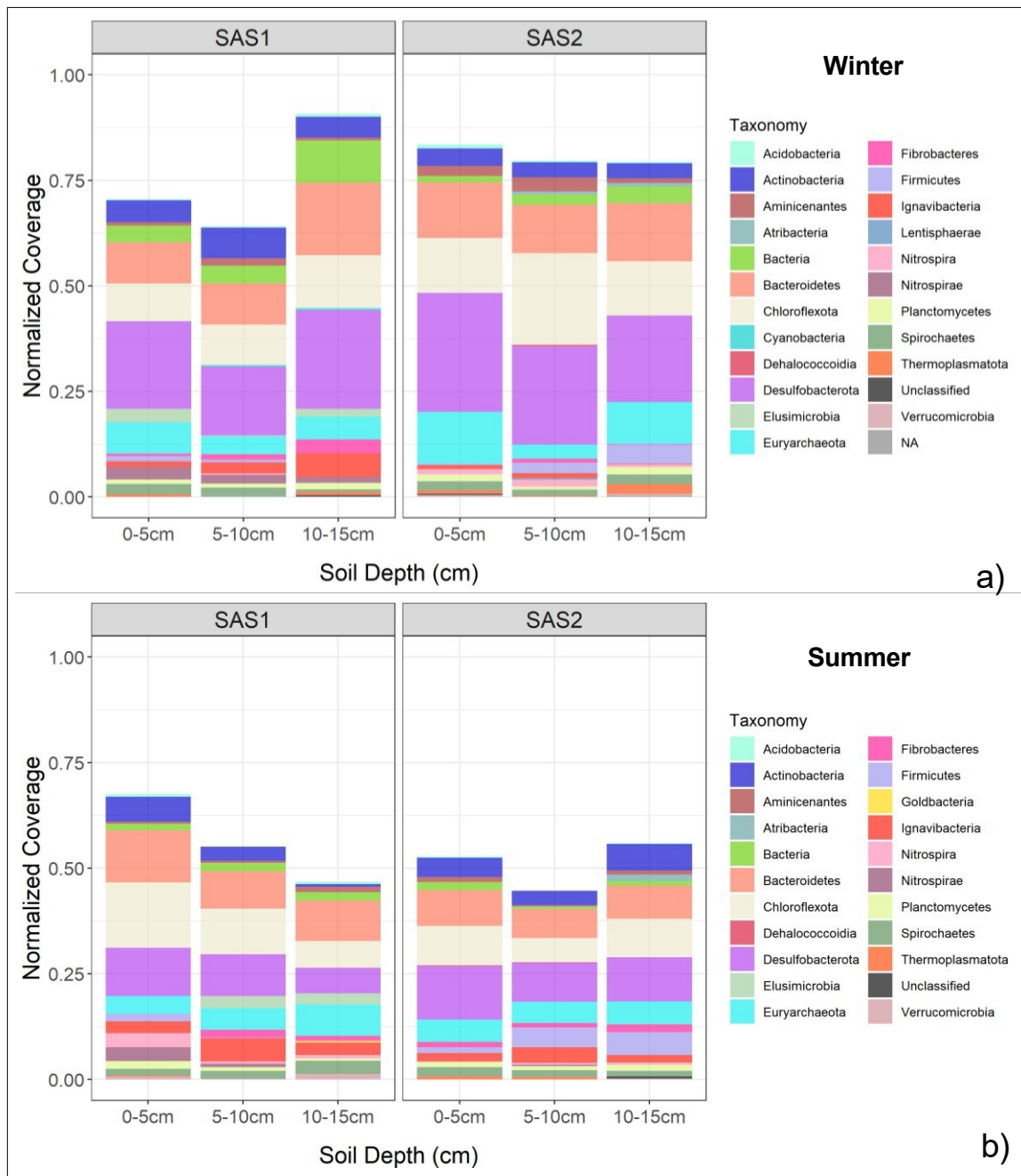


Figure 6. *hgcA* gene coverage across lakes, depths and seasons. Stacked bar plot (a) Winter and bar plot (b) Summer, each color represents a different identified taxonomy. Both plots display the taxonomic composition based on the normalized coverage of *hgcA* gene, across different soil depths (0-5cm, 5-10cm, and 10-15cm) at two sampling sites, SAS1 and SAS2.

Results showing the coverage levels of *hgcA*+ microorganisms, displayed in Figure 6, indicate that *hgcA*+ microorganisms were found in both lakes, at all analysed depths and throughout both seasons, winter and summer. The taxa that displayed overall higher *hgcA* gene coverages were Desulfobacterota, Bacteroidetes and Chloroflexota and Euryarchaeota.

In both seasons *hgcA* genes were present in the same phyla, differing in coverage values. However, some phyla were found to be unique in each of the lakes. In winter (Figure 6a), Cyanobacteria, Elusimicrobia and Nitrospirae *hgcA*⁺ genes were present only in lake SAS1 while Atribacteria, Dehalococcoidia, Lentisphaerae, and Verrucomicrobia *hgcA*⁺ genes were present exclusively in lake SAS2. Furthermore, in summer (Figure 6b), Elusimicrobia, Firmicutes and Goldbacteria *hgcA*⁺ were present in exclusively in lake SAS1 while Dehalococcoidia and Atribacteria *hgcA*⁺ only appeared in lake SAS2.

When comparing seasons (Figure 6a and 6b), it is noticeable that winter presented overall higher *hgcA* gene coverage values than summer. For instance, when considering the total coverage of the *hgcA* gene across all lakes in a given season, winter showed a coverage value of 4.68, while in summer, it remained at 3.23.

Seasonal shifts between the same lake also occur. For instance, *hgcA*⁺ Cyanobacteria and *hgcA*⁺ Lentisphaerae were present only during winter, while *hgcA*⁺ Goldbacteria was found only in summer.

3.3.2. *hgcB* coverage

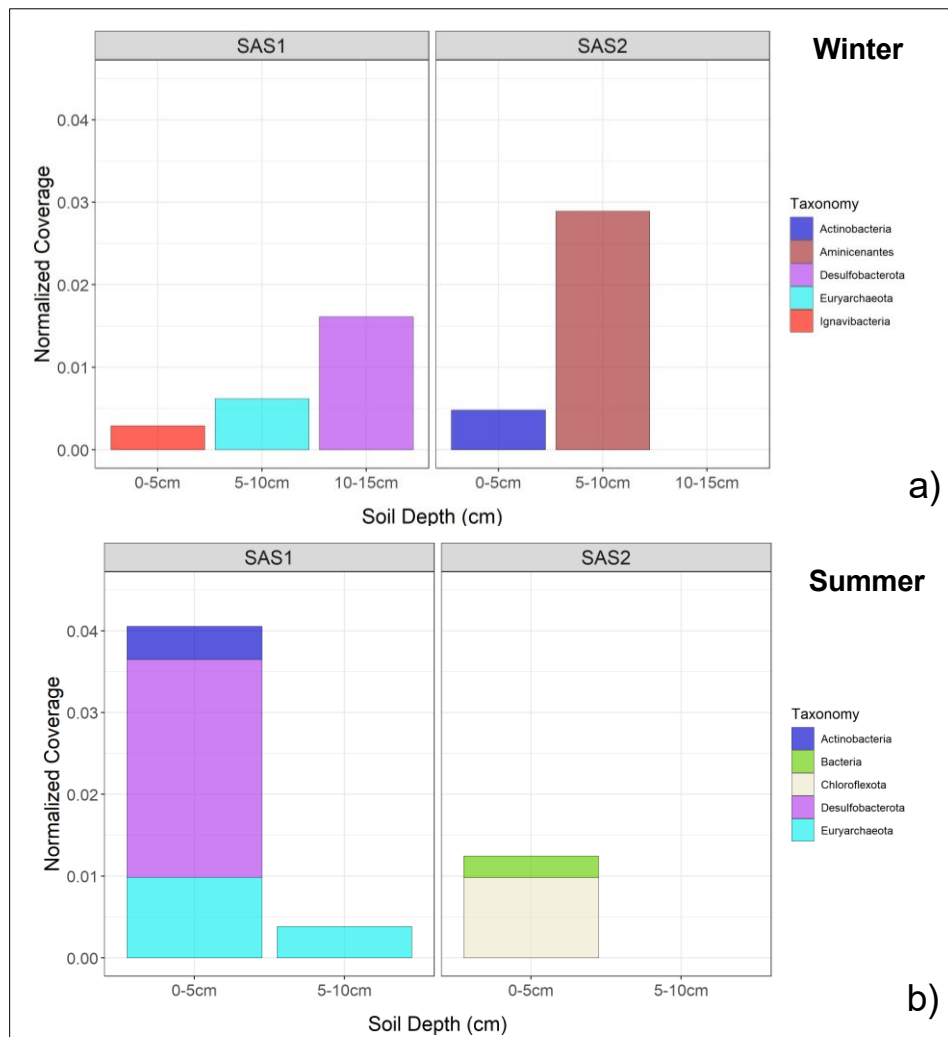


Figure 7. *hgcB* gene coverage across lakes, depths and seasons. Stacked bar plot (a) Winter and bar plot (b) Summer, each color represents a different identified taxonomy. Both plots display the taxonomic composition based on the normalized coverage of *hgcB* gene, across different soil depths (0-5cm, 5-10cm, and 10-15cm) at two sampling sites, SAS1 and SAS2.

hgcB true genes were scarce in all samples reflecting a low overall coverage across all lakes, depths, and seasons (Figure 7). Results showed that *hgcB*⁺ taxa differed greatly between lakes, with phyla *Ignavibacteria*, *Euryarchaeota*, *Desulfobacterota* present only in lake SAS1A and phyla *hgcB*⁺ *Aminicenantes* and *Chloroflexota* present only in lake SAS2. Among the identified phyla, *Actinobacteriota* *hgcB*⁺ were the only identified in both lakes.

SAS1 exhibited a greater diversity, and a generally higher coverage compared to SAS2. Furthermore, in winter (Figure 7a), *Aminicenantes* and *Desulfobacterota* showed higher coverage values while, in summer (Figure 7b), *Desulfobacterota* showed the highest coverage, followed by *Euryarchaeota*. Moreover, *Aminicenantes* and *Ignavibacteria* were only present in winter, whereas *Chloroflexota* appeared exclusively in summer.

3.3.3. *merA* coverage

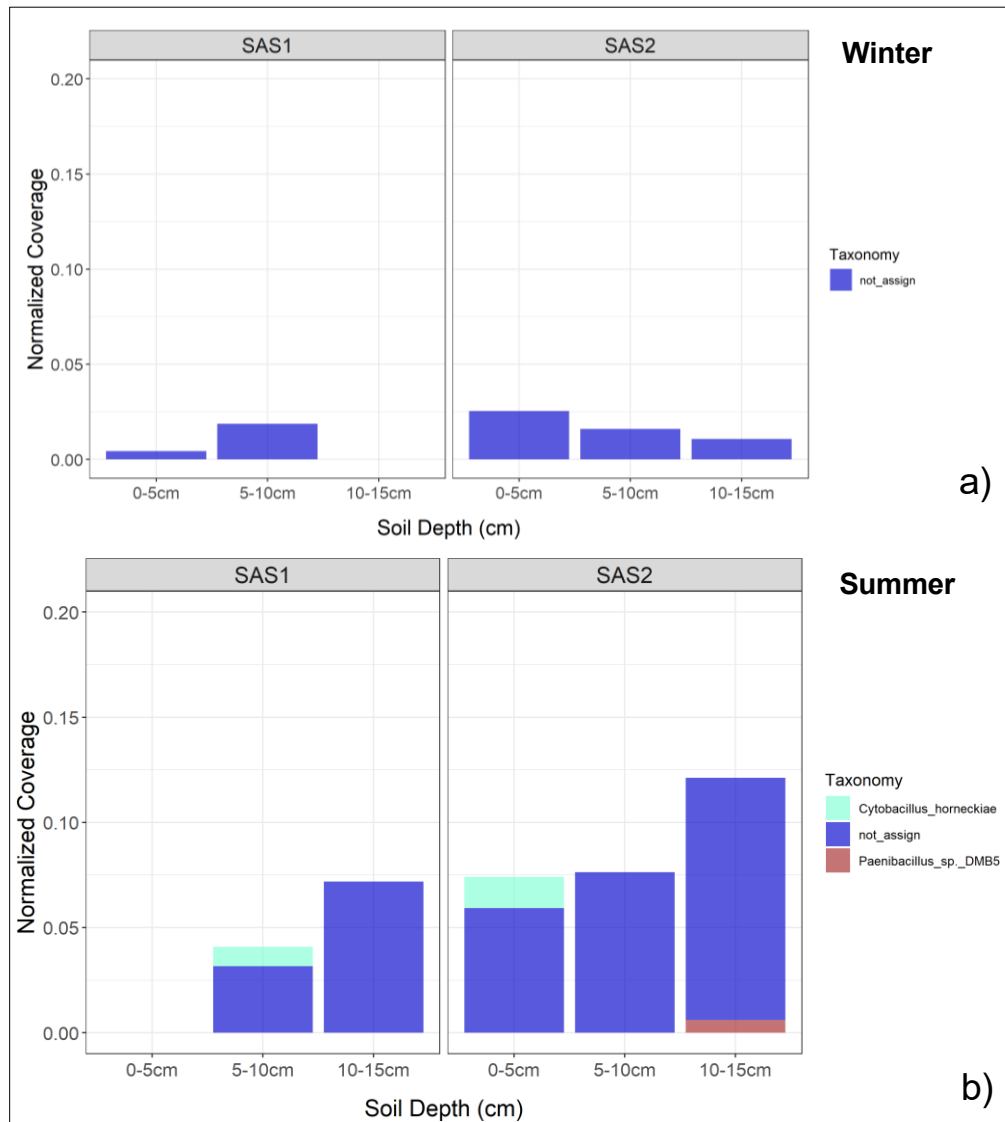


Figure 8. *merA* gene coverage across lakes, depths and seasons. The figure displays two stacked bar plots, bar plot (a) Winter and bar plot (b) Summer where each color represents a different microbial taxon. Both plots display the taxonomic composition based on the normalized coverage of *merA* gene, across different soil depths (0-5cm, 5-10cm, and 10-15cm) at two sampling sites, SAS1 and SAS2.

Results on *merA* and *merB* gene coverages showed that these genes were present in both lakes and in both seasons (Figure 8). *merA* coverage in winter (Figure 8a) was low in both lakes and no taxonomy could be assigned to these particular microorganisms. In summer (Figure 8b) *merA* gene coverage was higher and two species were identified, namely, *Cytobacillus horneckiae* and *Paenibacillus sp. DMB5*. We can also observe that coverage of the *merA* gene tended to decrease with depth in winter and increased with depth in summer in both lakes.

3.3.4. *merB* coverage

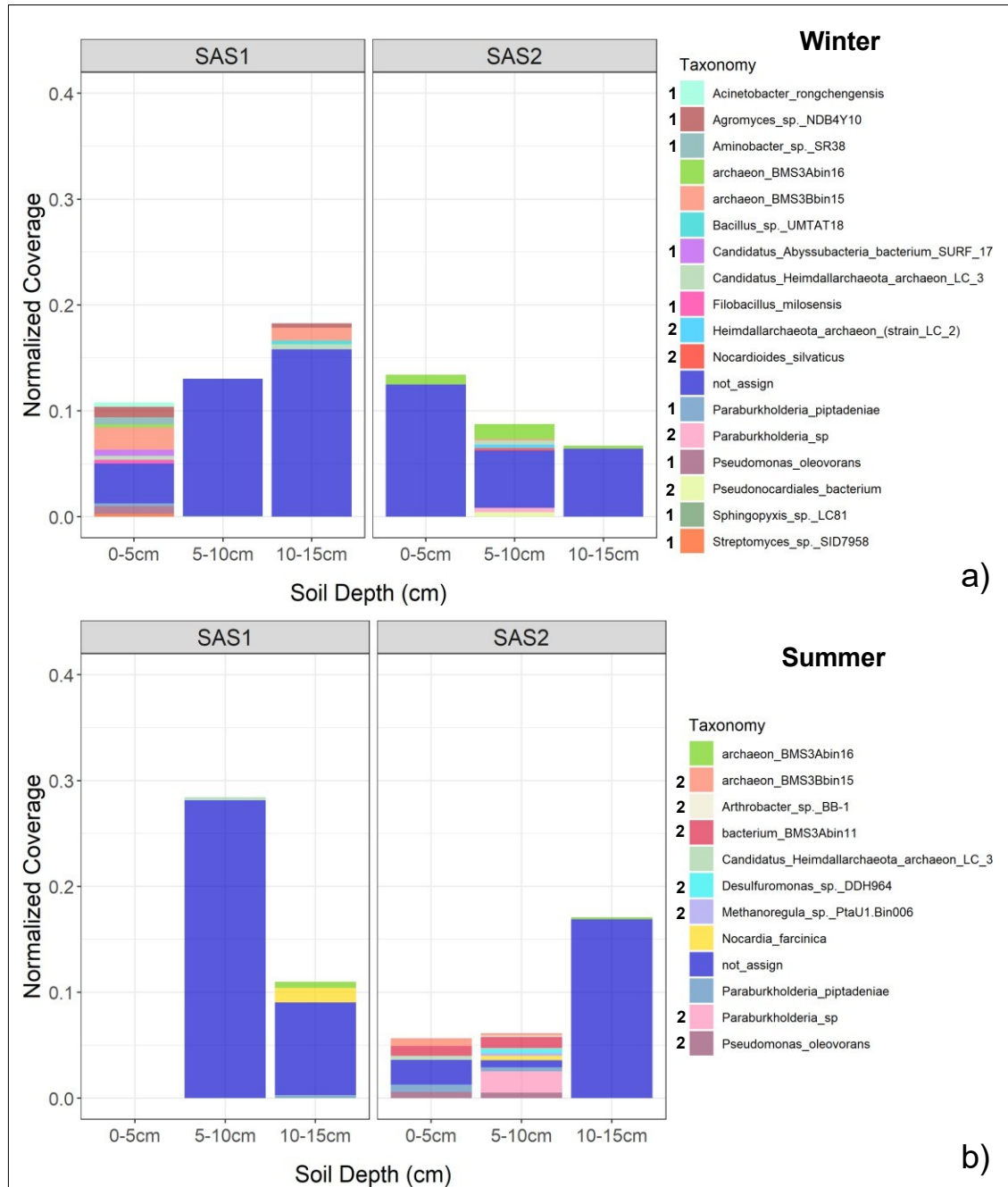


Figure 9. *merB* gene coverage across lakes, depths and seasons. Bar plot (a) Winter and bar plot (b) Summer, each color represents a different microbial taxon. Both plots display the taxonomic composition based on the normalized coverage of *merB* gene, across different soil depths (0–5 cm, 5–10 cm, and 10–15 cm) at two sampling sites, SAS1 and SAS2. Taxa found exclusively in a specific lake are labelled with the number of the lake where they were discovered (SAS1- 1 and SAS2- 2).

Results regarding the *merB* gene coverages (Figure 9) showed a higher degree of taxonomic assignment when compared to *merA*. Overall, *merB* coverage was higher than *merA* in both winter and summer.

In winter (Figure 9a), among all identified microorganisms, *merB*+ *archaeon BMS3Bbin15* had the highest coverage followed by *archaeon BMS3bin16* and *Agromyces*

sp. NDB4Y10. In lake SAS1 there was an increasing trend in *merB* coverage with depth, while the opposite occurred in lake SAS2. Differences were seen considering the taxonomic classification of the identified microorganisms between lakes. In lake SAS1 we identified *merB*+ *Actinobacter rongchengensis*, *Agromyces sp. NDB4y10*, *Aminobacter sp. NDB4Y10*, *Aminobacter sp. SR38*, *Bacillus sp. UMTAT18*, *Candidatus Abyssubacteria bacterium SURF 17*, *Filobacillus milosensis*, *Sphingopyxis sp. LC81* and *Streptomyces sp. SID7958*, whilst in lake SAS2 we exclusively observed *merB*+ *Heimdallarchaeota archeon strain LC 2*, *Nocardioides silvaticus*, *Paraburkholderia sp.* and *Pseudonocardiales bacterium*.

In summer (Figure 9b), in lake SAS1, fewer *merB*+ microorganisms were assigned a taxonomy, namely, *archeon BMS3Abin16*, *Candidatus heimdallarchaeota archeon LC3*, *Nocardia farcinica* and *Paraburkholderia piptadeniae*. In lake SAS2, however, a total of 11 microorganisms were identified (see figure for more details).

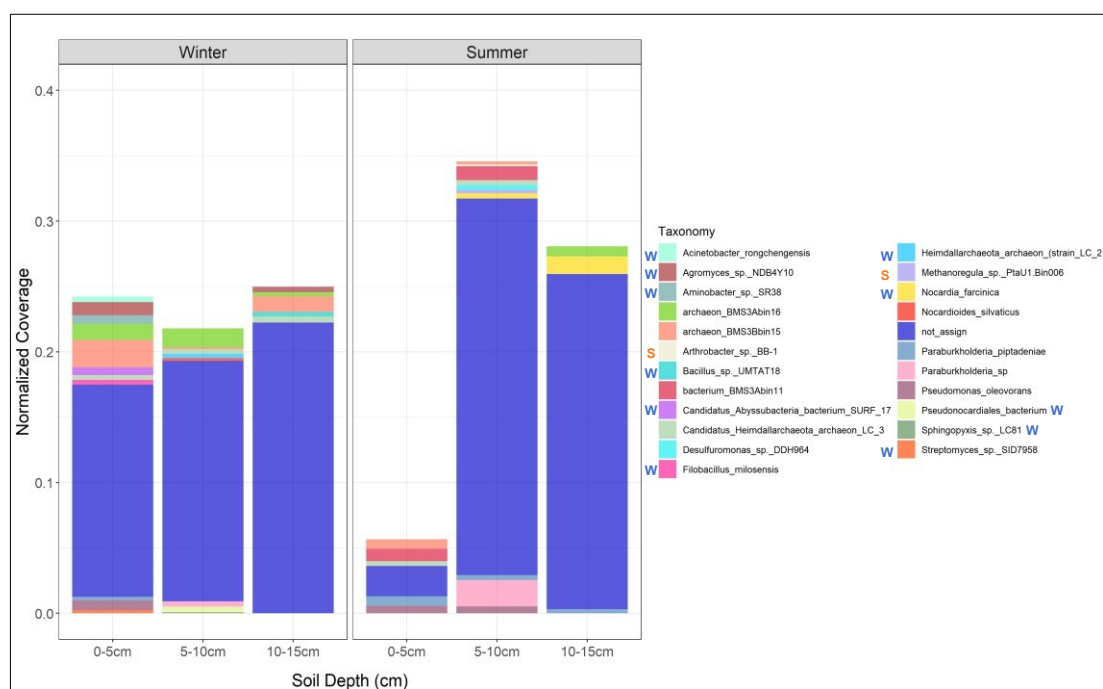


Figure 10. *merB* gene coverage comparative analysis between seasons. Stacked bar plot showing winter (left) and summer (right) normalized coverage of *merB* gene across different sediment depths. The y-axis represents normalized coverage, indicating the relative abundance of each taxon represented by different colors. Taxa found exclusively in a season are labelled accordingly (Winter- W; Summer- S).

Contrasting seasons (Figure 10), we observed that *merB* coverage was more uniform across depths in winter, while in summer, lowest value was observed at 0-5cm, followed by an increase in the following depths (5-10cm and 10-15cm). Considering season specificities, *merB*+ *Arthrobacter sp. BB-1*, *Bacterium BMS3Abin11*, *Desulfuromonas sp. DDH964*, and *Nocardia farcinica* were only present in summer, whereas *merB*+ *Acinetobacter rongchengensis*, *Agromyces sp. NDB4Y10*, *Aminobacter sp. SR38*,

Bacillus sp. UMTAT18, *Candidatus-Abyssobacterium bacterium SURF 17*, *Filobacillus milosensis*, *Heimdallarchaeota archaeon strain LC2*, *Nocardia silvaticus*, *Pseudonocardiales bacterium*, *Sphingopyxis sp. LC81*, and *Streptomyces sp. SID7958* were observed only in winter.

3.4. Environmental mercury data

The concentrations of total mercury (THg) and MMHg were measured across the different sediment depths (0-5cm, 5-10cm, 10-15cm) for both lakes (SAS1 and SAS2) during winter and summer 2022 (analysis performed by the group of Química Estrutural in IST, Lisbon). It is worth mentioning that THg measured also englobes MMHg concentration levels.

3.4.1. Total mercury

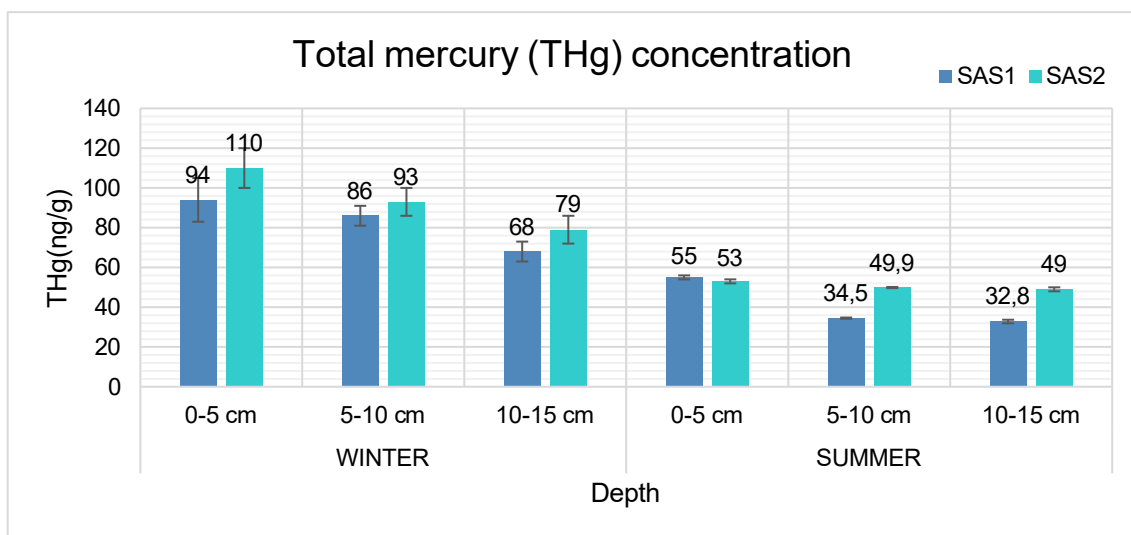


Figure 11. Total mercury (THg) concentrations. Winter (left) and summer (right) THg concentrations measured in (ng/g) at lake SAS1 (dark blue) and lake SAS2 (light blue). The bars represent the mean THg concentrations (ng/g), and the errors correspond to the standard deviation (SD) of the means.

As for THg (Figure 11), in Winter, for Lake SAS1, concentration was highest in the top sediment layer (0-5cm) with a mean of 94 ± 11 ng/g, decreasing with depth to 86 ± 5 ng/g at 5-10cm and further to 68 ± 5 ng/g at 10-15cm. In Lake SAS2, a similar trend was observed with the highest THg concentration in the top layer at 110 ± 10 ng/g, decreasing to 93 ± 7 ng/g at 5-10cm and 79 ± 7 ng/g at 10-15cm. Conversely, in summer, for Lake SAS1, THg concentrations were significantly lower comparing to winter, with the top layer having a mean of 55 ± 1 ng/g, decreasing to 34.5 ± 0.3 ng/g at 5-10cm and 32.8 ± 0.9 ng/g at 10-15cm. Similarly, Lake SAS2 exhibited lower THg levels in summer, with the top layer at 53 ± 1 ng/g, decreasing to 49.9 ± 0.3 ng/g at 5-10cm and 49 ± 1 ng/g at 10-15cm.

3.4.2. Monomethylmercury

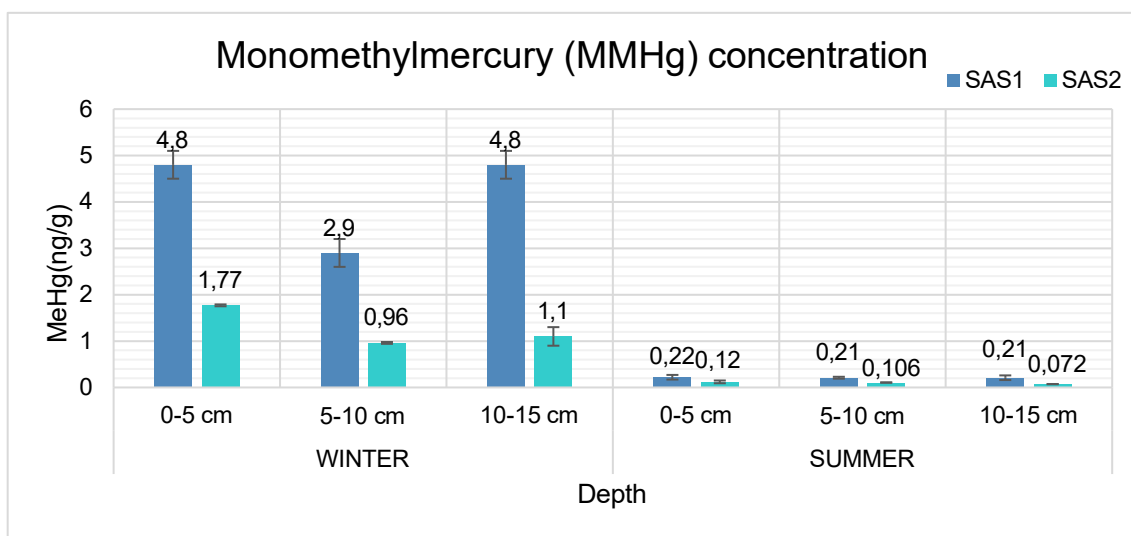


Figure 12. Monomethylmercury (MMHg) concentrations. Winter (left) and summer (right) MMHg concentrations measured in (ng/g) at lake SAS1 (dark blue) and lake SAS2 (light blue). The bars represent the mean MMHg concentrations (ng/g), and the errors correspond to the standard deviation (SD) of the means.

MMHg concentrations (Figure 12) during winter in lake SAS1 were highest in the top layer (0-5cm) at 4.8 ± 0.3 ng/g, decreasing to 2.9 ± 0.3 ng/g at 5-10cm, and increasing again to 4.8 ± 0.3 ng/g at 10-15cm. Likewise, in Lake SAS2, the highest MMHg concentration was also in the top layer at 1.77 ± 0.02 ng/g, decreasing to 0.96 ± 0.02 ng/g at 5-10cm, and slightly increasing to 1.1 ± 0.2 ng/g at 10-15cm. As for Summer in Lake SAS1, MMHg concentrations were significantly lower, with the top layer having a mean of 0.22 ± 0.05 ng/g and remaining relatively constant at 0.21 ± 0.02 ng/g at 5-10cm and 0.21 ± 0.05 ng/g at 10-15cm. In Lake SAS2, the top layer had a MMHg concentration of 0.12 ± 0.03 ng/g, decreasing to 0.106 ± 0.004 ng/g at 5-10cm and further to 0.072 ± 0.003 ng/g at 10-15cm.

Overall, THg concentrations were higher in winter than in summer across all sediment depths for both lakes. Noticeably, MMHg concentrations were around five times higher in winter than in summer. Both THg and MMHg generally decreased with depth, though some variations were observed, particularly in winter where MMHg levels increased in the deeper layers for Lake SAS1. While both lakes exhibit similar seasonal and depth-related trends, when compared to Lake SAS1, Lake SAS2 shows higher THg concentrations but lower concentrations of MMHg in both seasons.

4. Discussion

4.1. Taxonomic classification

Kaiju assigns a taxonomic classification to the reads, which are small ~150bp sequences, therefore results may be limited, with the percentage of unclassified reads increasing with taxonomical depth. To overcome this, we chose to use the data relative to phylum level, because it is more accurate and results in a higher percentage of identified reads. Furthermore, due to the relatively low abundance of certain taxa, we limited our analysis to phyla with a relative abundance greater than 0.5% (Figure 5).

Taxonomic classification of reads revealed a diverse set of microorganisms, some of which have been previously identified in Arctic ecosystems, particularly in thermokarst lakes. Two metabarcoding studies in Arctic permafrost are of particular interest. [Schostag et al. \(2015\)](#) evaluated the difference between winter and summer in the active layer of Svalbard permafrost highlighting the presence of Proteobacteria, Actinobacteria, Verrucomicrobia, Acidobacteria, and Gemmatimonadetes. Furthermore, they found seasonal differences in community structure, linking them to temperature fluctuations, freeze-thaw cycles, variations in photosynthetically active radiation and carbon (C) substrate availability ([Schostag et al., 2015](#)). The more recent metabarcoding study by [Scheel et al. \(2022\)](#) on 26,500-year-old thawing permafrost in Northeast Greenland identified a similar community structure, including Proteobacteria, Acidobacteria, Verrucomicrobia, Bacteroidetes, Actinobacteria and Chloroflexi among the more abundant phyla. Although the same phyla were found in our work, their abundance differed between these studies. We found Proteobacteria, Actinobacteria, Bacteroidetes, Euryarchaeota and Firmicutes to be the more abundant phyla in our sampled lakes. It is worth mentioning that in our study the presence of Euryarchaeota was considerable in both sites and seasons while the above-mentioned studies found low abundance of archaeal phyla. In particular, [Scheel et al. \(2022\)](#) found marginal levels of abundance for Archaeal phyla including Euryarchaeota. These differences can be accentuated by the different methodologies used, since the use of metabarcoding and metagenomics in environmental Arctic samples has shown to have different performances in diversity analysis due to the limited number of 16S rRNA genes identified in shotgun metagenomic procedures ([Pascoal et al, 2022](#)). More likely, the variations in the abundance of each phylum can be explained by the unique characteristics of the thermokarst lakes studied.

In a study by [Waldrop *et al.* \(2023\)](#) that evaluated microbial communities across various Arctic regions and that used a metagenomic approach, results were more similar to the ones we describe. With some differences in relative abundances, they also found Proteobacteria and Actinobacteria to be the major phyla, and the presence of Firmicutes, Bacteroidetes and Chloroflexi across the different areas of the study in relevant percentages. Furthermore, Archaeal communities, primarily methanogens from Euryarchaeota, were also observed, such as in our samples. The presence of this phylum is of extreme importance regarding mercury methylation since it includes many known methylators such as methanogens, as before mentioned. As [Waldrop *et al.* \(2023\)](#) highlighted, permafrost communities are sensitive to the age of permafrost, sedimentation, composition, and thawing, which could explain the above-mentioned variability in the relative abundances of each of the major phyla, since each site has its own history. Looking at the seasonal differences highlighted in our study, one of the most outstanding results was the 16% increase in the abundance of Proteobacteria from winter to summer (Figure 5). [Schostag *et al.* \(2015\)](#) referred an increase in Proteobacteria abundance as part of a bacterial community with a higher abundance of nitrogen-fixing bacteria and phototrophic organisms from June to October in the active layer of Svalbard permafrost (which corresponds to part of the summer months) since they are copiotrophic microorganisms, and, during summer months, C and nutrients become more available in this environment, which could explain the growth observed ([MacMillan *et al.*, 2015](#); [Vigneron *et al.*, 2019](#)). Likewise, summer environmental conditions from our study sites can also be enhancing then growth of these particular prokaryotic group.

Additionally, Euryarchaeota, Firmicutes, Planctomycetes, Chlorobi and Chloroflexi showed decreases in abundance from winter to summer which we consider worth mentioning. The high prevalence of Euryarchaeota during winter is of extreme importance, considering their role in Hg methylation. Their seasonal differences in abundance (-4.57%) may be related to strict anaerobiosis, characteristic of methanogens. The lower abundance of Euryarchaeota in summer months in the SAS lakes can possibly be explained by the increased oxygenation of the soil, due to thawing, that consequently reduces anaerobic niches, where they thrive ([Rivkina *et al.*, 2007](#)). As previously mentioned, Firmicutes were among the most abundant phyla in our samples across both seasons, with higher abundance in the winter decreasing in summer (-2.03%). In a study conducted by [Jansson & Taş \(2014\)](#), in the Canadian High Arctic permafrost, spore-forming Firmicutes strains were isolated, and the authors referred that the ability for some members of this phylum to form spores allows them to thrive in

harsher or more nutrient-limited environments, such as those found during winter in SAS lakes. Even with a decrease in abundance during summer, Firmicutes remained one of the most abundant phyla, likely be due to facultative anaerobic metabolism, which allows them to adapt to fluctuating O₂ levels associated with seasonal thawing. Planctomycetes also showed an interesting decrease in summer (-3.25%). They are known actors in biogeochemical cycles, such as nitrogen cycling and degradation of OM (Zheng *et al.*, 2024), possessing the ability to adapt to low O₂ environments which could favour them during winter over summer that tend to present more aerobic conditions. The decrease of (-1.75%) in Chlorobi abundance may be justified by their anaerobic metabolism capable of reducing sulphur compounds, which can be less abundant during summer months (Oiffer & Siciliano., 2009). Other phylum that we observed to decrease in abundance during summer was Chloroflexi (-2.25%). This includes thermophilic species that possess the ability to degrade complex OM (Ji *et al.*, 2020). During winter, their ability to adapt to freezing temperatures might offer them a competitive advantage, allowing them to be more active. With warmer temperatures, the increase in microbial competition from aerobes capable of degrading OM may explain the observed decrease in abundance. Overall, during summer, conditions became less favourable for anaerobic metabolism and the microbial community of SAS lakes shifted towards aerobic processes.

4.2. Mercury Cycle

4.2.1. *hgcA* and *hgcB* genes

Microorganisms found to possess the *hgcAB* gene cluster include members from the Bacteria domain, such as Proteobacteria, Desulfobacterota, Firmicutes, Chloroflexi, Bacteroidetes, Nitrospirae, Aminicenantes, Atribacteria, Chlorobi, Spirochaetes, Elusimicrobia, and members of the PVC superphylum (including Verrucomicrobia, Chlamydiae, and Lentisphaerae). Additionally, Archaea, such as Euryarchaeota, also possess this gene cluster (Parks *et al.*, 2013; Schaefer *et al.*, 2014; Gilmour *et al.*, 2018; Capo *et al.*, 2022). Well studied clades like Desulfobacterota, Firmicutes and Euryarchaeota are well known for their Hg methylation capability and to possess *hgcA* genes (Gilmour *et al.*, 2018).

Our results regarding *hgcA* gene coverage (Figure 6) led to an assigned taxonomy consistent with previous studies, particularly highlighting the known major role of Desulfobacterota, an SRB in Hg methylation (Gionfriddo *et al.*, 2020). Chloroflexota and Bacteroidetes also belong to major phyla found to be *hgcA*+, followed by Euryarchaeota and Actinobacteria. These results agree with studies from Capo *et al.*

(2023b) and Gionfriddo *et al* (2020). In spite of their findings showing different levels of coverage among different phyla, probably related to the specific conditions of their samples. In particular, Capo *et al* (2023b) found different leading taxa besides Desulfobacterota, such as Spirochaete, Acidobacteria and Verrucomicrobia having higher *hgcA* gene coverages. However, these results were obtained in lake sediments of a sulphate impacted lake in northern Minnesota. Podar *et al* (2015), analysed thawing permafrost soils and found similar community structures, however with higher abundance of methanogens, Firmicutes and Deltaproteobacteria. This is of particular interest, since a study by Gilmour *et al* (2018) found that methanogens carrying *hgcAB* gene cluster play an equally important role in Hg methylation similar to that of SRB and IRB. This gene cluster is present in *Methanomicrobia* and *Thermoplasmatota* (classes of Euryarchaeota). In light of this finding, Euryarchaeota in our sampling sites may represent one of the most abundant phyla contributing to Hg methylation.

hgcB assigned taxonomy resulted in small levels of coverage across all samples (Figure 7). This can be explained by the fact that, compared to *hgcA*, *hgcB* sequences are still less annotated in reference databases, resulting in poor assignment efficiencies. Also due to the application of a high stringency cutoff (Capo *et al.*, 2022) low number of *hgcB*+ microorganisms were identified. Nonetheless, as stated before, both genes are required for methylation to occur, due to their co-location in the genome (Parks *et al.*, 2013) and thus the presence of *hgcA* is enough to demonstrate potential of methylation. Based on this, we chose to conduct the discussion of the methylation potential of SAS lakes microbial communities, solely based on *hgcA* coverage results.

Seasonal differences regarding *hgcA*+ microorganisms taxonomy were not evident since the same taxa were identified in both winter and summer, except for Lentisphaerae in winter and Goldbacteria in summer, both presenting low coverage values. However, since no information on the literature was found regarding these microorganisms in particular, further investigation is required for evaluating these differences and investigating their significance. The principal difference between seasons lies in the coverage range of *hgcA* genes. Winter presented higher coverage values when compared to summer. This suggests that, during winter, there may be a higher level of Hg methylation, although an evaluation of the *hgcA* transcripts is needed to fully confirm this finding. Nonetheless, a novel study by Capo *et al.* (2023b) evaluating the expression levels of *hgcAB* genes in stratified brackish waters found that the presence or expression of the *hgcA* gene, along with Hg(II) speciation, improved the prediction of Hg(II) methylation in oxygen-depleted waters. Supporting these findings, the same study also

reported higher concentrations of MMHg in O₂ deficient environments (Capo *et al.*, 2023b). Thermokarst lakes, in winter, are known to possess ideal conditions for the formation of an anoxic environment, since ice and snow cover forms a barrier blocking atmospheric exchange and sunlight resulting in a rapid depletion of O₂ (Vigneron *et al.*, 2019). Furthermore, microorganisms that methylate, with exception from fermentative Firmicutes, are strictly anaerobic bacteria that use sulphate, iron and CO₂ as terminal electron acceptors (Gilmour *et al.*, 2018). Considering this, our high coverage values of *hgcA* genes in winter may be explained by the O₂ depletion during this season favouring an anaerobiotic community structure found in both SAS lakes. The probable higher Hg-methylation rates in winter are also supported by the increased winter MMHg concentrations found in our study. Similar findings have generally been observed in O₂ depleted environments and methylation is known to be a microbiological mediated process favoured in reducing environments (Capo *et al.*, 2022; Hudelson *et al.*, 2020).

4.2.2. *merA* and *merB* genes

In order to provide a simpler analysis of *merA* and *merB* results, taxonomy will be grouped by the correspondent phylum as described in Table 3.

Table 3. Distribution of *merA* and *merB* genes across microbial taxa, grouped by phylum.

	Phylum	Assigned taxonomy
<i>merA</i>+	Firmicutes	<i>Cytobacillus horneckiae</i> <i>Paenibacillus</i> sp. DMB5
	Candidatus Heimdallarchaeota	<i>Candidatus Heimdallarchaeota archeon</i> LC2 <i>Candidatus Heimdallarchaeota archeon</i> LC3
<i>merB</i>+	Actinobacteria	<i>Pseudonocardiales bacterium</i> <i>Agromyces</i> sp. NDB4Y10 <i>Nocardia silvaticus</i> <i>Streptomyces</i> sp. SID7958 <i>Nocardia farcinica</i> <i>Arthobacter</i> sp. BB-1
	Firmicutes	<i>Bacillus</i> sp. UMTAT18 <i>Filobacillus milosensis</i>
	Proteobacteria	<i>Actinobacter rongchengensis</i> <i>Sphingopyxis</i> sp. LC81 <i>Aminobacter</i> sp. SR38 <i>Paraburkholderia piptadeniae</i> <i>Paraburkholderia</i> sp. <i>Pseudomonas olevatorans</i>
	Thermodesulfobacteriota	<i>Desulforomonas</i> sp. DDH964
	Candidatus Hydrogenedentota	<i>Candidatus Abyssubacteria bacterium</i> SURF 17
	Euryarchaeota	<i>Methanoregula</i> sp. PtaU1. Bin006
	Unidentified Archaea	<i>BMS3Bbin15</i> (Archaeon) <i>BMS3bin16</i> (Archaeon)
	Other/Unclassified	<i>Bacterium BMS3Abin11</i>

Possibly due to the high stringency threshold applied to both *merA* and *merB* in order to select for true genes, our results showed a low number of sequences and consequently low coverage values in both seasons with some depths not having any coverage values attributed (Figures 8-10). Furthermore, taxonomy assigned to *merA* and *merB* had

limitations, as explained above (see methods). Nevertheless, coverage of *merA* was apparently higher in summer relative to winter. Seasonal shifts concerning *merB* coverages were less evident than *merA* since the former gene coverages were, overall, relatively similar among seasons. These differences will be further explained considering the concentrations of both THg and MMHg in the environment.

Previous findings regarding *merA*+ microorganisms, found that these are widely dispersed across different environments (Christakis *et al.*, 2021). The same study points that the *merA* gene was discovered in a large number of archaeal and bacterial phyla including Bacteroidetes, Firmicutes, Actinobacteria, and Proteobacteria. Alas, we could only identify two *merA*+ microorganisms that belong to Firmicutes phylum. A recent study by Rincón-Tomás *et al.* (2024), found that, despite its known participation in Hg methylation, Desulfobacterota may also play an interesting role in Hg reduction, due to the high presence of *merA* genes they observed in this group. Desulfobacterota were one of the most abundant *hgcA*+ phyla in our samples and although the presence of *merA* and *hgcAB* wasn't found in the same microorganisms (Rincón-Tomás *et al.*, 2024) it would be important for future studies to explore their dual role in Hg transformations. This could provide a better understanding of how the presence of this phylum impacts both the formation and elimination of MMHg.

Taxonomy regarding *merB*+ microorganisms yielded better results. Different phyla were identified, in accordance with previous works, including known *merB*+ phyla such as Proteobacteria, Firmicutes, Actinobacteria and Euryarchaeota (Boyd & Barkay, 2012; Christakis *et al.*, 2021; Rincón-Tomás *et al.*, 2024) among other phyla - Candidatus Heimdallarchaeota (Christakis *et al.*, 2021), Thermodesulfobacteriota (SRB) and Candidatus Hydrogenedentota, a rare candidate phylum usually with low environmental abundances in spite of its ubiquitousness (Skoog & Bosak, 2024).

Most identified *mer* operons sequentially encode both MerA and MerB or only MerA (Krout *et al.*, 2022). While this would be important to consider in future studies, our work did not evaluate the possible simultaneous presence of both *merA* and *merB* within the same genome. Therefore, it remains unclear whether these genes belong to the same organism or function independently. Nevertheless, recently, MerB-only encoding genomes have been identified and are thought to be as common as *merA*-only microorganisms (Krout *et al.*, 2022; Barkay & Gu 2022). Given the role of *merB* in the first step of MMHg demethylation by converting it to Hg(II), its presence in the environment is of extreme value. This conversion helps detoxify the environment by reducing MMHg toxicity, which eventually may benefit ecosystems health. Therefore,

microorganisms known to possess *mer* genes could potentially play a crucial role in bioremediation of mercury-contaminated environments as previously mentioned in a study by [Dash & Das, \(2012\)](#).

4.2.3. Seasonal dynamics of mercury concentrations

THg and MMHg concentrations found in our study are relatively high when comparing to other studies, for instance in the subarctic Fennoscandia summer soils ([Tarbier *et al.*, 2021](#)) and in arctic and subarctic ponds ([MacMillan *et al.*, 2015](#)).

When analysing THg concentrations in both seasons we see that winter holds the higher concentrations in both lakes (Figure 11). The same trend is visible regarding MMHg concentrations (Figure 12). Therefore, we can postulate that there is probably a higher methylation rate in winter considering that *MMHg/THg* concentration ratio is ten times higher in winter (0.03) than in summer (0.003). As seen in previous sections, *hgcA* gene coverages also showed higher values during winter, further supporting the hypothesis that methylation rates are higher than demethylation rates during this season in these lakes. In spite of this, we observed a high coverage of *merB* genes during winter, suggesting that demethylation may still be occurring in these lakes to help balance the MMHg concentrations. This could be explained by the high concentrations of THg in winter, as Hg demethylation is influenced by Hg availability in the environment [Nascimento *et al.* \(2003\)](#). Increased Hg levels can activate the *mer* operon, initiating the demethylation pathway. However, coverage of *merA* was low in winter, which suggests that demethylation is dependent on *merB*. Nonetheless, a study by [Krout *et al.* \(2022\)](#) suggested that intracellular production of Hg(II), as a result of demethylation, leads to higher toxicity compared to MMHg itself if Hg(II) is not reduced or neutralized. Therefore, *merB* must be accompanied by additional mechanisms to either reduce, sequester, or redistribute produced Hg(II). Without these, demethylation of MMHg might cause increased cell toxicity and lysis ([Krout *et al.*, 2022](#)). Other non-*mer*-mediated demethylation pathways can be occurring in the environment to counterbalance the formation of Hg(II). Abiotic mechanisms have been documented to aid in Hg(II) reduction, in particular reduction by humic acids associated with natural organic matter (NOM) ([Christakis *et al.*, 2021](#)). These transformations commonly occur in anoxic environments high in NOM, which appears to be the case of our study sites during the winter ([Christakis *et al.*, 2021](#)). Although this can explain reduction of Hg(II) independently of *mer*, another hypothesis described by [Lin *et al.* \(2023\)](#), based on the metabolism of SRB, is worth mentioning, given the high abundance of SRBs found in SAS lakes. When these bacteria break down sulphate, they release H₂S as a byproduct. This reacts with iron in the soil

or sediments leading to the formation of highly reactive iron sulphides. The reaction between MMHg and these iron sulphides transforms MMHg into dimethylmercury ((CH₃)₂Hg) that is efficiently volatilized by the microorganisms. Having all this in consideration the most probable cause for high MMHg in winter is due to higher methylation, supported by the higher coverage values of the *hgcA* gene. Nonetheless, possible *merB*-mediated demethylation in winter should also be taken in consideration.

In summer, despite lower *hgcA* coverages, the potential for methylation is still present in SAS lakes. However, a significant loss of MMHg occurs, probably due to environmental changes caused by thawing and the influx of O₂ into the soil. This shift toward aerobic conditions can affect the methylators that are typically more active under anaerobic conditions, making methylation less dominant and increasing demethylation in the lakes, leading to a subsequent loss of MMHg. This is supported by the higher coverage values of both *merA* and *merB* genes during the warmer months. However, although MMHg pool can be greatly affected by *mer*-dependent demethylation, further analysis are required to understand to what extent this occurs within SAS lakes, since the high stringency cutoffs applied may have masked the true coverages of *mer* genes in our study. Additionally, various other MMHg demethylation pathways have been identified (Barkay & Gu, 2022) which could also contribute to MMHg loss during the warmer months. Thermokarst lake sediments are anoxic, dark environments with abundant labile DOM from surrounding vegetation leeches (Folhas *et al.*, 2020), that may favour non-*mer*-dependent demethylation mechanisms as previously discussed (see introduction). If, in fact, transcripts of *merA* and *merB* had been evaluated in these sites we could better comprehend their impact in MMHg transformations. Besides, an evaluation of the lake's DOM, sulphide and O₂ concentrations would allow us to investigate if *mer*-dependent demethylation plays a higher role in MMHg depletion of the SAS lakes rather than *mer*-independent and abiotic demethylation. Another possible cause for the loss of MMHg during summer could be its release into the water column as a consequence of thawing. This was previously observed by (Lehnher *et al.*, 2012) that found that MMHg concentrations in water were partially controlled by MMHg production in sediments.

Overall, our results support the idea that there is a shift in the microbial community from winter to summer, accompanied by a change on the predominant Hg-cycle pathway, from methylation in winter to demethylation in summer. The transition to aerobic conditions in the summer may be what enables these transformations since the availability of O₂ can, potentially, favour MMHg demethylation processes and inhibit

anaerobic Hg-methylation. In future studies transcription of genes should be accessed in order to better comprehend the weight *hgc* and *mer* genes have in Hg transformations.

4.2.4. Contrasts between the thermokarst lakes

Differences between lakes were only evident in THg concentrations and MMHg concentrations. No statistical tests were made regarding differences between gene coverages. SAS1 overall presents lower levels of THg and higher levels of MMHg in both seasons compared to SAS2. Based on these results we can hypothesise a higher methylation rate in lake SAS1 than SAS2. [Hudelson *et al* \(2020\)](#), demonstrated that the maximum depth of the lake had a stronger influence on methylation than Hg concentrations or OM content. In their study the deeper lakes showed greater demethylation potential, and the shallowest lakes had higher methylation potential. Given the study lake depths, SAS1 is the shallowest lake (1.3 m) compared to SAS2 (2.5 m). For lake SAS1 methylation ratio (*MMHg/THg*) values were 0.05 (winter) and 0.006 (summer), and for lake SAS2 0.01 (winter) and 0.002 (summer). These suggest that our results are in accordance with [Hudelson *et al* \(2020\)](#) findings; lake SAS1, the shallowest lake, had a higher methylation ratio while the deeper lake SAS2 had a lower methylation rate. Furthermore, as previously observed, lake SAS1 is closer than SAS2 to the nearest palsa, which could lead to a greater import of OM that also impacts methylation which could explain these differences ([Folhas *et al.*, 2020](#)).

5. Conclusion

This master's thesis was part PERMAMERC project that aimed to improve understanding of Hg biogeochemistry in permafrost thaw systems and its impacts on Arctic ecology, ecosystems, and human health. Permafrost soils are thought to be huge reservoirs for Hg, and the continuous thawing is releasing this compound into the surrounding environment, posing a huge threat to human and wildlife. Therefore, it is important to thoroughly comprehend the pathways of the Hg cycle, and how climate change is affecting them, and the consequent MMHg release. Microorganisms play a crucial role in Hg transformations; therefore, the goal of this thesis was to evaluate how the community structure of thermokarst lakes of northern Canada is related to Hg methylation and demethylation processes and further evaluate the differences between winter and summer seasons.

An array of microorganisms previously known to be involved in Hg cycle were identified in both lakes, at all depths and seasons. We identified several phyla as *hgcA*⁺ and showing high coverages of this gene, including Desulfobacterota, Chloroflexota, Bacteroidetes, Euryarchaeota and Actinobacteria. We observed a higher *hgcA*⁺ gene coverage and higher taxonomic diversity in winter, when compared to summer and postulated an increased methylation activity during winter. Considering demethylation pathways, we showed mer-mediated demethylation potential was higher in summer, and proposed this was due to the environmental shift from anaerobic to aerobic conditions. Additionally, microorganisms with demethylation potential were present in accordance with previous studies and included, Firmicutes, Proteobacteria, Actinobacteria; Thermodesulfobacteriota; Candidatus Hydrogenedentota and Euryarchaeota.

A meta transcriptomic analysis would be a beneficial addition to this study since it could provide information and prove functionality of the genes in study, since eDNA analysis only, although important for gene and pathways assessment, is not enough to prove gene functionality. Finally, these results highlight the complexity of Hg cycle dynamics in thermokarst lakes. The information here gathered along with other works by the PERMAMERC project have a big potential for providing valuable information on the Hg cycle. In a world where global temperatures continue to increase, it is essential to understand how climate change is affecting Arctic ecosystems, since they have a higher vulnerability, to anticipate and mitigate environmental risks, in particular the case of this study related to Hg release.

References

- Abo-Elkhier, M. M., Abd Elwahaab, M. A., & Abo El Maaty, M. I. (2019). Measuring similarity among protein sequences using a new descriptor. *BioMed Research International*, 2019, 2796971. <https://doi.org/10.1155/2019/2796971>
- Andrews, S. (2010). FASTQC. *A quality control tool for high throughput sequence data*. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
- National Research Council (US) Committee on Metagenomics: Challenges and Functional Applications. (2007). *The new science of metagenomics: Revealing the secrets of our microbial planet*. National Academies Press. <https://www.ncbi.nlm.nih.gov/books/NBK54011/>
- Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., Sneddon, M. W., Henderson, M. L., Riehl, W. J., Murphy-Olson, D., Chan, S. Y., Kamimura, R. T., Kumari, S., Drake, M. M., Brettin, T. S., ... Yu, D. (2018). KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology*, 36(7), 566–569. <https://doi.org/10.1038/nbt.4163>
- Barberán, A., Ramirez, K. S., Leff, J. W., Bradford, M. A., Wall, D. H., & Fierer, N. (2014). Why are some microbes more ubiquitous than others? Predicting the habitat breadth of soil bacteria. *Ecology Letters*, 17(7), 794–802. <https://doi.org/10.1111/ele.12282>
- Barkay, T., & Gu, B. (2022). Demethylation—The Other Side of the Mercury Methylation Coin: A Critical Review. *ACS Environmental Au*, 2(2), 77–97. <https://doi.org/10.1021/acsenvironau.1c00022>
- Blunden, J., Boyer, T., & Bartow-Gillies, E. (2023). State of the Climate in 2022. *Bulletin of the American Meteorological Society*, 104(9), S1–S516. <https://doi.org/10.1175/2023BAMSStateoftheClimate.1>
- 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>
- Boyd, E., & Barkay, T. (2012). The Mercury Resistance Operon: From an Origin in a Geothermal Environment to an Efficient Detoxification Machine. *Frontiers in Microbiology*, 3. <https://doi.org/10.3389/fmicb.2012.00349>

- 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), 14255. <https://doi.org/10.1038/ncomms14255>
- Bravo, A. G., Kothawala, D. N., Attermeyer, K., Tessier, E., Bodmer, P., Ledesma, J. L. J., Audet, J., Casas-Ruiz, J. P., Catalán, N., Cauvy-Fraunié, S., Colls, M., Deininger, A., Evtimova, V. V., Fonvielle, J. A., Fuß, T., Gilbert, P., Herrero Ortega, S., Liu, L., Mendoza-Lera, C., ... Amouroux, D. (2018). The interplay between total mercury, methylmercury and dissolved organic matter in fluvial systems: A latitudinal study across Europe. *Water Research*, 144, 172–182. <https://doi.org/10.1016/j.watres.2018.06.064>
- Burn, C. R. (2023). Permafrost. Em *Reference Module in Earth Systems and Environmental Sciences*. Elsevier. <https://doi.org/10.1016/B978-0-323-99931-1.00107-0>
- Capo, E., Cosio, C., Gascón Díez, E., Loizeau, J.-L., Mendes, E., Adatte, T., Franzenburg, S., & Bravo, A. G. (2023b). Anaerobic mercury methylators inhabit sinking particles of oxic water columns. *Water Research*, 229, 119368. <https://doi.org/10.1016/j.watres.2022.119368>
- Capo, E., Feng, C., Bravo, A. G., Bertilsson, S., Soerensen, A. L., Pinhassi, J., Buck, M., Karlsson, C., Hawkes, J., & Björn, E. (2022). Expression Levels of *hgcAB* Genes and Mercury Availability Jointly Explain Methylmercury Formation in Stratified Brackish Waters. *Environmental Science & Technology*, 56(18), 13119–13130. <https://doi.org/10.1021/acs.est.2c03784>
- Capo, E., Peterson, B. D., Kim, M., Jones, D. S., Acinas, S. G., Amyot, M., Bertilsson, S., Björn, E., Buck, M., Cosio, C., Elias, D. A., Gilmour, C., Goñi-Urriza, M., Gu, B., Lin, H., Liu, Y.-R., McMahon, K., Moreau, J. W., Pinhassi, J., ... Gionfriddo, C. M. (2023a). A consensus protocol for the recovery of mercury methylation genes from metagenomes. *Molecular Ecology Resources*, 23(1), 190–204. <https://doi.org/10.1111/1755-0998.13687>
- Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics (Oxford, England)*, 34(17), i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Christakis, C. A., Barkay, T., & Boyd, E. S. (2021). Expanded Diversity and Phylogeny of mer Genes Broadens Mercury Resistance Paradigms and Reveals an Origin for MerA Among Thermophilic Archaea. *Frontiers in Microbiology*, 12. <https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.682605>

- Crevecoeur, S., Vincent, W. F., Comte, J., & Lovejoy, C. (2015). Bacterial community structure across environmental gradients in permafrost thaw ponds: Methanotroph-rich ecosystems. *Frontiers in Microbiology*, 6, 192. <https://doi.org/10.3389/fmicb.2015.00192>
- Danecek, P., Bonfield, J. K., Liddle, J., Marshall, J., Ohan, V., Pollard, M. O., Whitwham, A., Keane, T., McCarthy, S. A., Davies, R. M., & Li, H. (2021). Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2), giab008. <https://doi.org/10.1093/gigascience/giab008>
- Dash, H. R., & Das, S. (2012). Bioremediation of mercury and the importance of bacterial *mer* genes. *International Biodeterioration & Biodegradation*, 75, 207–213. <https://doi.org/10.1016/j.ibiod.2012.07.023>
- 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 & Technology*, 47(10), 4967–4983. <https://doi.org/10.1021/es305071v>
- Edgar, R. C. (2004). MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5), 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Folhas, D., Duarte, A. C., Pilote, M., Vincent, W. F., Freitas, P., Vieira, G., Silva, A. M. S., Duarte, R. M. B. O., & Canário, J. (2020). Structural Characterization of Dissolved Organic Matter in Permafrost Peatland Lakes. *Water*, 12(11), Artigo 11. <https://doi.org/10.3390/w12113059>
- Gilmour, C. C., Bullock, A. L., McBurney, A., Podar, M., & Elias, D. A. (2018). Robust Mercury Methylation across Diverse Methanogenic Archaea. *mBio*, 9(2), e02403-17. <https://doi.org/10.1128/mBio.02403-17>
- Gionfriddo, C. M., Wymore, A. M., Jones, D. S., Wilpiseski, R. L., Lynes, M. M., Christensen, G. A., Soren, A., Gilmour, C. C., Podar, M., & Elias, D. A. (2020). An Improved *hgcAB* Primer Set and Direct High-Throughput Sequencing Expand Hg-Methylator Diversity in Nature. *Frontiers in Microbiology*, 11. <https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2020.541554>
- Gionfriddo, C., Capo, E., Peterson, B., Lin, H., Jones, D., Bravo, A. G., Bertilsson, S., Moreau, J., McMahon, K., Elias, D., & Gilmour, C. (2021). *Hg-MATE-Db.v1.01142021*. <https://doi.org/10.25573/serc.13105370>
- Goñi-Urriza, M., Corsellis, Y., Lancelaur, L., Tessier, E., Gury, J., Monperrus, M., & Guyoneaud, R. (2015). Relationships between bacterial energetic metabolism, mercury

- methylation potential, and *hgcA/hgcB* gene expression in *Desulfovibrio dechloroacetivorans* BerOc1. *Environmental Science and Pollution Research*, 22(18), 13764–13771. <https://doi.org/10.1007/s11356-015-4273-5>
- Ha, E., Basu, N., Bose-O'Reilly, S., Dórea, J. G., McSorley, E., Sakamoto, M., & Chan, H. M. (2017). Current progress on understanding the impact of mercury on human health. *Environmental Research*, 152, 419–433. <https://doi.org/10.1016/j.envres.2016.06.042>
- Hartmuth, K., Boettcher, M., Wernli, H., & Papritz, L. (2022). Identification, characteristics and dynamics of Arctic extreme seasons. *Weather and Climate Dynamics*, 3(1), 89–111. <https://doi.org/10.5194/wcd-3-89-2022>
- Hu, D., Gao, Y. H., Yao, X. S., & Gao, H. (2020). Recent advances in dissecting the demethylation reactions in natural product biosynthesis. *Current opinion in chemical biology*, 59, 47–53. <https://doi.org/10.1016/j.cbpa.2020.04.014>
- Hudelson, K. E., Drevnick, P. E., Wang, F., Armstrong, D., & Fisk, A. T. (2020). Mercury methylation and demethylation potentials in Arctic lake sediments. *Chemosphere*, 248, 126001. <https://doi.org/10.1016/j.chemosphere.2020.126001>
- Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11, 119. <https://doi.org/10.1186/1471-2105-11-119>
- IPCC. (2023). *Climate change 2023: Synthesis report. Contribution of working groups I, II and III to the sixth assessment report of the Intergovernmental Panel on Climate Change* (H. Lee & J. Romero, Eds.). IPCC. <https://doi.org/10.59327/IPCC/AR6-9789291691647>
- Jansson, J. K., & Taş, N. (2014). The microbial ecology of permafrost. *Nature Reviews. Microbiology*, 12(6), 414–425. <https://doi.org/10.1038/nrmicro3262>
- Ji, M., Kong, W., Liang, C., Zhou, T., Jia, H., & Dong, X. (2020). Permafrost thawing exhibits a greater influence on bacterial richness and community structure than permafrost age in Arctic permafrost soils. *The Cryosphere*, 14(12), 3907–3916. <https://doi.org/10.5194/tc-14-3907-2020>
- Khamis, A., Raoult, D., & La Scola, B. (2004). rpoB Gene Sequencing for Identification of *Corynebacterium* Species. *Journal of Clinical Microbiology*, 42(9), 3925–3931. <https://doi.org/10.1128/JCM.42.9.3925-3931.2004>

- Kim, H.-M., Jeon, S., Chung, O., Jun, J. H., Kim, H.-S., Blazyte, A., Lee, H.-Y., Yu, Y., Cho, Y. S., Bolser, D. M., & Bhak, J. (2021). Comparative analysis of 7 short-read sequencing platforms using the Korean Reference Genome: MGI and Illumina sequencing benchmark for whole-genome sequencing. *GigaScience*, 10(3), giab014. <https://doi.org/10.1093/gigascience/giab014>
- Klapstein, S. J., Ziegler, S. E., & O'Driscoll, N. J. (2018). Methylmercury photodemethylation is inhibited in lakes with high dissolved organic matter. *Environmental Pollution*, 232, 392–401. <https://doi.org/10.1016/j.envpol.2017.09.049>
- Krout, I. N., Scrimale, T., Vorojeikina, D., Boyd, E. S., & Rand, M. D. (2022). Organomercurial Lyase (MerB)-Mediated Demethylation Decreases Bacterial Methylmercury Resistance in the Absence of Mercuric Reductase (MerA). *Applied and Environmental Microbiology*, 88(6), e00010-22. <https://doi.org/10.1128/aem.00010-22>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Lehnherr, I., St. Louis, V. L., & Kirk, J. L. (2012). Methylmercury Cycling in High Arctic Wetland Ponds: Controls on Sedimentary Production. *Environmental Science & Technology*, 46(19), 10523–10531. <https://doi.org/10.1021/es300577e>
- 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, X., Cui, Y., Ma, D., Song, D., & Liu, L. (2022). Vertical distribution of bacterial community diversity in the Greater Khingan Mountain permafrost region. *Ecology and Evolution*, 12(7), e9106. <https://doi.org/10.1002/ece3.9106>
- Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Lin, H., Moody, E. R. R., Williams, T. A., & Moreau, J. W. (2023). On the Origin and Evolution of Microbial Mercury Methylation. *Genome Biology and Evolution*, 15(4), evad051. <https://doi.org/10.1093/gbe/evad051>
- Lindsey, R., & Scott, M. (2022, October 18). Climate change: Arctic sea ice summer minimum. NOAA Climate.gov. <https://www.climate.gov/news-features/understanding-climate/climate-change-arctic-sea-ice-summer-minimum>

- MacMillan, G. A., Girard, C., Chételat, J., Laurion, I., & Amyot, M. (2015). High Methylmercury in Arctic and Subarctic Ponds is Related to Nutrient Levels in the Warming Eastern Canadian Arctic. *Environmental Science & Technology*, 49(13), 7743–7753. <https://doi.org/10.1021/acs.est.5b00763>
- Matsen, F. A., Kodner, R. B., & Armbrust, E. V. (2010). pplacer: Linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*, 11(1), 538. <https://doi.org/10.1186/1471-2105-11-538>
- Matveev, A., Laurion, I., Deshpande, B. N., Bhiry, N., & Vincent, W. F. (2016). High methane emissions from thermokarst lakes in subarctic peatlands. *Limnology and Oceanography*, 61(S1), S150–S164. <https://doi.org/10.1002/lno.10311>
- Meng, J., Fan, J., Bhatt, U. S., & Kurths, J. (2023). Arctic weather variability and connectivity. *Nature Communications*, 14(1), 6574. <https://doi.org/10.1038/s41467-023-42351-x>
- Meng, S., Peng, T., Liu, X., Wang, H., Huang, T., Gu, J.-D., & Hu, Z. (2022). Ecological Role of Bacteria Involved in the Biogeochemical Cycles of Mangroves Based on Functional Genes Detected through GeoChip 5.0. *mSphere*, 7(1), e00936-21. <https://doi.org/10.1128/msphere.00936-21>
- Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, 7(1), 11257. <https://doi.org/10.1038/ncomms11257>
- MINAMATA CONVENTION ON MERCURY TEXT AND ANNEXES. (2019). www.mercuryconvention.org
- Nascimento, A. M. A., & Chartone-Souza, E. (2003). Operon mer: Bacterial resistance to mercury and potential for bioremediation of contaminated environments. *Genetics and Molecular Research: GMR*, 2(1), 92–101. http://www.funpecrp.com.br/gmr/year2003/vol1-2/sim0005_full_text.htm
- Navgire, G. S., Goel, N., Sawhney, G., Sharma, M., Kaushik, P., Mohanta, Y. K., Mohanta, T. K., & Al-Harrasi, A. (2022). Analysis and Interpretation of metagenomics data: An approach. *Biological Procedures Online*, 24(1), 18. <https://doi.org/10.1186/s12575-022-00179-7>
- Neuwirth, E. (2022). RColorBrewer: ColorBrewer palettes. <https://CRAN.R-project.org/package=RColorBrewer>

- Oiffer, L., & Siciliano, S. D. (2009). Methyl mercury production and loss in Arctic soil. *Science of The Total Environment*, 407(5), 1691–1700. <https://doi.org/10.1016/j.scitotenv.2008.10.025>
- Oren, A. (2024). *On validly published names, correct names, and changes in the nomenclature of phyla and genera of prokaryotes: A guide for the perplexed*. *npj Biofilms and Microbiomes*, 10, Article 20. <https://doi.org/10.1038/s41522-024-00494-9>
- Otoni, J. R., Oliveira, V. M. de, & Passarini, M. R. Z. (2022). Chapter 1 - Microbes in thawing permafrost: Contributions to climate change. Em A. Kumar, J. Singh, & L. F. R. Ferreira (Eds.), *Microbiome Under Changing Climate* (pp. 1–28). Woodhead Publishing. <https://doi.org/10.1016/B978-0-323-90571-8.00001-8>
- Palangi, S., & Abyaneh, H. (2017). *Impacts of climate change on permafrost as a threat to the future of the Earth*. https://www.researchgate.net/publication/318672969_Impacts_of_Climate_Change_on_PERMAFROST_as_a_threat_to_the_future_of_the_Earth
- Parks, J. M., Johs, A., Podar, M., Bridou, R., Hurt, R. A., Smith, S. D., Tomanicek, S. 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 (New York, N.Y.)*, 339(6125), 1332–1335. <https://doi.org/10.1126/science.1230667>
- Pascoal, F., Costa, R., Assmy, P., Duarte, P., & Magalhães, C. (2022). Exploration of the types of rarity in the Arctic Ocean from the perspective of multiple methodologies. *Microbial Ecology*, 84(1), 59–72. <https://doi.org/10.1007/s00248-021-01821-9>
- Pavithra, K. G., SundarRajan, P., Kumar, P. S., & Rangasamy, G. (2023). Mercury sources, contaminations, mercury cycle, detection and treatment techniques: A review. *Chemosphere*, 312, 137314. <https://doi.org/10.1016/j.chemosphere.2022.137314>
- 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>
- Potter, S. C., Luciani, A., Eddy, S. R., Park, Y., Lopez, R., & Finn, R. D. (2018). HMMER web server: 2018 update. *Nucleic Acids Research*, 46(W1), W200–W204. <https://doi.org/10.1093/nar/gky448>

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.R-project.org/>

Rantanen, M., Karpechko, A., Lipponen, A., Nordling, K., Hyvarinen, O., Ruosteenoja, K., Vihma, T., & Laaksonen, A. (2022). The Arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, 3, 168. <https://doi.org/10.1038/s43247-022-00498-3>

Rincón-Tomás, B., Lanzén, A., Sánchez, P., Estupiñán, M., Sanz-Sáez, I., Bilbao, M. E., Rojo, D., Mendibil, I., Pérez-Cruz, C., Ferri, M., Capo, E., Abad-Recio, I. L., Amouroux, D., Bertilsson, S., Sánchez, O., Acinas, S. G., & Alonso-Sáez, L. (2024). Revisiting the mercury cycle in marine sediments: A potential multifaceted role for *Desulfobacterota*. *Journal of Hazardous Materials*, 465, 133120. <https://doi.org/10.1016/j.jhazmat.2023.133120>

Rivkina, E., Shcherbakova, V., Laurinavichius, K., Petrovskaya, L., Krivushin, K., Kraev, G., Pecheritsina, S., & Gilichinsky, D. (2007). Biogeochemistry of methane and methanogenic archaea in permafrost. *FEMS Microbiology Ecology*, 61(1), 1–15. <https://doi.org/10.1111/j.1574-6941.2007.00315.x>

Schaefer, J. K., Rocks, S. S., Zheng, W., Liang, L., Gu, B., & Morel, F. M. M. (2011). Active transport, substrate specificity, and methylation of Hg(II) in anaerobic bacteria. *Proceedings of the National Academy of Sciences*, 108(21), 8714–8719. <https://doi.org/10.1073/pnas.1105781108>

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 & Technology*, 38(16), 4304–4311. <https://doi.org/10.1021/es049895w>

Schaefer, K., Elshorbany, Y., Jafarov, E., Schuster, P. F., Striegl, R. G., Wickland, K. P., & Sunderland, E. M. (2020). Potential impacts of mercury released from thawing permafrost. *Nature Communications*, 11(1), Artigo 1. <https://doi.org/10.1038/s41467-020-18398-5>

Schaefer, K., Lantuit, H., Romanovsky, V. E., Schuur, E. A. G., & Witt, R. (2014). The impact of the permafrost carbon feedback on global climate. *Environmental Research Letters*, 9(8), 085003. <https://doi.org/10.1088/1748-9326/9/8/085003>

- Scheel, M., Zervas, A., Jacobsen, C. S., & Christensen, T. R. (2022). Microbial Community Changes in 26,500-Year-Old Thawing Permafrost. *Frontiers in Microbiology*, 13, 787146. <https://doi.org/10.3389/fmicb.2022.787146>
- Schloerke, B., Cook, D., Larmarange, J., Briatte, F., Marbach, M., Thoen, E., Elberg, A., & Crowley, J. (2024). *GGally: Extension to ggplot2*. <https://ggobi.github.io/ggally/>
- Schostag, M., Stibal, M., Jacobsen, C. S., Bælum, J., Taş, N., Elberling, B., Jansson, J. K., Semenchuk, P., & Priemé, A. (2015). Distinct summer and winter bacterial communities in the active layer of Svalbard permafrost revealed by DNA- and RNA-based analyses. *Frontiers in Microbiology*, 6, 399. <https://doi.org/10.3389/fmicb.2015.00399>
- Schuster, P. F., Schaefer, K. M., Aiken, G. R., Antweiler, R. C., Dewild, J. F., Gryziec, J. D., Gusmeroli, A., Hugelius, G., Jafarov, E., Krabbenhoft, D. P., Liu, L., Herman-Mercer, N., Mu, C., Roth, D. A., Schaefer, T., Striegl, R. G., Wickland, K. P., & Zhang, T. (2018). Permafrost Stores a Globally Significant Amount of Mercury. *Geophysical Research Letters*, 45(3), 1463–1471. <https://doi.org/10.1002/2017GL075571>
- Skoog, E. J., & Bosak, T. (2024). Predicted metabolic roles and stress responses provide insights into candidate phyla Hydrogenedentota and Sumerlaeota as members of the rare biosphere in biofilms from various environments. *Environmental Microbiology Reports*, 16(1), e13228. <https://doi.org/10.1111/1758-2229.13228>
- Tarbier, B., Hugelius, G., Kristina Sannel, A. B., Baptista-Salazar, C., & Jonsson, S. (2021). Permafrost Thaw Increases Methylmercury Formation in Subarctic Fennoscandia. *Environmental Science & Technology*, 55(10), 6710–6717. <https://doi.org/10.1021/acs.est.0c04108>
- The UniProt Consortium. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, 51(D1), D523–D531. <https://doi.org/10.1093/nar/gkac1052>
- Thomas, T., Gilbert, J., & Meyer, F. (2012). Metagenomics—A guide from sampling to data analysis. *Microbial Informatics and Experimentation*, 2, 3. <https://doi.org/10.1186/2042-5783-2-3>
- VanInsberghe, D., Arevalo, P., Chien, D., & Polz, M. F. (2020). How can microbial population genomics inform community ecology? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 375(1798), 20190253. <https://doi.org/10.1098/rstb.2019.0253>

Vigneron, A., Lovejoy, C., Cruaud, P., Kalenitchenko, D., Culley, A., & Vincent, W. F. (2019). Contrasting Winter Versus Summer Microbial Communities and Metabolic Functions in a Permafrost Thaw Lake. *Frontiers in Microbiology*, 10. <https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2019.01656>

Waldrop, M. P., Chabot, C. L., Liebner, S., Holm, S., Snyder, M. W., Dillon, M., Dudgeon, S. R., Douglas, T. A., Leewis, M.-C., Walter Anthony, K. M., McFarland, J. W., Arp, C. D., Bondurant, A. C., Taş, N., & Mackelprang, R. (2023). Permafrost microbial communities and functional genes are structured by latitudinal and soil geochemical gradients. *The ISME Journal*, 17(8), 1224–1235. <https://doi.org/10.1038/s41396-023-01429-6>

Watson, C. T., Matsen, F. A., Jackson, K. J. L., Bashir, A., Smith, M. L., Glanville, J., Breden, F., Kleinstein, S. H., Collins, A. M., & Busse, C. E. (2017). Comment on «A Database of Human Immune Receptor Alleles Recovered from Population Sequencing Data». *Journal of Immunology (Baltimore, Md.: 1950)*, 198(9), 3371–3373. <https://doi.org/10.4049/jimmunol.1700306>

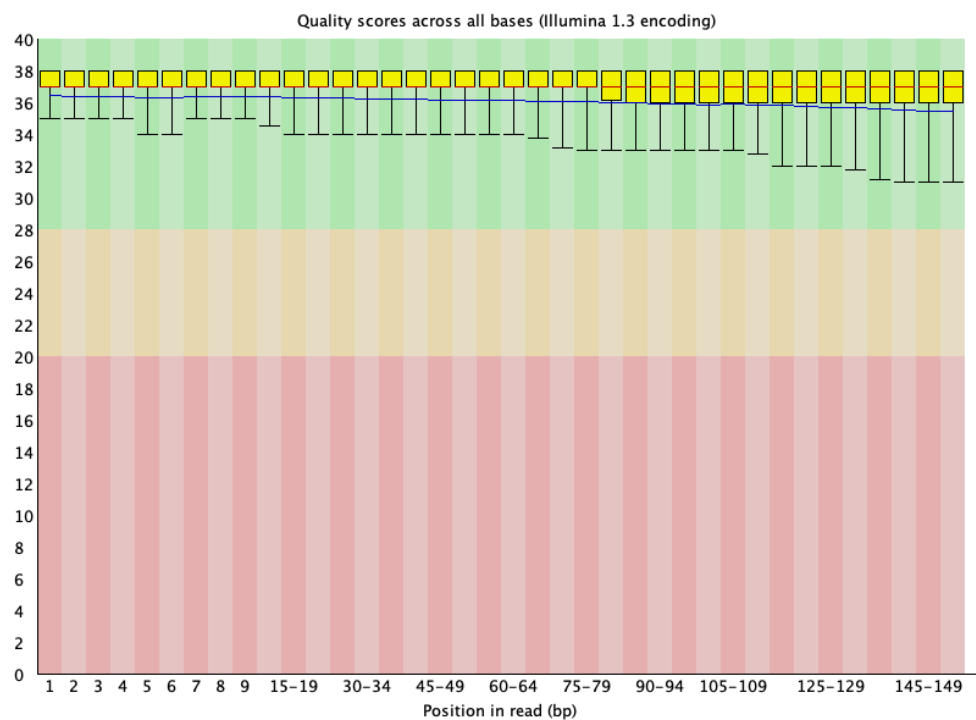
Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation* (R package version 1.1.4). <https://github.com/tidyverse/dplyr>, <https://dplyr.tidyverse.org>

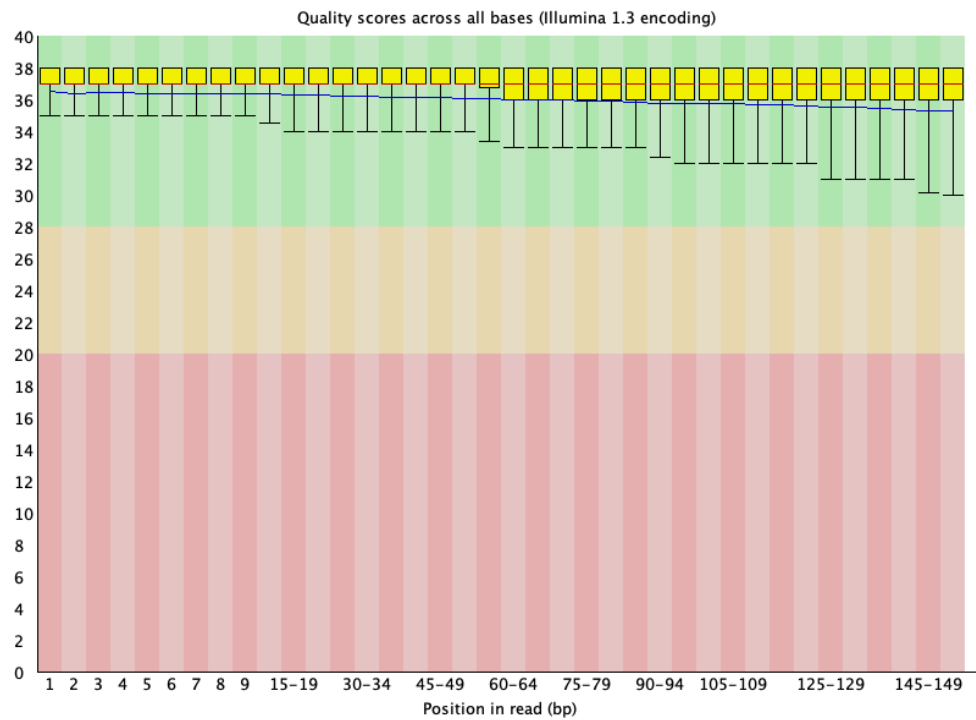
Wilke, C. (2024). *cowplot: Streamlined plot theme and plot annotations for 'ggplot2'* (R package version 1.1.3). <https://wilkelab.org/cowplot/>

Zheng, R., Wang, C., Liu, R., Cai, R., & Sun, C. (2024). Physiological and metabolic insights into the first cultured anaerobic representative of deep-sea *Planctomycetes* bacteria. *eLife*, 12, RP89874. <https://doi.org/10.7554/eLife.89874>

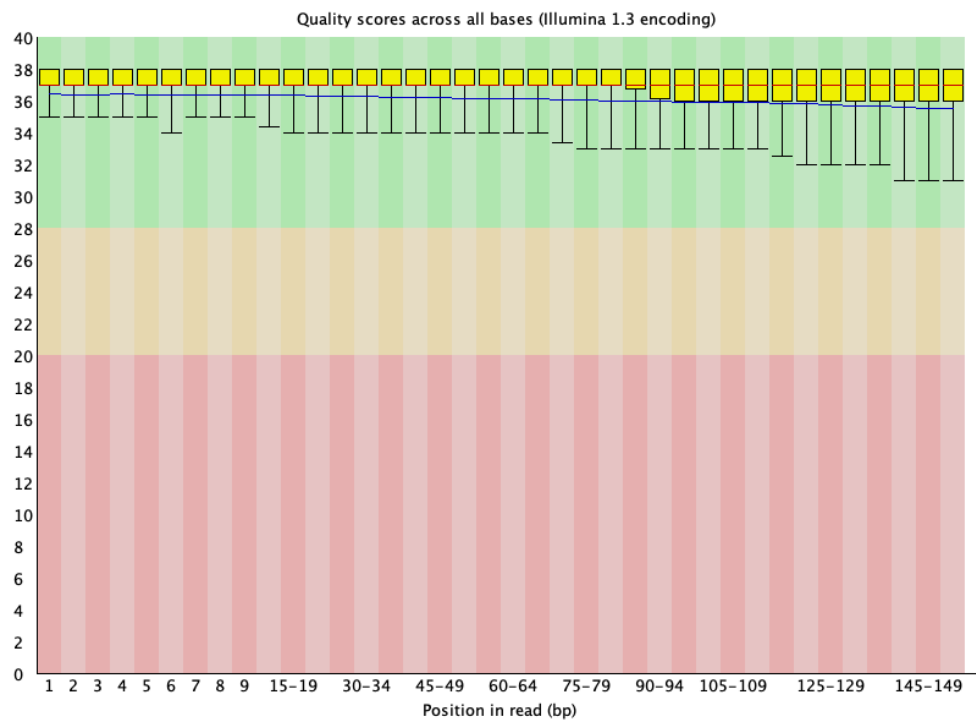
Supplementary files



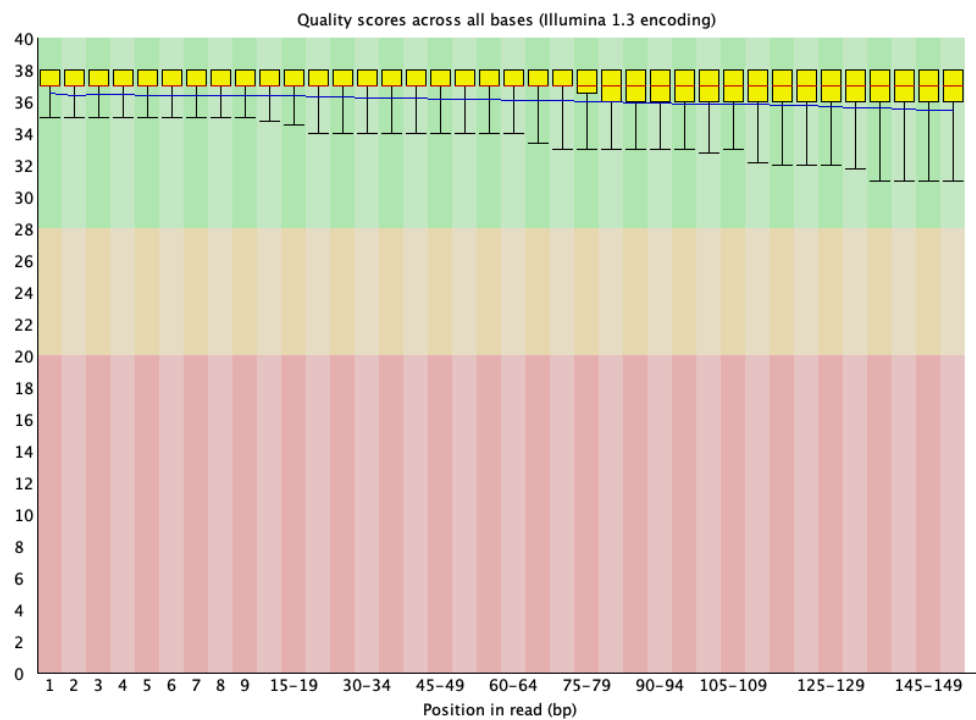
Supplementary figure 1. *Per base sequence quality for Winter PE22_2 sample.*



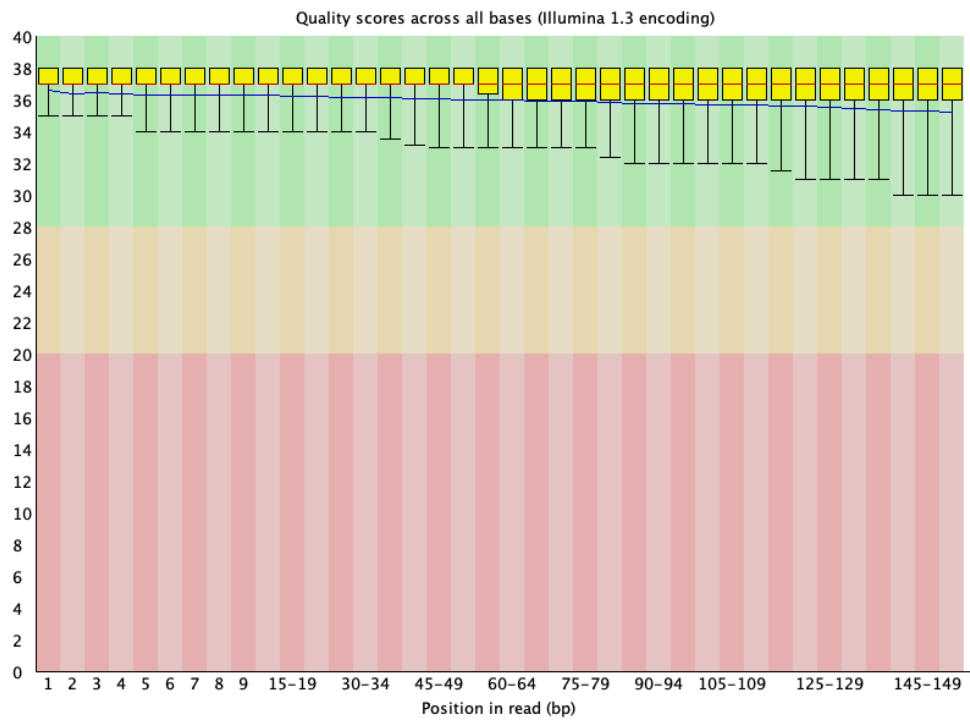
Supplementary figure 2. *Per base sequence quality for Winter PE22_3 sample.*



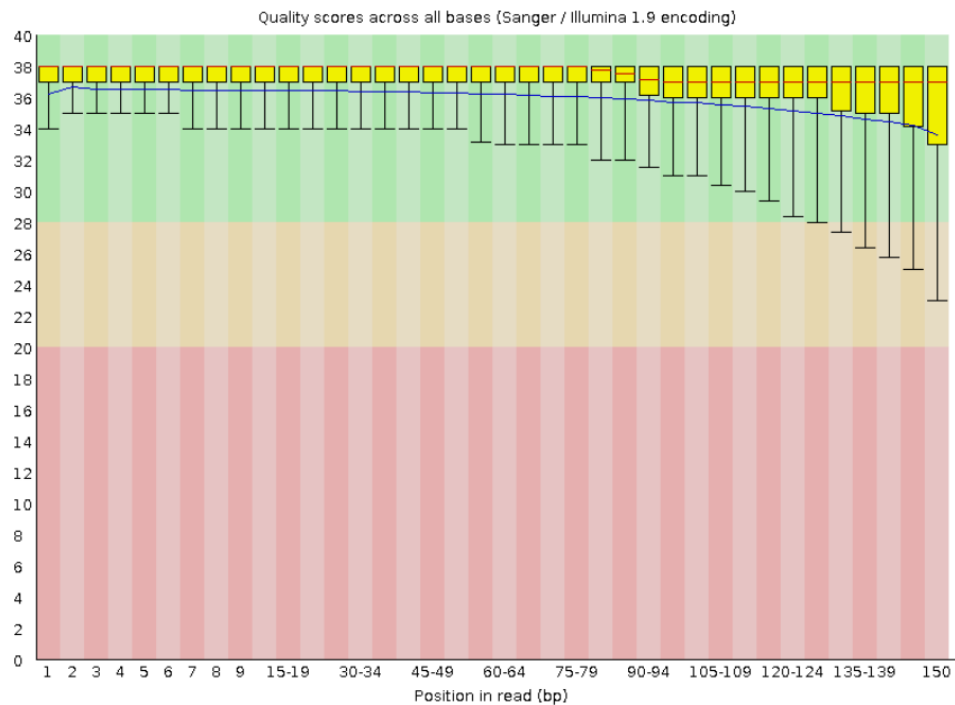
Supplementary figure 3. *Per base sequence quality for Winter PE22_9 sample.*



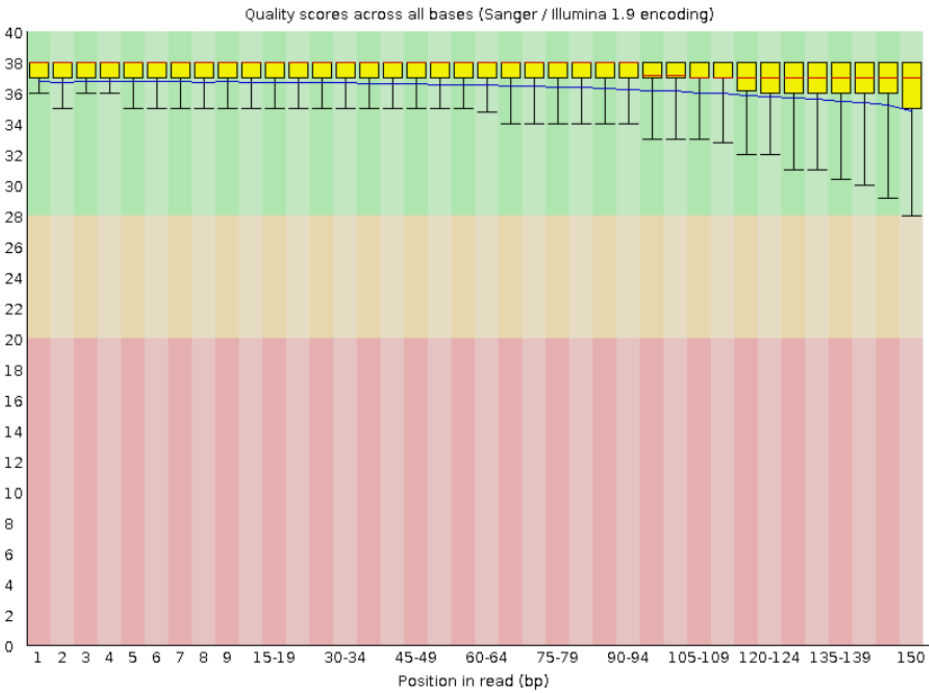
Supplementary figure 4. *Per base sequence quality for Winter PE22_10 sample.*



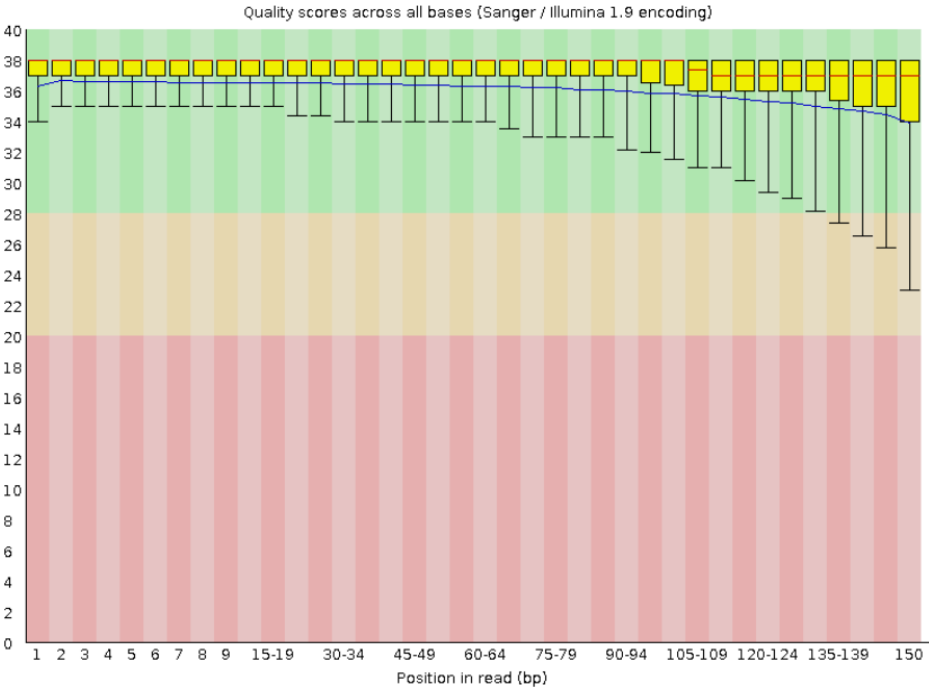
Supplementary figure 5. Per base sequence quality for Winter PE22_11 sample.



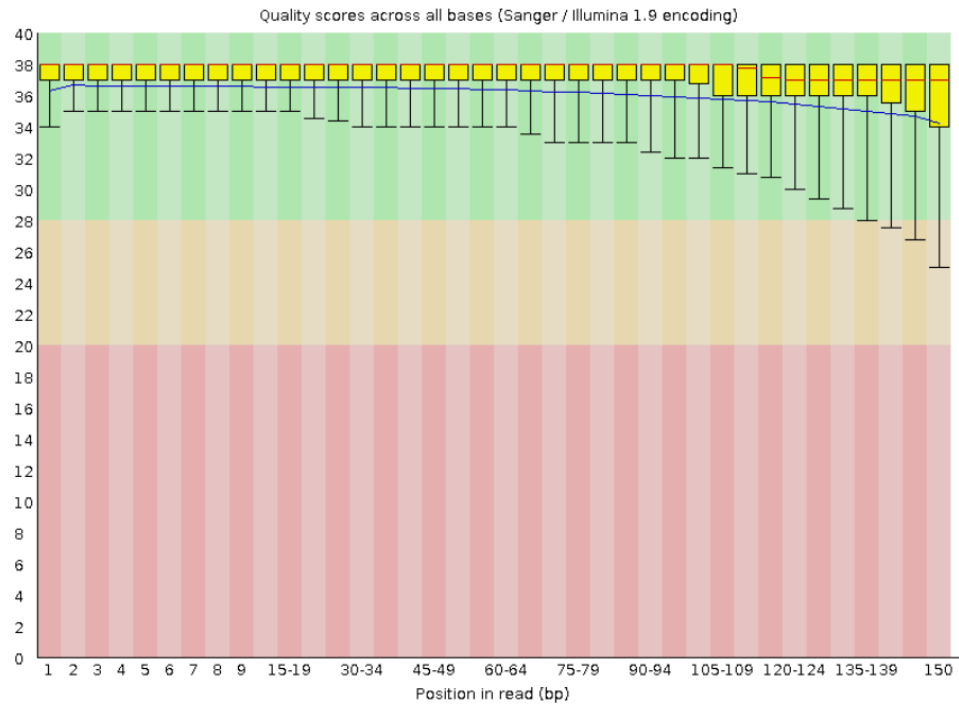
Supplementary figure 6. Per base sequence quality for Summer PE22_29 sample.



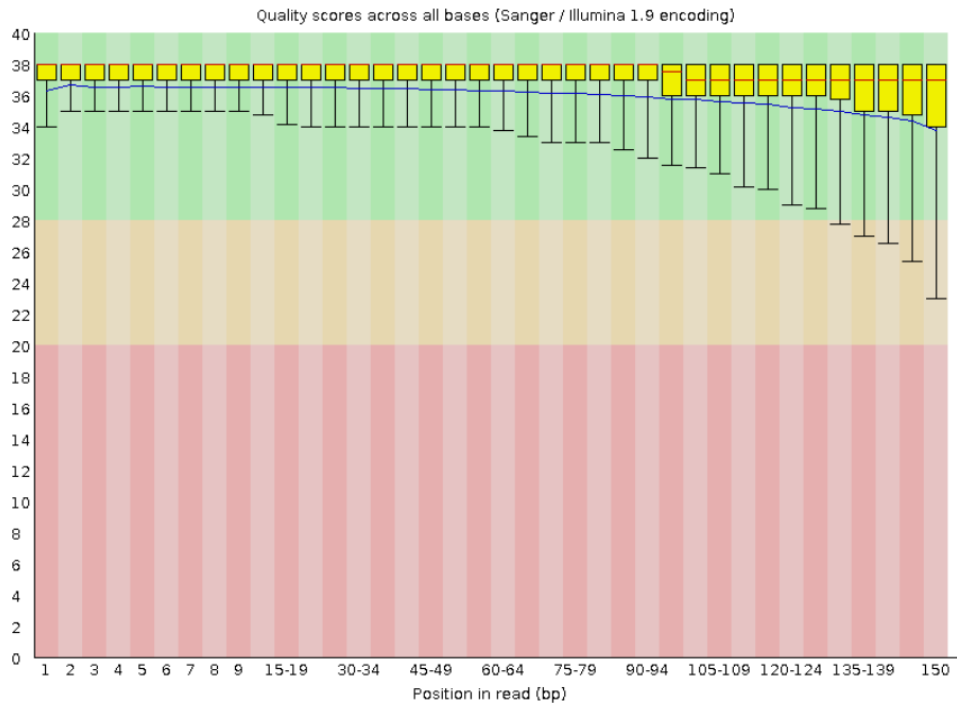
Supplementary figure 7. Per base sequence quality for Summer PE22_31 sample.



Supplementary figure 8. Per base sequence quality for Summer PE22_39 sample.



Supplementary figure 9. Per base sequence quality for Summer PE22_41 sample.



Supplementary figure 10. Per base sequence quality for Summer PE22_43 sample.