



**LA VIE DANS LA GLACE :
Adaptations génomiques des psychrophiles dans la glace du lac le plus nordique du
Canada**

Maxime Larose

Mémoire présenté à l'Université du Québec à Chicoutimi en vue de l'obtention du grade de
Maître ès sciences (M. Sc.) en ressources renouvelables

Québec, Canada

© Maxime Larose, 2024

RÉSUMÉ

La glace de lac isole les lacs qui en sont couverts du reste de leur environnement. Le réchauffement climatique a perturbé la stabilité de la glace pérenne sur les lacs d'eau douce en Arctique, comme en témoignent les premiers événements de fonte complète enregistrés depuis 2011 sur le lac le plus nordique du Canada, le lac Ward Hunt. Les assemblages microbiens dans ces habitats autrefois stables sont donc particulièrement vulnérables au réchauffement futur dans l'Arctique. Pourtant, à ce jour, la glace de lac est un compartiment peu étudié de la cryosphère. Le séquençage par amplicon du gène 16S de l'ARNr a révélé la présence de microbes diversifiés et actifs dans la glace du lac Ward Hunt, et un séquençage métagénomique profond produisant des génomes assemblés à partir de métagénomes a été utilisé pour étudier leur potentiel fonctionnel, c'est-à-dire comprendre les adaptations utilisées par ces organismes pour vivre dans cet habitat précaire. Nous avons constaté que les abondances relatives des microorganismes et des gènes variaient selon la couche de glace (glace en surface, au centre et en contact avec la colonne d'eau) et étaient distinctes de l'eau du lac sous-jacente. Aussi, les fonctions associées à la psychrophilie (telles que la résistance au froid) ont été détectées en plus grand nombre dans la glace de surface. La combinaison de génomes assemblés par métagénomes et les transcrits du gène 16S nous a permis d'identifier des fonctions potentiellement activement exprimées dans la glace du lac au moment de l'échantillonnage. Cette étude montre la diversité du microbiome de la glace d'eau douce d'un lac du Haut Arctique, dont la couverture de glace est vulnérable à un réchauffement soutenu, ce qui impliquerait une perte d'habitat pour ces microbes uniques et leurs adaptations au froid.

ABSTRACT

Warming has disrupted the stability of perennial freshwater ice, as evidenced by the first recorded ice off events that have occurred since 2011 on Ward Hunt Lake (WHL), the northernmost lake in Canada. Microbial assemblages in these once-stable habitats are therefore especially vulnerable to further warming in the Arctic, yet lake ice is an understudied compartment of the cryosphere. Amplicon sequencing of the 16S rRNA gene revealed the presence of diverse and active microbes in WHL ice, and deep shotgun metagenomic sequencing producing metagenome-assembled genomes was used to investigate their functional potential. We found that the relative abundances of microbes and genes varied by ice layer (surface ice, middle ice and bottom ice in contact with the water column) and were distinct from underlying lake water, and that psychrophily-associated functions (such as cold-shock) were found in greater counts in surface ice. Combining metagenome-assembled genomes and 16S transcripts allowed us to identify functions potentially actively expressed in lake ice at time of sampling. This study shows the diversity of the freshwater ice microbiome of a High Arctic lake, whose ice cover is vulnerable to further warming, implying habitat loss for these unique microbes and their adaptations to cold.

TABLE DES MATIÈRES

RÉSUMÉ	ii
ABSTRACT.....	iii
TABLE DES MATIÈRES	iv
LISTE DES TABLEAUX	v
LISTE DES FIGURES	vi
LISTE DES ABRÉVIATIONS	viii
REMERCIEMENTS.....	ix
AVANT-PROPOS.....	x
CHAPITRE 1: Introduction générale	1
CHAPITRE 2: LIFE IN ICE: Genomic Adaptations of Psychrophiles in the Ice of Canada's Northernmost Lake	8
2.1 INTRODUCTION	8
2.2 MATERIAL AND METHODS.....	12
2.2.1 Sample Collection.....	12
2.2.2 Sample processing	14
2.2.3 Extraction and sequencing.....	14
2.2.4 Read processing.....	15
2.2.5 Psychrophilic trait annotation	16
2.2.6 Binning of metagenome-assembled genomes (MAGs)	17
2.2.7 Statistical methods.....	17
2.2.8 Data availability.....	18
2.3 RESULTS.....	18
2.3.1 Taxonomic annotation	18
2.3.2 Functional potential	24
2.4 DISCUSSION.....	30
CHAPITRE 3: Tests préliminaires de décontamination de carottes de glace	41
3.1 INTRODUCTION	41
3.2 MATÉRIEL ET MÉTHODES	44
3.3 RÉSULTATS ET DISCUSSION	47
3.4 CONCLUSION	48
CHAPITRE 4: Conclusion générale	50
BIBLIOGRAPHIE.....	56
ANNEXE 1: Supplementary material (Chapitre 2)	64
ANNEXE 2: Quantification de l'ADN dans la glace de lacs boréaux et arctique.....	67

LISTE DES TABLEAUX

- Table 1.** Metagenome-assembled genomes (MAGs) from the co-assembly, ordered by expected completion of the genome (%).27
- Tableau 2.** Résultats des décontaminations pour les solutions contrôle, hypochlorite de sodium 5% et éthanol 95% par bain et par vaporisation. La limite de détection pour la quantification de l'ADN selon la trousse utilisée est de 0.005 ng/µl, pour un minimum de 0.1 ng d'ADN dans l'échantillon. Les valeurs non détectables (en deçà de ce seuil) ont été indiquées comme telles dans le tableau.47
- Table S1.** Mean length of sequenced reads, the number of forward (R1) and reverse (R2) reads per sample and the number of contigs obtained after assembly using Megahit. The co-assembly uses the total reads from all eight samples..64
- Table S2.** Total number of raw reads per amplicon sequencing samples and the resulting number of reads before and after DADA2 processing.64
- Table S3.** Psychrophilic genes, consensus sequence lengths, HMMER expected matches and whether BLASTx correctly identified our best HMMER hits or not. "NoMatch" means HMMER did not find the gene in our contigs.65
- Tableau S4.** Quantification de l'ADN dans les échantillons après extraction pour trois lacs boréaux et le lac Ward Hunt en Arctique. La limite de détection pour la quantification de l'ADN selon la trousse utilisée est de 0.005 ng/µl, pour un minimum de 0.1 ng d'ADN dans l'échantillon. Les valeurs non détectables (en deçà de ce seuil) ont été indiquées comme telles dans le tableau.68

LISTE DES FIGURES

- Figure 1.** Location of WHL in the Canadian High Arctic. **A.** Location of Ward Hunt Island (WHI), Nunavut. **B.** Location of WHI at the mouth of the Disraeli fiord (northern Ellesmere Island). **C.** Photograph of WHL in July 2022. **D.** An ice core divided into three parts (top, middle, bottom) following extraction with Kovacs Mark II system (coring barrel on the right).....13
- Figure 2.** Relative abundance of bacteria at the phylum-level in lake ice and the water column under the ice from triplicate ice cores for **A.** shotgun metagenomic sequencing, **B.** amplicon sequencing of the 16S rDNA gene and **C.** amplicon sequencing of 16S rRNA transcripts (cDNA). Gaps show samples for which sequencing yielded too few reads and were removed, or that were not sequenced (**Table S1**). Phyla representing less than 0.07% (from shotgun sequencing) and 0.01% (from amplicon sequencing) of relative abundance across all samples were combined into “Others” for visualization purposes. Organisms grouped in “(Bacteria)” are bacterial organisms for whom phylum assignment could not be resolved.19
- Figure 3.** Relative abundance of bacterial genera in lake ice and water from under the ice from triplicate ice cores for **A.** shotgun metagenomic sequencing, **B.** amplicon sequencing of the 16S rDNA gene and **C.** amplicon sequencing of 16S rRNA transcripts (cDNA). The twelve most abundant genera across all samples are shown, while the others were combined into “Other”.23
- Figure 4.** Principal coordinate analysis from Bray-Curtis dissimilarity index showing bacterial community patterns of across ice layers and water for **A.** 16S amplicon rDNA sequencing and **B.** 16S amplicon cDNA sequencing.24
- Figure 5.** Heatmap of normalized abundance of psychrophily-associated genes in ice core replicates identified with HMMER for three groups of functional adaptations: solute uptake (red), protection against toxic oxidative reaction products (brown), and protection against cold denaturation (black). Reads were normalized per sample, dendrograms show distance-based clusters of similar samples (right) and genes (top). The Z-score is a measure of how many standard deviations there are between the abundance of a gene in a sample and the mean abundance of that gene across all samples.....26
- Figure 6.** **A.** Distribution of mean coverage of high completion MAGs in ice layers and water (number of hits/total number of contigs) **B.** Presence/absence of psychrophily-associated genes (presented in **Figure 5** and **Table S3**) detected in high-completion MAGs.....28

Figure 7. Morceaux de carotte de glace de lac après 60 secondes d’immersion dans une solution 40% éthanol et colorant alimentaire vert. **A.** Glace blanche de surface et **B.** glace noire à l’interface de la colonne d’eau, après immersion. Aucune trace de coloration n’est visible dans les carottes, indiquant que les solutions de décontamination n’ont pas pénétré l’intérieur des carottes.45

Figure S1. Relative abundance of *Commamonadaceae* in ice and water below ice from triplicate ice cores for **A.** shotgun metagenomic sequencing, **B.** amplicon sequencing of the 16S rDNA gene and **C.** amplicon sequencing of 16S rRNA transcripts (cDNA). Gaps show samples for which sequencing yielded too few reads (**Table S1**). Genera representing less than 0.01% (shotgun sequencing) and 0.01% (amplicon sequencing) of relative abundance were combined into “Others”.66

LISTE DES ABRÉVIATIONS

ADN / DNA	Acide désoxyribonucléique / <i>Deoxyribonucleic acid</i>
ADNc / cDNA	ADN complémentaire / <i>Complementary DNA</i>
ADNr / rDNA	ADN ribosomique / <i>Ribosomal DNA</i>
ARN / RNA	Acide ribonucléique / <i>Ribonucleic acid</i>
ARNr / rRNA	ARN ribosomique / <i>Ribosomal RNA</i>
LIA	<i>Last Ice Area</i>
MAG	Génomes assemblés à partir de métagénomes / <i>Metagenome-assembled genome</i>
Mb	<i>Megabit</i>
NCBI	<i>National Center for Biotechnology Information</i>
PERMANOVA	Analyse de variance multivariée permutacionnelle / <i>Permutational multivariate analysis of variance</i>
PCoA	Analyse en coordonnées principales / <i>Principal Coordinate Analysis</i>
WHI	Île Ward Hunt / <i>Ward Hunt Island</i>
WHL	Lac Ward Hunt / <i>Ward Hunt Lake</i>
Zmax	Profondeur maximale

REMERCIEMENTS

Je tiens à exprimer toute ma reconnaissance à ma directrice de recherche, Dr. Catherine Girard pour son soutien indéfectible. Ses précieux conseils, sa grande disponibilité et sa générosité ont été instrumentaux dans la complétion de ce travail de recherche. Je tiens également à remercier Catherine pour les nombreuses opportunités d'apprentissage et de développement qui m'ont été offertes tout au long de ce projet.

J'aimerais aussi remercier les individus et laboratoires qui ont contribué de façon significative à mon projet : Pénélope Blackburn Desbiens, Alexander I. Culley, Warwick F. Vincent, le laboratoire Gépromic de l'UQAC, l'Aquatic Laboratory de l'UQAC, Milla Rautio, Noémie Gaudreault, Crysta Rhains, Claire Fournier et Pascal Tremblay.

Enfin, je dois souligner le support financier et logistique des organismes suivants : Centre d'études nordiques (CEN), Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), des Fonds de recherche Nature et technologies du Québec (FRQNT), Programme du plateau continental polaire (PPCP), Programme de formation scientifique dans le Nord (PFSN), Fonds Nouvelles Frontières en recherche (FNFR), Sentinelle Nord, Hamlet of Resolute Bay, Parcs Canada, ArcticNet, Calcul Québec, Nunavut Research Institute, Groupe de Recherche Interuniversitaire en Limnologie (GRIL).

AVANT-PROPOS

Les microorganismes dominent les courts réseaux trophiques aquatiques de l'Arctique (Vincent et al., 2011), et malgré leurs conditions froides et oligotrophes combinées avec une saisonnalité extrême, les lacs polaires abritent des microbiomes diversifiés (Cavaco et al., 2019 ; Comte et al., 2018 ; Girard et al., 2023). De plus, la glace de lac elle-même peut constituer un habitat microbien, car des organismes microbiens aquatiques peuvent y être incorporées pendant la période de gel (Santibáñez et al., 2019). La glace de lac est un compartiment sous-estimé de la cryosphère étant donné sa structure plus simple que la glace de mer - pourtant, il a été démontré qu'elle abrite des communautés extrémophiles (Bondarenko et al., 2012), et qu'elle peut soutenir de la production primaire, comme démontré dans un lac boréal au printemps (Kivilä et al., 2023), ce qui suggère que les organismes de la glace de lac peuvent être actifs dans certaines conditions. La vie dans la cryosphère est dominée par des microbes psychrophiles (Bourquin et al., 2022), et une grande partie de ce que nous savons de ceux-ci provient d'études sur les glaciers, la glace de mer, la neige et le pergélisol. On sait cependant toujours peu de choses sur l'identité ou l'activité des microbes dans la glace des lacs, qu'elle soit saisonnière ou pérenne. Avec le réchauffement rapide du Nord et l'attrition de la glace dans cette région, il est urgent de documenter la diversité méconnue des communautés microbiennes de la glace de lac arctique.

Ce mémoire présente la première caractérisation microbienne de la glace du lac le plus septentrional du Canada, et est divisé en quatre chapitres. Le premier chapitre est une introduction générale qui met en place les connaissances existantes sur la microbiologie des lacs polaires et de la cryosphère. Le second chapitre, qui porte le même titre que le présent mémoire, présente l'étude des communautés de la glace du lac Ward Hunt (Nunavut) sous forme d'article scientifique, il est donc rédigé en anglais en vue d'une soumission à un journal scientifique pour révision par les pairs. Le troisième chapitre expose les résultats de tests préliminaires de décontamination de carottes de glaces qui ont informé le design expérimental de ce projet. Le quatrième chapitre est une conclusion générale. En annexe se trouvent des données supplémentaires qui seront jointes à l'article scientifique lors de sa soumission ainsi que des données expérimentales sur la glace de lacs boréaux, comparés au lac Ward Hunt.

CHAPITRE 1: Introduction générale

On compte environ 117 millions de lacs sur terre (Verpoorter et al., 2014) dont la moitié sont situés au nord du 45° parallèle (Campbell et al., 2005; Hampton et al., 2017). À l'échelle globale, 86% de tous les lacs subissent une période de gel en surface annuellement (Korver et al., 2024). Le couvert de glace affecte le transfert d'énergie, de chaleur, de lumière et de matière entre le lac et son environnement en isolant la colonne d'eau d'échanges avec l'atmosphère, d'apports via les écoulements dans le paysage et en protégeant les eaux de surface du brassage par le vent (Cavaliere et al., 2021). Les conditions dans la colonne d'eau d'un lac couvert de glace sont donc fortement contrastées à celles d'un lac sans couvert. On sait aussi que la composition chimique et biologique de la glace diffère de celle de la colonne d'eau. Les particules y sont retenues, mais les composés dissous en sont quant à eux exclus (Imbeau et al., 2021). On peut donc attribuer à la glace une fonction de stockage de composés labiles qui sont rendus disponibles aux organismes de la colonne d'eau lors de la période de fonte. C'est donc dire que la glace détient un rôle prédominant dans le cycle biochimique des écosystèmes qui en sont recouverts (Imbeau et al., 2021).

Selon la topographie et la bathymétrie du lac, la présence d'affluents, la latitude, le climat et l'année, la durée du couvert de glace peut varier de quelques jours à toute l'année, et son épaisseur de moins d'un centimètre à plusieurs mètres (Cavaliere et al., 2021). L'opacité de la glace peut aussi grandement varier. En Arctique, la glace de lac est généralement formée d'eau de lac gelée (glace noire) qui est très transparente, alors qu'en zones plus tempérées, les précipitations sous forme de neige recouvrent la glace

de lac et s'y lient pour former une couche de glace plus opaque, appelée glace blanche ou glace de neige, qui superpose la glace noire (Leppäranta, 2010). Cette glace blanche limite énormément le passage de la lumière vers les couches inférieures de la glace et vers la colonne d'eau alors que la glace noire, formée à l'interface de la colonne d'eau, laisse davantage pénétrer les rayons du soleil (Imbeau et al., 2021).

Si on a longtemps cru que les lacs couverts de glace étaient en dormance et que l'activité biologique sous cette glace était limitée, un intérêt grandissant de la communauté scientifique pour la limnologie des régions froides nous montre que la situation est toute autre (Hampton et al., 2017). En effet, des études ont démontré que les communautés microbiennes aquatiques persistent sous la glace, et varient en abondance et en diversité dans la colonne d'eau au fil des saisons; dans certains cas, la richesse spécifique de la colonne d'eau pourrait même augmenter sous le couvert de glace (Butler et al., 2019). Ces communautés sont soutenues sous la glace, qu'il s'agisse de lacs à couvert saisonnier (Tran et al., 2018) ou pérenne (Comte et al., 2018). Mais il n'y a pas que la colonne d'eau qui sert d'habitat; la glace constitue un habitat en soi pour des microorganismes actifs et adaptés à ses conditions extrêmes (Santibáñez et al., 2019). Les conditions environnementales spécifiques à la glace comme le froid et l'oligotrophie (Collins and Margesin, 2019) impliquent des adaptations uniques de la vie microbienne psychrophile (Anesio and Bellas, 2011; Ruuskanen et al., 2020), qui est définie par un optimum de croissance variant entre des températures sous le point de congélation jusqu'à environ 10° C. On s'attend donc à trouver dans la glace des communautés similaires à celles de la colonne d'eau, mais adaptées aux conditions difficiles de la cryosphère.

Les adaptations susmentionnées sont le résultat de la combinaison de plusieurs fonctions biologiques. Par exemple, les basses températures inhibent normalement les réactions chimiques et enzymatiques. Pour maintenir leurs réactions catalysées par des enzymes, les organismes doivent opérer une translation de l'optimum d'activité enzymatique. Ils y parviennent en déstabilisant la protéine ou son site actif, lui permettant d'être active à des températures qui ne le permettent normalement pas chez des organismes mésophiles (Feller and Gerday, 2003). D'autres adaptations incluent des fonctions qui permettent l'utilisation de lipides membranaires alternatifs (Hassan et al., 2020a) qui induisent une plus grande fluidité pour l'acquisition des nutriments (Raymond-Bouchard and Whyte, 2017).

Ces adaptations physiologiques sont encodées par le génome des microorganismes psychrophiles, qui partagent un bagage génomique chromosomique commun (*core genome*) plus restreint que d'autres organismes mésophiles, et dépendent davantage d'éléments génétiques mobiles (« mobilome ») pour répondre aux conditions de leur environnement (Vigneron et al., 2022). Ces éléments mobiles incluent les plasmides, qui sont des morceaux circulaires de code génétique que les microorganismes peuvent se partager par transfert horizontal (ou latéral), d'un individu à un autre, plutôt que par transfert vertical, d'un parent vers son descendant (Ciok et al., 2018). Les plasmides confèrent typiquement un avantage sélectif à son porteur, et permettent la propagation d'adaptations ou de résistances (notamment aux antibiotiques, aux métaux lourds) dans les populations bactériennes. La résultante de tels transferts est la spécialisation locale des microorganismes, qui partagent davantage de fonctions métaboliques avec des individus de groupes phylogénétiques éloignés,

mais évoluant dans le même milieu qu'eux, et qui s'échangent donc les mêmes éléments génétiques (Vigneron et al., 2022). Ainsi, les conditions difficiles de la glace favorisent ces échanges de matériel génétique et contribuent à la vision de la cryosphère comme étant un « point chaud » de l'évolution microbienne (Anesio and Bellas, 2011). Ces communautés sont d'un intérêt important en Arctique, puisque bon nombre de lacs y sont couverts de glace toute l'année, et ce depuis des décennies – les communautés qui y vivent ont donc pu s'y spécialiser au fil du temps, à un degré beaucoup plus important que dans les lacs à couvert saisonnier, dont la glace disparaît chaque année. Malgré cette spécialisation locale et l'instabilité croissante de cet habitat (Vincent, 2010), les communautés de la glace de lac demeurent peu connues (Anesio and Laybourn-Parry, 2012).

La durée moyenne du couvert de glace s'est vue réduite d'environ 12 jours au cours du siècle dernier en raison des changements climatiques (Magnuson et al., 2000); une réduction du nombre de lacs à la surface desquels un couvert de glace saisonnier se forme est donc inévitable (Sharma et al., 2019). Ces changements, qui peuvent également provoquer de l'intermittence ou une disparition complète des couverts de glace, risquent d'influencer considérablement les régimes thermiques et de lumière dans la colonne d'eau et d'avoir des répercussions lors des saisons suivantes (Cavaliere et al., 2021). Dans le Haut Arctique, où les lacs demeurent normalement continuellement gelés, on constate une diminution de l'épaisseur et de l'étendue de la glace des lacs et cette dernière peut même parfois disparaître complètement en été (Bégin et al., 2021a; Lehnerr et al., 2018; Michelutti et al., 2020; Paquette et al., 2015). Lorsque la glace persiste, ses marges s'amincissent, ce qui entraîne une augmentation

de la pénétration lumineuse et une productivité primaire accrue en zone littorale (Bégin et al., 2021a). Poussée à son extrême, la disparition du couvert de glace de ces lacs pourrait augmenter le brassage et la turbidité de la colonne d'eau, limitant la productivité primaire et altérant le régime biologique de ces systèmes (Bégin et al., 2021a). La dégradation des couverts de glace pérennes risque donc d'avoir des conséquences profondes sur les lacs arctiques, qui sont majoritairement oligotrophes.

Le lac le plus nordique du Canada, le lac Ward Hunt, a vu l'épaisseur et la superficie de son couvert glace diminuer très rapidement à partir du milieu des années 2000 pour disparaître complètement à l'été 2011, puis à nouveau en 2012 (Paquette et al., 2015) et en 2016 (Bégin et al., 2021a). Des données historiques qui remontent aux années 1950 montrent que la glace se maintenait alors à environ quatre mètres d'épaisseur en été (Hattersley-Smith et al., 1955; Paquette et al., 2015), témoignant de l'ampleur de l'attrition de la glace sur ce lac. Si le couvert s'est maintenu depuis l'automne 2016, son épaisseur en juillet 2022 n'était en moyenne que de 1,52 m.

Cette nouvelle réalité de la précarité de la glace sur les lacs à hautes latitudes risque de rapprocher fonctionnellement la glace des lacs arctiques de celle des lacs boréaux à couvert de glace saisonnier, et il est possible que les adaptations uniques permettant aux organismes de vivre dans la glace pérenne disparaissent avec ces dernières. En effet, il y a un risque d'observer une diminution des adaptations fonctionnelles aux conditions de la glace, au profit de communautés plus généralistes issues de la colonne d'eau. Nous pourrions alors assister à une homogénéisation latitudinale des fonctions microbiennes de la glace. Malgré un nombre croissant d'études en limnologie hivernale, peu d'études s'intéressent à la glace de lac comme

habitat. Pourtant, en considérant le nombre élevé de lacs qui gèlent et les transformations importantes que subissent ces écosystèmes, ainsi que le risque de perte de fonctions associé, il apparaît important de mieux comprendre quels sont les impacts de ces changements sur les communautés microbiennes de la glace de lacs arctiques.

L'objectif principal du présent projet était de comparer les communautés microbiennes et les gènes qui leur permettent de se développer dans les conditions extrêmes de la glace de lac. Ce projet marque, à notre connaissance, le premier séquençage fonctionnel de la glace de lac arctique. Il nous apparaissait nécessaire de caractériser cette glace multiannuelle sachant que le site d'étude a subi de grandes perturbations au cours de 20 dernières années, incluant les premières fontes complètes de son couvert de glace, mais aussi la rupture de la plate-forme de glace Ward Hunt située à proximité, exposant davantage le lac à la déposition atmosphérique d'aérosols issus de l'Océan Arctique ouvert (Vincent et al., 2011).

Les objectifs spécifiques du projet étaient de caractériser les communautés de la glace de lac par trois approches de séquençage complémentaires : la première, le séquençage métagénomique de l'ADN, qui consiste à séquencer tout l'ADN d'un échantillon, permet d'identifier les membres de communautés microbiennes ainsi que leurs fonctions métaboliques. La seconde approche, par amplicon sur l'ADNr consiste à amplifier un marqueur taxonomique pour produire une description plus précise de la composition taxonomique des communautés des échantillons. Pour la troisième méthode de séquençage, par amplicon également, c'est l'ARNr qui est séquençé dans le but d'identifier quels organismes produisent des transcrits d'ARN, et sont donc actifs au moment de l'échantillonnage. Ensemble, ces trois méthodes permettent d'identifier

les gènes d'adaptation portés par les microorganismes de la glace de lac, tout en essayant de comprendre leur distribution entre les différentes couches de glace et comment la glace se distingue de l'eau qu'elle recouvre. Nous avons choisi une approche en biologie moléculaire basée sur la génomique, car les microorganismes environnementaux sont difficiles à mettre en culture et donc à étudier par les approches traditionnelles de microbiologie (Yarza et al., 2014). Les approches de séquençage par amplicon seules permettent d'identifier les groupes connus, mais ne nous renseignent pas sur les gènes qu'ils portent, au contraire de l'approche métagénomique qui permet l'identification des fonctions métaboliques des organismes, que l'on connaisse l'identité taxonomique de ces groupes ou non. La combinaison de l'approche par amplicon de l'ARN, qui renseigne sur l'activité métabolique, et métagénomique, qui renseigne sur le potentiel fonctionnel, nous permettra d'inférer quelles fonctions détectées dans la glace ont le potentiel d'être active au moment de l'échantillonnage.

CHAPITRE 2: LIFE IN ICE: Genomic Adaptations of Psychrophiles in the Ice of Canada's Northernmost Lake

2.1 INTRODUCTION

There are approximately 117 million lakes on Earth (Verpoorter et al., 2014) and approximately 50% of those are located north of the 45° parallel and therefore experience a period of surface freezing during the year (Campbell et al., 2005; Hampton et al., 2017), lasting up to 12 months in the High Arctic (Imbeau et al., 2021). Lake ice affects the transfer of energy, heat, light and matter between the environment and the lake by insulating the water column from the atmosphere, landscape and mixing by wind (Cavaliere et al., 2021). Conditions in the water column of an ice-covered lake therefore contrast sharply with those of an ice-free lake and may vary seasonally depending on duration of ice cover. Seasonal variations in terms of ice cover are the lowest in high latitudes lakes that have perennial or multi-annual ice cover. The chemical and biological composition of ice differs from that of the water column; particles are retained, but dissolved compounds are excluded (Imbeau et al., 2021; Zhou et al., 2023). Ice therefore has a function of storing labile compounds which are made available to organisms in the water column during the melting period, thus playing a predominant role in lake biogeochemistry (Imbeau et al., 2021).

In the Arctic, lake ice is overall more transparent than at temperate latitudes due to low precipitations, leading to a thinner snow-originating layer (white ice), which is opaque and limits the passage of light (Leppäranta, 2010). Indeed, High Arctic lakes are mostly covered by black ice (Robinson et al., 2021) formed of frozen lake water,

which is highly transparent and allows greater light penetration (Imbeau et al., 2021). Although the thickness of multi-year ice can limit light transmission (Katlein et al., 2015), the water column of ice-capped polar lakes still receives enough light to support photosynthesis during the polar summer (Mohit et al., 2017), especially in snow-free conditions (Imbeau et al., 2021) like those of polar deserts where precipitations are low.

Microorganisms dominate the truncated aquatic food webs of the Arctic (Vincent et al., 2011), and although they experience cold oligotrophic conditions with extreme seasonality, polar lakes harbour diverse microbiomes (Cavaco et al., 2019; Comte et al., 2018; Girard et al., 2023). Furthermore, lake ice itself can constitute a microbial habitat, as cells can be incorporated from lake water into ice during freezing (Santibáñez et al., 2019). Lake ice is an overlooked compartment of the cryosphere, given its simpler structure than sea ice - still, it has been found to harbour extremophile communities (Bondarenko et al., 2012), and can host primary production, as demonstrated in a boreal lake in the spring (Kivilä et al., 2023), suggesting lake ice organisms may be active under certain conditions. Life in the global cryosphere is dominated by cold-dwelling microbes (Bourquin et al., 2022), and much of what we know of these psychrophiles comes from studies on glaciers, sea ice, snow, and permafrost; little is known on the identity or activity of microbes in lake ice, be it seasonal or perennial.

Conditions in multi-year lake ice such as low temperatures, oligotrophy, osmotic stress and exposure to UV radiation (Collins and Margesin, 2019) likely necessitate adaptations of psychrophilic microbial life, similar to those that have been documented in other compartments of the cryosphere (Anesio and Bellas, 2011; Ruuskanen et al.,

2020). These adaptations may include altered cellular components, such as enzymes modified for lower temperatures (Feller, 2003), fatty acids (Hassan et al., 2020b), and antifreeze proteins (Collins and Margesin, 2019) to promote membrane fluidity, pigments for protection from damaging UV radiation (Halbach et al., 2022) and high-efficiency nutrient uptake systems to compete in this nutrient-depleted habitat (Edwards et al., 2013). Psychrophilic microorganisms are characterized by a common genomic background (core genome) that is more restricted than other mesophilic organisms and depends more on mobile genetic elements (“mobilome”) to respond to their environment (Liu et al., 2022; Vigneron et al., 2022), granting them greater genetic flexibility to endure harsher conditions. These include plasmids shared horizontally between microbes, like within the *Polaromonas* genus, a psychrophile ubiquitous in the cryosphere which relies on plasmids encoding DNA repair enzymes as well as enhanced transport and energy production systems (Ciok et al., 2018). The result of such transfers is the local specialization of microbes, which may share more metabolic functions with individuals from phylogenetically distant groups evolving in the same environment as them and therefore exchange the same genetic elements (Girard et al., 2023; Vigneron et al., 2022). Viruses also appear to be drivers of genetic diversity through horizontal gene transfer in the cryosphere (Anesio and Bellas, 2011; Dinsdale et al., 2008). In ice, where life consists mostly of microbes, viruses exert the major top-down control; viral-mediated exchange of genetic material may be increased, promoting diversification in hosts (Anesio and Bellas, 2011), and psychrophiles like *Polaromonas* carry CRISPR-Cas systems showing their means of protection from viral infection (Ciok et al., 2018). Microbial life in ice is therefore dynamic, with poorly studied communities of highly adapted psychrophiles, and their habitat is increasingly

unstable. Recently, Lake Hazen, the world's largest High Arctic lake, has seen its ice regime shift considerably, with open-water conditions now occurring in the summer (Lehnherr et al., 2018; Michelutti et al., 2020). Ward Hunt Lake, approximately ~150 km north-west of Hazen, has also exhibited sporadic summer ice-out conditions since 2011 (Paquette et al., 2015). This transition from perennial ice cover to multiyear and eventually seasonal cover of High Arctic lakes impacts water stratification (Paquette et al., 2015; Veillette et al., 2010), but also lead to the disappearance of a cryospheric habitat for lake ice microbes. It is therefore urgent to better document these communities.

Located on Ward Hunt Island, north of Ellesmere Island (83.087111, -74.145556), Ward Hunt Lake (WHL) is the northernmost lake in Canada. It is an ultraoligotrophic lake with a historically perennial ice cover that has seen the extent and thickness of its ice decrease rapidly from 2008 onward, going from a stable four metres thick measured from 1953 (Hattersley-Smith et al., 1955) to 2004, to a complete disappearance during the summer of 2011 (Paquette et al., 2015). Open water episodes followed during the summers of 2012 and 2016 (Bégin et al., 2021a; Paquette et al., 2015). These episodes of open water imply more turbidity, higher planktonic primary production and oxygen fluctuations in the following fall and winter (Bégin et al., 2021a), underscoring the profound role of ice cover on polar lake regimes. Previous microbial studies in Ward Hunt Island showed rich cyanobacteria-dominated mats thriving at the bottom of the lake under ice (Mohit et al., 2017) and high watershed connectivity of lacustrine communities to terrestrial water tracks (Comte et al., 2018).

Disappearing ice on Ward Hunt Lake could result in the loss of ice-adapted organisms and of their adaptive functions that may not survive in lake water following ice-off, as previous works have hinted at strong niche filtering leading to ecological sorting of microbes in high latitude lakes (Comte et al., 2018; Girard et al., 2023). Uncovering the communities of WHL ice and their specific genes is crucial at this transitory stage of ice cover duration and thickness. Here, we present a metagenomic characterization of the microbial communities of six-year-old multiyear ice on Ward Hunt Lake. Our goal was to contrast taxonomic and functional make-up of the microbiomes of lake ice across different strata, from surface (ice in contact with air) to bottom (ice in contact with the water column), and to compare it to under ice conditions in lake water.

2.2 MATERIAL AND METHODS

2.2.1 Sample Collection

Samples were collected in July 2022 on Ward Hunt Lake, Nunavut (83.087111, -74.145556) (**Figure 1**). Data from automated cameras show that the last ice-off event occurred in 2016, meaning the ice cover at the time of sampling was multi-year, and approximately six years old (NEIGE, 2023). At the time of sampling, ice thickness varied from 137 to 192 cm (mean 152 cm) across the lake and was characterized by a thin frazil layer (4-14 cm) overlaying black ice. Three ice core replicates (155 cm in length) were collected above the lake's maximum depth ($Z_{\max} = 9.6$ m). Cores were collected using the Mark II coring system (Kovacs Enterprise, Roseburg, OR, USA) and split into three layers on a clean surface with an 70% ethanol-washed saw, based on ice quality and for ease of processing. These layers, which were measured on site,

are hereafter referred to as “top” (in contact with air), “middle”, and “bottom” ice (in contact with the water column). To minimize exogenous contamination, sterile gloves were used for all sample handling. Samples were then melted in sterile Whirl-Paks® (Filtration Group, Pleasant Prairie, WI, USA), in the dark on the field. Surface water samples were also collected at Zmax using a Kemmerer Water Sampler (Wildco, Buffalo, NY, USA) and stored in 4 L Cubitainer containers that had previously been washed with 2% (vol/vol) Contrad™ 70 liquid detergent (DeconLabs, King of Prussia, PA, USA) and 10% (vol/vol) ACS-grade HCl (Sigma-Aldrich, St. Louis, MO, USA) and MilliQ water.

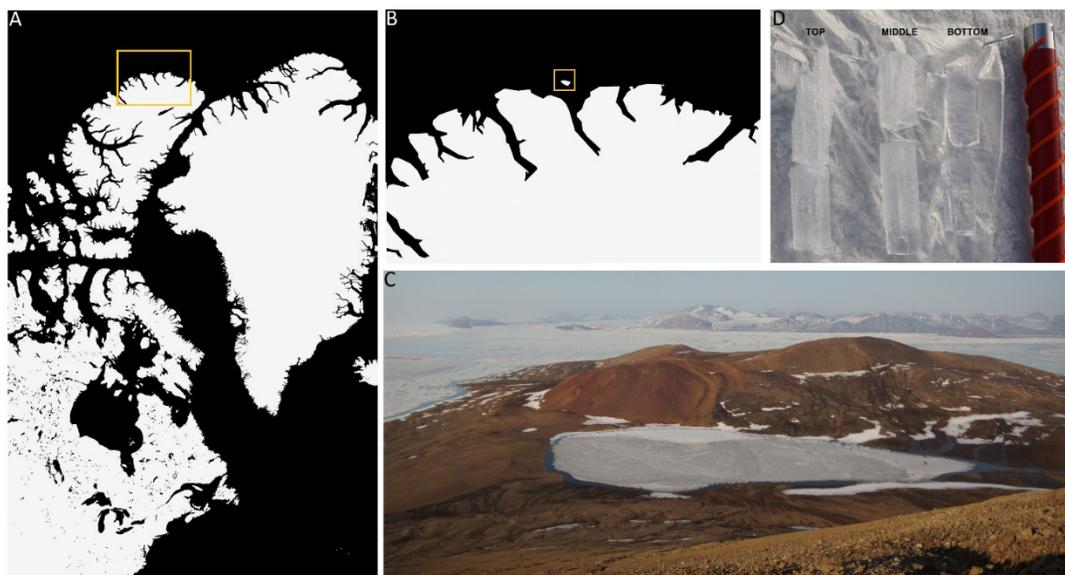


Figure 1. Location of WHL in the Canadian High Arctic. **A.** Location of Ward Hunt Island (WHI), Nunavut. **B.** Location of WHI at the mouth of the Disraeli fiord (northern Ellesmere Island). **C.** Photograph of WHL in July 2022. **D.** An ice core divided into three parts (top, middle, bottom) following extraction with Kovacs Mark II system (coring barrel on the right).

2.2.2 Sample processing

Melted ice and lake water were filtered onto Sterivex™ filters (0.22µm) (MilliporeSigma, Burlington MA, USA) using a Masterflex L/S peristaltic pump (Masterflex, Gelsenkirchen, Germany). Masterflex tubing was cleaned prior to filtrations by cycling Contrad 70 detergent and MilliQ water through the system. RNA-later® (Sigma-Aldrich) (2 ml) was used to protect nucleic acids until their extraction. Sterivex™ filters were then kept at -80 °C until departure from the field in an Arctic Express™ dry shipper (Thermo Fisher Scientific, Waltham, MA, USA). Once returned to the lab, they were kept at -80 °C.

2.2.3 Extraction and sequencing

DNA extractions were performed with the AllPrep DNA/RNA mini kit (QIAGEN, Hilden, Germany) as described in Cruaud et al., 2017. Library preparation and shotgun metagenomic sequencing were done at the Integrated Microbiome Resource (IMR, Dalhousie University, Halifax, NS, Canada) on an Illumina NextSeq2000 sequencer at 4X depth, which yielded an average of 1,367,398 reads/sample (**Table S1**). Reverse transcription of RNA to cDNA, library preparation (DNA & cDNA) and amplicon-based sequencing of the V3-V4 regions of the 16S rRNA marker gene were carried out at the *Plateforme d'analyses génomiques* of the Institute of integrative biology and systems (IBIS, Université Laval, Quebec, QC, Canada) on an Illumina MiSeq using primers 341F and 805R (Herlemann et al., 2011) as described in Blackburn-Desbiens et al., (*in prep*). DNA sequencing yielded

1,492,673 reads (with a sample average of 57,410) while cDNA sequencing yielded 1,829,249 reads (with a sample average of 70,355).

2.2.4 Read processing

Raw shotgun reads were trimmed for quality using Trimmomatic 0.39 (Bolger et al., 2014) using SLIDINGWINDOW:4:20 CROP:145 HEADCROP:15 AVGQUAL:20 and default for the rest of parameters. Reads were assembled by two approaches: single-sample assembly and co-assembly using Megahit 1.2.9 (Li et al., 2016), with only resulting contigs of 1000bp or longer being kept for downstream analysis (Alneberg et al., 2014). Amplicon-based reads were processed with DADA2 1.28.0 (Callahan et al., 2016) in R 4.3.0 (R Core Team, 2021). Primers and low-quality reads were trimmed using trimLeft=c(17,21), truncLen=c(280,260), maxN=0, maxEE=c(3,4), truncQ=2, and chimeras were removed, as described in Blackburn-Desbiens et al. (*in prep*).

Microbial taxonomy was assigned to raw shotgun reads using PhyloFlash 3.4.1 (Gruber-Vodicka et al., 2020) for metagenomes and the SILVA v138 database (Quast et al., 2012) for amplicon reads. Since the sequencing of one of the bottom samples (Bottom C) yielded very few reads (**Table S1**), it was removed from this analysis. Taxonomy for amplicon reads was assigned from the SILVA v138 database (assignTaxonomy() and addSpecies()), and chloroplasts and low-abundance taxa were removed in phyloseq (McMurdie and Holmes, 2013), as described in Blackburn-Desbiens et al. (*in prep*). The *Escherichia-Shigella* genus, which can indicate

contamination from sample handling, was also removed due to high presence in cDNA (RNA) samples, attributable to the low yield in these samples.

2.2.5 Psychrophilic trait annotation

A list of genes associated with psychrophily was collated from a review on genomic adaptations of cold-dwelling prokaryotes (Bowman, 2017), including solute uptake, protection against toxic oxidative reactions, and protection against cold denaturation (**Table S2**). Those functions were then searched for in the National Center for Biotechnology Information (NCBI) Nucleotide database (Kans, 2013) for deposited gene sequences and, in case of very few hits, from the NCBI Gene database to allow for enough sequences to represent broader sequence diversity. Gene sequences were downloaded using NCBI E-utilities `esearch` 17.1 (Sayers, 2009). Resulting fasta files for these genes were manually curated to retain only desired sequences (by removing synthetic constructs, chloroplasts, mRNA, homonyms). After curation, sequences were aligned using Muscle 3.8.31 (Edgar, 2004) with the parameter `-maxiters 2` for files larger than 200 KB (for genes `katG`, `sodA`, `sodB`). Alignments were manually curated using Seaview 5 (Gouy et al., 2020) to ensure alignment consistency and to trim to a consensus region. Afterwards, profiles of homologous sequences were built from these alignments using a Hidden Markov model with HMMER 3.2.1 (Eddy, 2018), and were then aligned with our contigs to detect the presence of those genes in our samples. The contig with the most significant score for each of the genes was then searched with BLASTx 2.14.0 using standard databases (non-redundant protein sequence database) (McGinnis and Madden, 2004) to confirm the identity of hits. A permissive E-value of $1.0e-15$ was chosen as a cut-off as hits below but close to this threshold were correctly

identified by BLASTx, whereas hits with higher E-values never matched with the query genes in BLASTx. Genes whose contigs were not correctly identified by BLASTx, regardless of their HMMER scores, were not retained. An abundance table of occurrences for each gene in each sample was produced. These were normalized by the total number of contigs in the samples and are presented in a heatmap made with `gplots` 3.1.3 (Warnes et al., 2022) in R. Adaptation genes were grouped into solute uptake, protection against toxic oxidative reaction products, and protection against cold denaturation mechanisms for ease of visualizations.

2.2.6 Binning of metagenome-assembled genomes (MAGs)

Binning was performed on contigs produced from the co-assembly using Anvi'o v7.1 (Eren et al., 2020) and following the Metagenomic Workflow (Eren et al., 2015), using Kaiju 1.6.2 (Menzel et al., 2016) to import taxonomy into Anvi'o's contigs database. 27 bins were obtained, of which ten could be reported as Metagenome-assembled genomes (MAGs) (<2Mb in size and <10% redundancy). Another bin was manually refined (*Flavobacterium*) to obtain an acceptable level of redundancy. One bin that had a slightly higher level of redundancy than expected (+2.68%) (*Polaromonas*) was kept for further assessment (Eren et al., 2015).

2.2.7 Statistical methods

Taxonomic composition matrices were converted from absolute to relative abundance to compare means of taxonomic groups across ice layers and with water. Permutational multivariate analyses of variance (PERMANOVA) with $1e+6$ permutations were performed in `vegan` 2.6.4 (Oksanen et al., 2001) to determine the

contribution ice layer in structuring microbial communities, with a significant threshold at $\alpha = 0.05$. A principal Coordinate Analysis (PCoA) was also built with `vegan` using a Bray-Curtis dissimilarity matrix to visualize community-assembly patterns across the dataset.

2.2.8 Data availability

Shotgun metagenomic sequencing data will be deposited at NCBI Sequence Read Archive (SRA). Amplicon sequencing data are available under the SRA BioProject PRJNA1119994.

2.3 RESULTS

2.3.1 Taxonomic annotation

WHL ice contained diverse bacterial communities at the phylum level. Taxonomic annotation of shotgun reads produced a greater number of low-abundance phyla (**Figure 2A**), as shown by the greater proportion of phyla grouped into “Others”. Pseudomonata accounted for nearly half of all metagenomic reads in ice (average of 52% across all layers), but only 31% in water under ice. Bacteroidota were the second most prevalent phylum with higher abundance in water (8-22%) and top ice (average of 18%) than in middle and bottom ice (approximately 4%). Actinomycetota accounted for 3-7% of reads in all samples. Cyanobacteria represented 2-12% (average of 6%) of reads and was the second most abundant phylum in bottom ice. Bacillota was found in all samples between 2% and 8% and Patescibacteria averages 2%. Other phyla each accounted on average for less than 1% of reads.

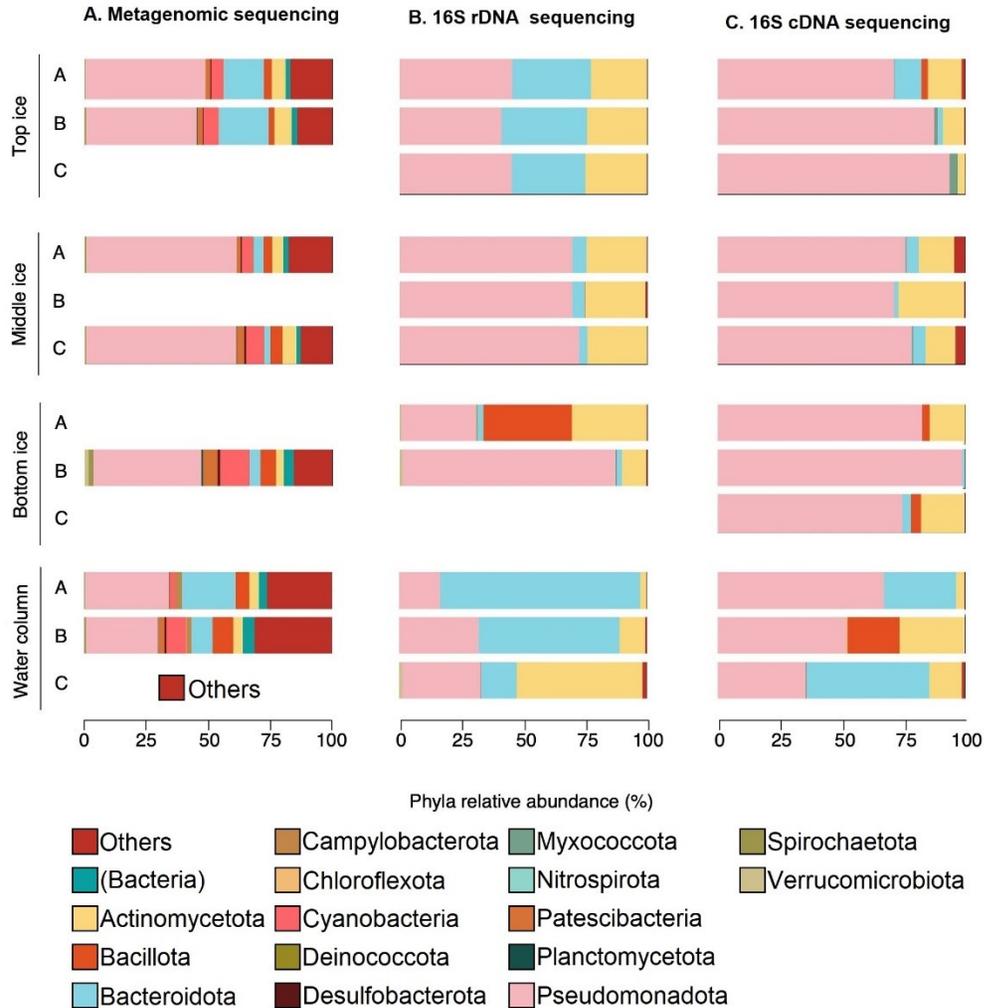


Figure 2. Relative abundance of bacteria at the phylum-level in lake ice and the water column under the ice from triplicate ice cores for **A.** shotgun metagenomic sequencing, **B.** amplicon sequencing of the 16S rDNA gene and **C.** amplicon sequencing of 16S rRNA transcripts (cDNA). Gaps show samples for which sequencing yielded too few reads and were removed, or that were not sequenced (**Table S1**). Phyla representing less than 0.07% (from shotgun sequencing) and 0.01% (from amplicon sequencing) of relative abundance across all samples were combined into “Others” for visualization purposes. Organisms grouped in “(Bacteria)” are bacterial organisms for whom phylum assignment could not be resolved.

Taxonomic assignment for rDNA amplicons produced fewer low-abundance phyla (**Figure 2B**), with a lower proportion of organisms grouping as “Others” (~20%). Pseudomonadota was the most represented phylum overall, similarly to shotgun metagenomic sequencing, accounting for 41-46% of top ice communities, and

constituting up to 73% of assemblages in middle ice. Bottom ice showed heterogeneous taxonomic composition between replicates, with Pseudomonadota representing 86% of the population in one sample and 30% in the other. Pseudomonadota were less abundant in water (17-32% of reads). Bacteroidota and Actinomycetota made up most of the rest of the community with an average of 23% each. In the top layer, Bacteroidota represents 32% of reads, but only 5% and 2% of middle and bottom layers respectively. Water samples varied widely with Bacteroidota dominating two samples (81%, 57%) and being third most abundant in the other replicate (14%). The distribution of Actinomycetota was more even across ice samples, averaging 23%. Similarly to Bacteroidota, the relative abundance of Actinomycetota varied greatly across water samples, between 2% and 51%. Dominant phyla were therefore generally consistent across both shotgun and amplicon sequencing data (**Figure 2A & B**). In both cases, Pseudomonadota was the dominant phylum with higher representation in ice than in water. Bacillota was only found in one sample (Bottom A) and accounted for 36% of its reads. Verrucomicrobiota represented 1% of bottom ice samples and water (replicate C). Other groups like Planctomycetota, Cyanobacteria and Deinococcota are found in only one sample each and accounted for less than 1% of reads.

Taxonomic annotation of 16S rRNA transcripts (cDNA) at the phylum level was dominated by Pseudomonadota, at greater proportions than those found in 16S rDNA (**Figure 2B & C**). Pseudomonadota accounted for up to 94% of transcripts detected in ice. The relative abundance of this phylum is greater than 70% in all ice samples and varied between 36% and 67% for the water replicates. Remaining transcripts mostly belonged to Actinomycetota (13%) and Bacteroidota (4%), the latter

varying widely between samples and accounting for 50% of one water replicate (Water C).

Across all sequencing approaches, Pseudomonadota were consistently the most abundant phyla. Within this phylum, 11-27% (shotgun), 12-41% (rDNA amplicon), 0-51% (cDNA amplicon) of these organisms belonged to *Comamonadaceae*, a psychrophilic Pseudomonadota family. From taxonomic annotation of shotgun metagenomic reads, *Polaromonas* largely dominated *Comamonadaceae* with between 48% and 63% of reads in ice and 30% in water (**Figure S1A**). A notable proportion (22-41%) of taxa were either unassigned at the genus level or detected in very low abundances. In amplicon rDNA and cDNA datasets, *Polaromonas* was again dominant, representing over half the reads across all samples within its family (**Figure S1B & C**). *Paucibacter* was another genus from the *Comamonadaceae* family that showed a constant presence (5% in shotgun rDNA, 15% in amplicon rDNA and 10% in cDNA), mostly in ice (**Figure S1B & C**).

To explore the microbial composition of WHL ice at a finer taxonomic resolution, we examined the most abundant genera detected from both sequencing methods (rDNA) (**Figure 3**). Metagenomic reads annotation produced a high proportion of “Others” or unassigned taxa (average of 69%) (**Figure 3B**). Despite this, the dominant groups from this taxonomic annotation were shared with rDNA amplicon assignments (**Figure 3B**). Apart from the overall dominant *Polaromonas* (25%) we found that *Hymenobacter* accounted for 31% of the diversity in top ice and 15% average across all samples. *Parafrigoribacterium* was the second most abundant genus with a relative abundance of 17% and a higher presence in top and middle ice.

Actinobacterium accounted for 15% of the overall taxonomy with a higher presence in the middle layer (26%). Again, the bottom layers again showed heterogeneity between replicates, but *Polaromonas* and *Sphingomonas* remained the most abundant genera. Similarly to bottom ice, community composition in water collected under the ice showed high heterogeneity; *Flavobacterium* was the dominant genus, with 13%, 57%, and 81% of the diversity in the samples. *Polaromonas* varied between 9% and 24% of reads detected in water. *Flavobacterium* was not found in higher than 2% proportion in ice samples. *Parafrigoribacterium*, while highly abundant in the amplicon-based taxonomy, was not detected in the metagenomic taxonomy. *Paucibacter* accounted for 4% of the reads, with a prevalence in middle ice (average of 7%). Taxonomic assignments of cDNA transcripts showed similar patterns to the rDNA sequencings with *Polaromonas* as the overall most identified genus (27%), *Flavobacterium* almost exclusively in the water (0%, 29%, 48%), *Hymenobacter* mostly in the top layer (10%, 2%, 0%), and *Paucibacter* being prevalent in middle ice (average of 11%).

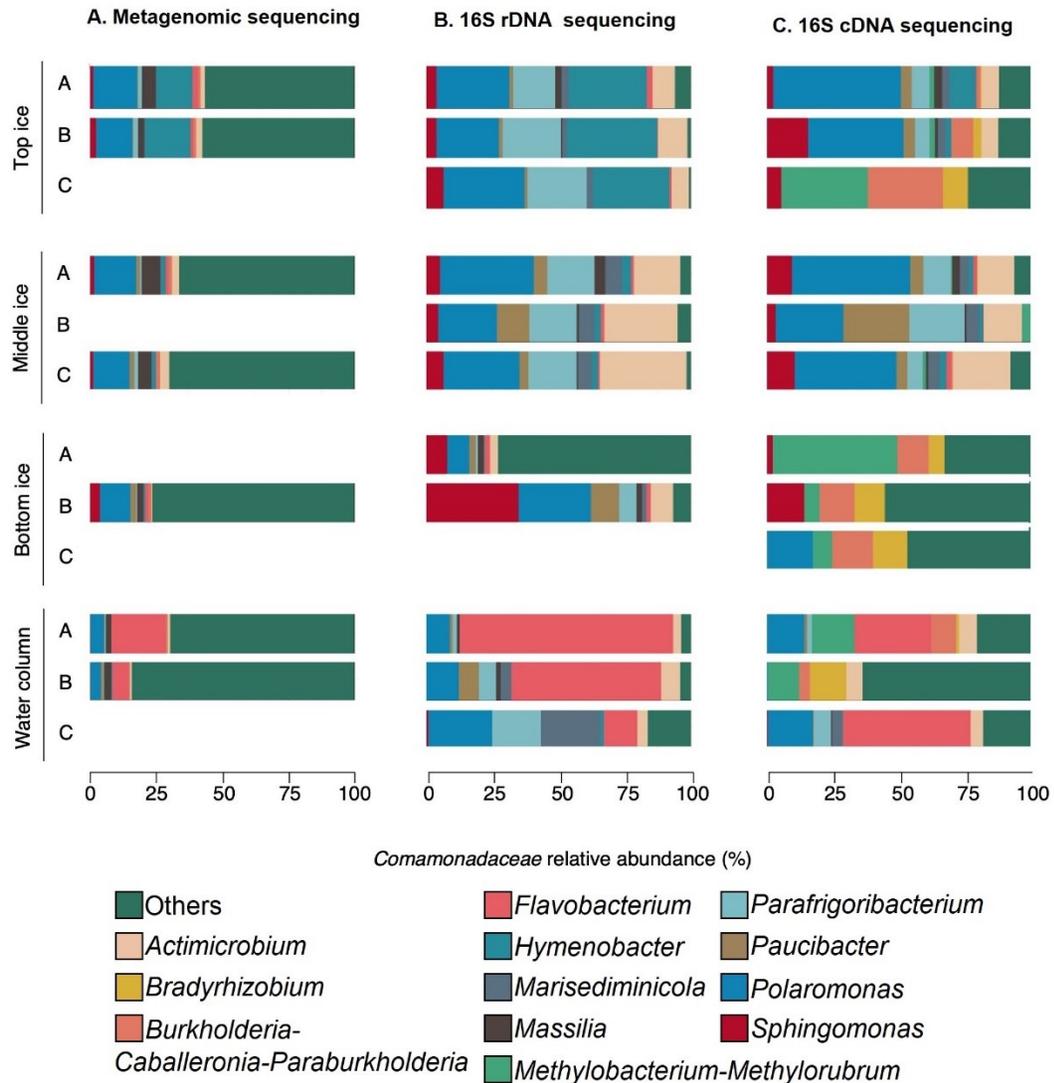


Figure 3. Relative abundance of bacterial genera in lake ice and water from under the ice from triplicate ice cores for **A.** shotgun metagenomic sequencing, **B.** amplicon sequencing of the 16S rDNA gene and **C.** amplicon sequencing of 16S rRNA transcripts (cDNA). The twelve most abundant genera across all samples are shown, while the others were combined into “Other”.

A PCoA from Bray-Curtis dissimilarity allowed us to compare bacterial assemblages in ice and in water column while considering community-level diversity (**Figure 4**). For the rDNA PCoA (**Figure 4A**), ice layers were strongly stratified along Axis 1 (40% variance explained), and more weakly along Axis 2 (19% variance explained). Here, the bottom layer of ice distinguished itself from the other layers, with

both samples being very different from each other. We could see no clear pattern emerge from the cDNA PCoA (**Figure 4B**). A PERMANOVA was performed to assess the effect of the strata (Top, Middle, Bottom, Water) on the community, which showed no significant effect ($p = 0.1277$ for rDNA, $p = 0.6086$ for cDNA). Another PERMANOVA was performed on the sample type (ice or water) for the rDNA samples ($p = 0.07891$ for rDNA, $p = 0.2318$ for cDNA). However, low sampling number limited statistical power for analyses.

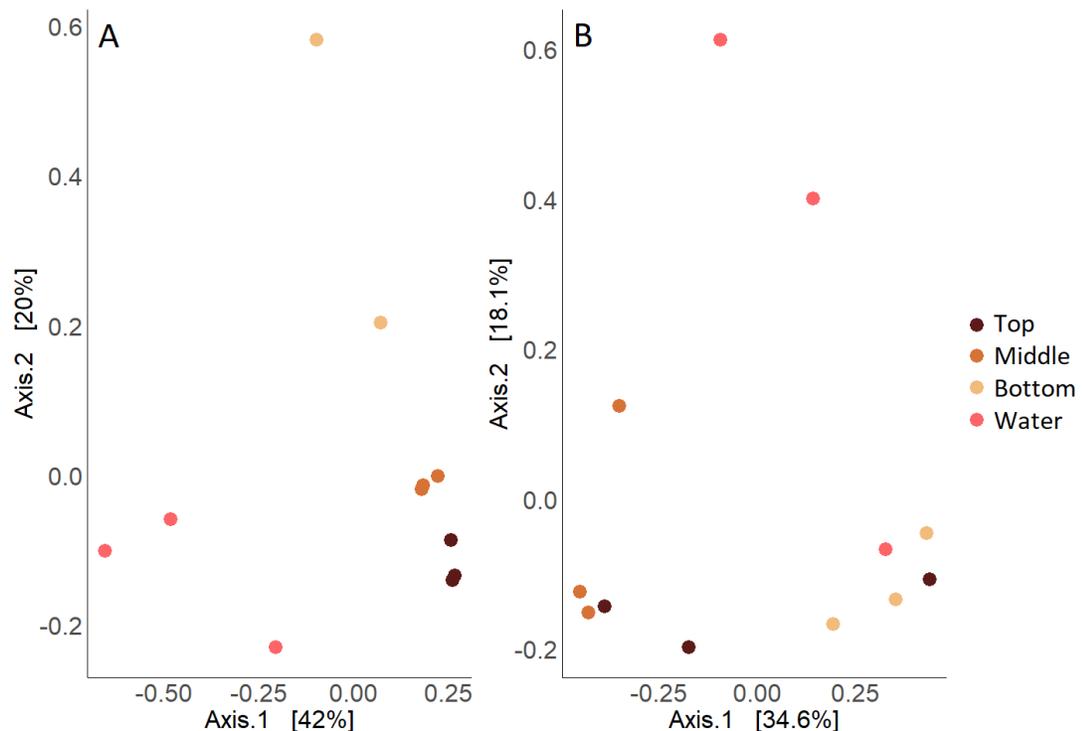


Figure 4. Principal coordinate analysis from Bray-Curtis dissimilarity index showing bacterial community patterns of across ice layers and water for **A.** 16S amplicon rDNA sequencing and **B.** 16S amplicon cDNA sequencing.

2.3.2 Functional potential

Psychrophilic genes involved in solute uptake, protection against toxic oxidative reaction, and protection against cold denaturation were generally

proportionally less abundant in water under the ice, as well as in the bottom layer of lake ice than in upper ice layers (**Figure 5**). Most psychrophilic genes were detected in proportionally greater abundance in top ice and in the middle sections, with one middle sample clustering with the top layers (Middle A) and one (Middle C) clustering with the water and bottom layer, while still showing greater presence of genes than those (**Figure 5**). Oxidative stress genes *sodA* and *sodB* were found in greater proportion in one of the water samples (Water A). A gene encoding a cold-shock protein (*cspA*) was detected only in two samples, one top and one middle, with the top sample presenting the gene in three orders of magnitude higher than the middle one. The middle samples did not cluster together, but one solute uptake gene (*proW*) was detected in greater abundance inside the middle samples. Oxidative stress genes *efeB*, *katE*, *btuE* and *katG*, as well as solute uptake *otsA* were mostly found in Top samples (**Figure 5**).

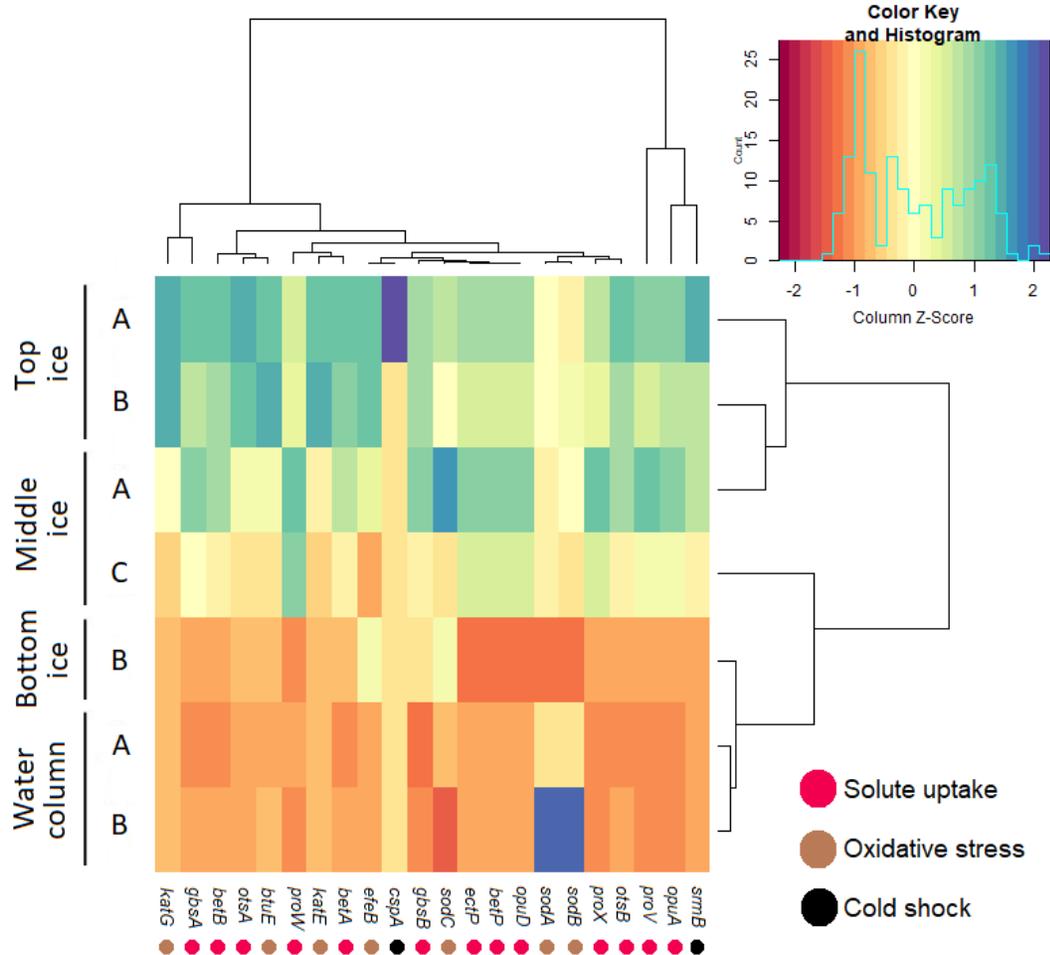


Figure 5. Heatmap of normalized abundance of psychrophily-associated genes in ice core replicates identified with HMMER for three groups of functional adaptations: solute uptake (red), protection against toxic oxidative reaction products (brown), and protection against cold denaturation (black). Reads were normalized per sample, dendrograms show distance-based clusters of similar samples (right) and genes (top). The Z-score is a measure of how many standard deviations there are between the abundance of a gene in a sample and the mean abundance of that gene across all samples.

27 bins were assembled from the metagenomic co-assembly, with each bin representing a putative genome for a single organism reconstituted from metagenomically-produced genome fragments (**Table 1**). Of these, eleven met the criteria to be classified as MAGs, which included high completion (> 2Mb), and low

redundancy (<10%*). Among them we found many genera that were detected in high abundance in taxonomic annotations (**Figure 3**), including two *Polaromonas* and two *Hymenobacter* MAGs. *Flavobacterium*, *Paucibacter*, *Massilia* and *Sphingomonas* were also found across our three sequencing methods, strongly suggesting that these genera are present and active throughout lake ice.

Table 1. Metagenome-assembled genomes (MAGs) from the co-assembly, ordered by expected completion of the genome (%).

Taxonomic annotation	Total size (Mb)	Completion (%)	Number of contigs	Redundancy (%)	Number of annotated genes
<i>Sphingomonas</i> genus	3.49	98.59	295	2.82	3 589
<i>Polaromonas</i> genus	3.49	95.77	374	12.68	3 559
<i>Polaromonas</i> genus	3.69	76.06	460	2.82	3 709
<i>Microbacteriaceae</i> family	2.01	71.83	385	5.63	2 133
<i>Hymenobacter</i> genus	4.97	64.79	504	0.00	4 813
<i>Paucibacter</i> genus	3.57	49.30	606	2.82	3 564
<i>Flavobacterium</i> genus	2.60	46.48	619	1.41	2 689
<i>Burkholderiaceae</i> family	2.93	39.44	509	7.04	3 076
<i>Massilia psychrophila</i> sp.	2.26	35.21	539	1.41	2 398
<i>CAILRJ01</i> (<i>Burkholderiales</i>)	3.19	30.99	576	4.23	3 274
<i>Hymenobacter</i> genus	2.36	28.17	554	0.00	2 350

Most MAGs were constructed from contigs from the upper and middle layers of ice from one of our cores (core A) (**Figure 6A**), while the bottom layer did not contribute many contigs to MAG reconstruction. *Flavobacterium* was however an exception, as this MAG was mostly reconstructed from contigs assembled in water under the ice and was the most abundant and active genus found in the water column (**Figure 3**).

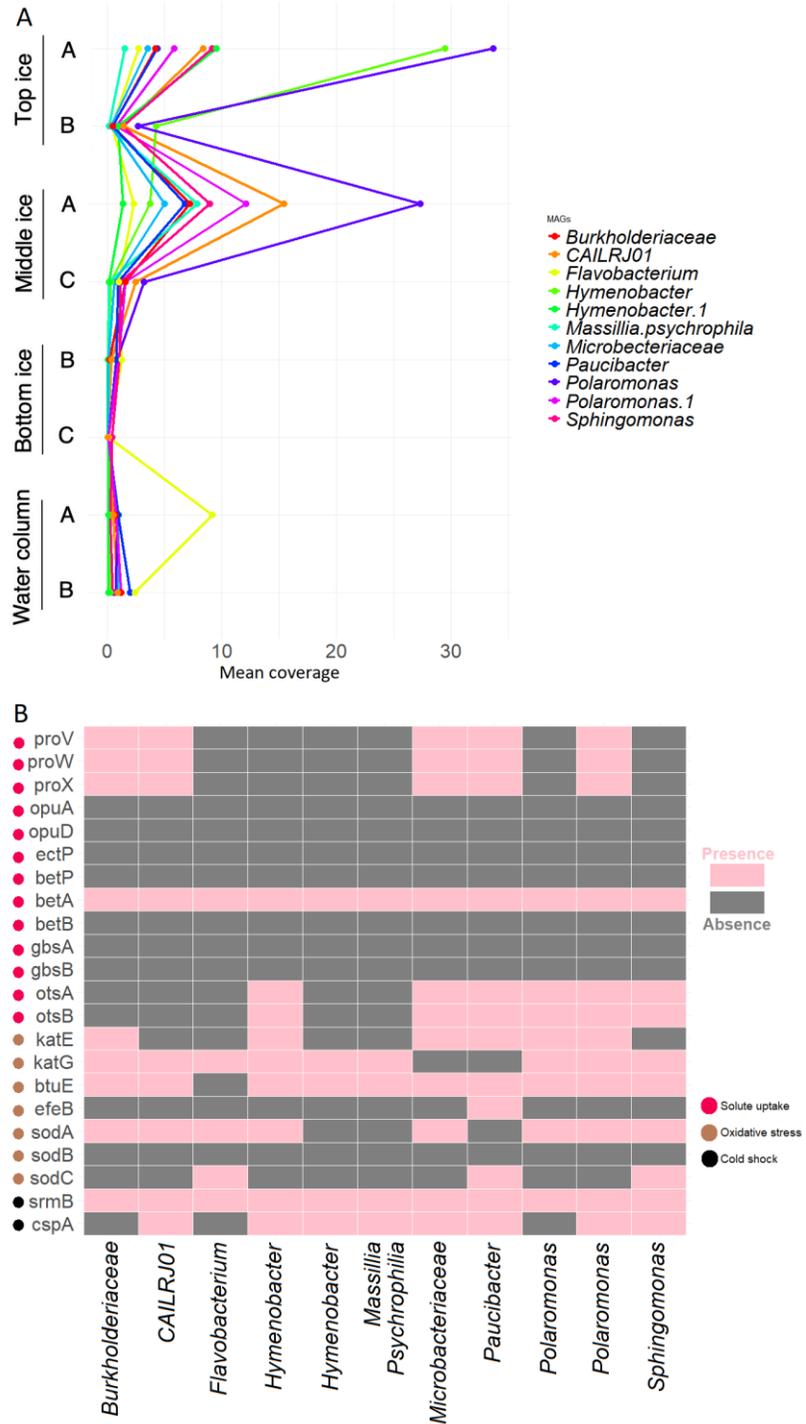


Figure 6. A. Distribution of mean coverage of high completion MAGs in ice layers and water (number of hits/total number of contigs) **B.** Presence/absence of psychrophily-associated genes (presented in **Figure 5** and **Table S3**) detected in high-completion MAGs.

We investigated whether some of the psychrophily-associated genes we identified through Hidden-Markov models on our contigs (**Figure 5**) were also found on the high-completion MAGs, i.e. if they were found on contigs that were recruited into the assembly of these MAGs. We found that the contigs recruited into MAGs did not all contain the adaptation genes we targeted, such as solute uptake genes *opuA*, *opuB*, *ectP*, *betP*, *betB*, *gbsA* and *gbsB*, and oxidative stress gene *sodB* which were absent from the 11 high-confidence MAGs (**Figure 6B**). Only *srmB* and *betA* were detected in all the MAGs. Both pairs of *Polaromonas* and *Hymenobacter* showed some differences in their psychrophilic adaptations. Out of the 22 genes that were found within our samples, both the solute uptake genes *otsA* and *otsB* along with both oxidative stress adaptations *katE* and *sodA* were present in one *Hymenobacter* MAG and absent from the other. Another four genes (cold-shock *cspA* and solute uptake *proV*, *proW* and *proX*) are present in one of the *Polaromonas* MAG and not in the other. The *Flavobacterium* MAG, which was mostly recruited from contigs assembled in the water column, did not distinguish itself much from the other MAGs in terms of the presence and absence of adaptations – it did, however, share a rare gene (*sodC*, an oxidative stress adaptation) with two other MAGs (*Paucibacter*, *Sphingomonas*). It had, along with two others (*Hymenobacter*, *Massilla psychrophilia*), the lowest number of adaptations from the HMMER detection profiles, with five of the 22 genes. Other MAGs had an average of 8.5 genes found in MAGs. Amongst the 13 genes associated with solute uptake, seven were not found in any MAGs while only one of the oxidative stress genes was not identified. Only three MAGs did not carry both cold-shock genes *cspA* and *srmB*.

2.4 DISCUSSION

Lake ice is a feature of most lakes, with over 85% of lakes experiencing at least some period of ice cover every year (Korver et al., 2024). At greater latitudes, this ice-covered period can extend year-round, making lake ice a stable habitat in freshwater systems. However, lake ice has received less attention from microbiologists than other compartments of the cryosphere (such as glaciers, sea ice and permafrost), which have been shown to harbour diverse microbial communities (Bourquin et al., 2022). Our data present the microbial composition of lake ice from Ward Hunt Lake, Canada's northernmost freshwater system, which at the time of sampling had maintained its ice covered for six continuous years.

The Pseudomonadota phylum was the most abundant in almost all our samples across every sequencing method. Actinomycetota (in ice) and Bacteroidota (in water) were the other two major phyla in our samples. While it was the first sequencing of Ward Hunt Lake's ice cover, its water column showed Pseudomonadota and Bacteroidota as the most abundant phyla, similarly to what has been observed by others (Comte et al., 2018). *Polaromonas*, a ubiquitous psychrophile (Ciok, et al. 2018) that was found on almost every continent, in various cryospheric environments (Darcy et al., 2011), was the overall most abundant genus in our ice samples. This genus has been shown to be amongst the most abundant on glacier surfaces (Gawor, et al. 2016), another terrestrial freshwater cryospheric environment. Previous results from nearby ice habitats (including lake ice) showed *Polaromonas* as significantly more characteristic of ice than water (Girard et al., 2023). In this study, *Polaromonas* accounted for between ~20% and ~40% of the genera abundances of lake ice. Their

presence in the ice of WHL, but also at lower but detectable abundances in our water samples may indicate that *Polaromonas* is integrated into ice from freezing lake water. However, at the time of sampling, the ice cover on WHL was thinning, and it is more likely that *Polaromonas* found in water are either lacustrine ecotypes (Girard et al., 2023), or originated from melting ice. In any case, the presence of this genus in both habitats suggests *Polaromonas* might maintain itself in the water column even if the ice cover on WHL melts entirely. Niche adaptation of *Polaromonas* has been observed elsewhere (Gawor et al., 2016; Girard et al., 2023), indicating this taxon maybe especially resilient to habitat loss.

Flavobacterium is both the most abundant and most active genus in water samples and has been found before in epilithic biofilms, accompanied by *Polaromonas* (Busi et al., 2022). Our detection of *Flavobacterium* in the water column of WHL is consistent with its previous isolation from a cold oligotrophic lake (Nam et al., 2017). *Flavobacterium* was also found in our ice samples, but in very low proportions, which was surprising considering its high presence detected by others in glacier ice (Miteva, 2008). *Hymenobacter* was mostly found in top ice, but was also found throughout ice cores and in the cDNA sequencing (from RNA) but in lesser proportions, indicating that this could be a genus with lower activity inside the ice, and mostly active at the surface. This observation is supported by other results that show similar proportion of presence and activity for *Hymenobacter* (Bradley et al., 2022). This could be due to *Hymenobacter* originating from white ice and percolating into lower ice strata via cracks, instead of being included during freezing of black ice at the water column

interface. This is consistent with previous isolations of members of the *Hymenobacter* genus from snow, permafrost and aerosols (Klassen and Foght, 2011).

As most of the ice on WHL is black ice, which originates from lake water freezing and accreting on the bottom layer of the ice cover (Leppäranta, 2010) we expected lake water to be more similar to bottom ice. However, it appears that some psychrotolerant taxa, such as *Flavobacterium* (Nam et al., 2017) are not as well represented in ice as expected. Indeed, *Flavobacterium* were instead overrepresented in water under ice. This genus could be excluded from ice during freezing (which could lead to its concentration in water at the ice interface), or cells may be incorporated into ice but not survive. A mesocosm experiment on inclusion of bacteria into lake ice showed that bacterial aggregates at the water-ice interface when cells are viable, which could promote incorporation into ice, while dead cells were excluded (Santibáñez et al., 2019). Sequencing of cDNA from WHL revealed that while *Flavobacterium* were proportionally less active than what was expected from their detection counts using rDNA (**Figure 3**), an important proportion of them were alive in water under ice, suggesting they had the potential to form these aggregates and be included in ice. Their low presence in ice is therefore a divergence from what was found in the above-mentioned research. *Flavobacterium* may therefore be included inside the ice, but might not survive in it.

Actimicrobium is a newly discovered genus within the *Oxalobacteraceae* family that was first isolated in Antarctica's coastal seawater, which is fed by glacier ice melt (Kim et al., 2011). The genus is cold tolerant thanks to its high fatty acid content (Kim et al., 2011). Its high presence and activity in our ice samples hints at a bipolar

distribution of ecotypes for this genus, as evidence by niche adaptation in other genera to lake ice conditions (Girard et al., 2023). *Parafrigoribacterium* is another poorly known genus (Kämpfer et al., 2000) that we detected in ice. This aerobic taxon has been identified in glacial water channels and is also closely linked to life in ice (Varliero et al., 2021), further showing the similarities between lake ice and other compartments of the cryosphere, even those with different origins (like glaciers).

Cyanobacteria are common features of Arctic freshwater microbiomes, and have been found in WHL before (Jungblut et al., 2010; Mohit et al., 2017). Cyanobacteria can be filamentous and live on benthic or icy substrates, or be planktonic and thrive in the water column of lakes (Quesada and Vincent, 2012). We therefore expected to detect at least some members of this phyla in lake water and in lake ice on WHL. We found Cyanobacteria in shotgun metagenomic sequences, but not in amplicon-based analyses, despite this being the method used by previous studies who reported on the phylum's presence in WHL. The absence of cyanobacteria in amplicon samples may be explained by the fact that amplicon sequencing targeting the 16s V3 V4 regions is not very effective at capturing cyanobacteria diversity, whereas metagenomic sequencing doesn't involve the primers responsible for this limitation (Huber et al., 2019). The cyanobacteria that we found through the metagenomic assay were distributed throughout lake ice, and not restricted to the surface. This may be due to the good penetration of light through black ice, which may sustain photosynthesis even in this difficult environment. The absence of "active" Cyanobacteria throughout ice may be due to the aforementioned limitations of primer pairs, or be due to these cells actually being functionally inactive, and simply frozen during ice accretion from the

water column. The presence of intact photosynthetic pigments or deeper sequencing of RNA transcripts would provide further context on potential activity of cyanobacteria in the ice cover of WHL, and how ice-out may impact production in the system.

Metagenomic sequencing doesn't amplify a taxonomic marker, and instead sequences all available genomic fragments. It was therefore expected that taxonomic assignments produced from metagenomes would differ from amplicon assignments, obtained through the sequencing of a taxonomic marker gene (V3-V4 16S) (**Figure 2 & 3**). The high proportion of "others" means that we either captured more rare taxa or potentially more errors emerged from the assignment, but it showed an overall very similar pattern to the amplicon taxonomic assignment. The metagenomic taxonomy is also useful for downstream analyses and used to confirm annotations performed on amplicon sequences. A previous study performed on High Arctic lacustrine sediments showed that taxonomic assignments can differ between metagenomic and amplicon sequencing in terms of proportions of dominant groups, but also in the presence or absence of groups found by either method (Ruuskanen et al., 2020). A combination of both methods allows us to obtain high-resolution taxonomic composition information (from 16 rDNA) along with functional potential (from shotgun metagenomic sequencing), and to link specific functions back to the organism that carries it on its genome.

Still, sequencing methods based on DNA only (whether they be shotgun- or amplicon-based) can exclusively be used to report on the presence of genomic material: the activity of said organism cannot be inferred based on this detection. This is an important point when studying ice and its underlying water column, as they are extreme

environments, it is therefore expected that some taxa detected in these habitats may form endospores and be dormant, or in other cases, not be alive at all and simply be detected through the conservation of free DNA fragments in the ice matrix. Methods based only on DNA could therefore overestimate the importance of dead or dormant organisms, which do not contribute directly to microbial communities. To curb this bias, we included sequencing of rRNA, reverse transcribed into cDNA, which allowed us to detect actively replicating cells (those producing new ribosomes for protein translation). While a recent study on glacial surfaces reported that the majority of bacteria detected through 16S rDNA sequencing were active (also detected via rRNA) (Bradley et al., 2022) this had never been tested before on lake ice, which differs in its structure significantly. Our taxonomic results show that 16s rDNA and rRNA sequencing produce taxonomies with similar compositions, but that the proportions can vary greatly between the two (**Figures 2 & 3, B & C**). Pseudomonadota makes up more of the activity than the presence in all the samples, suggesting that relying on rDNA alone would lead us to underestimate the potential role of this phylum in lake ice. While comparisons of rDNA and rRNA may be influenced by extraction yield (ie. if DNA was more successfully extracted from samples than RNA, for example), these discrepancies in taxonomic annotation of rRNA and rDNA reads show that the presence of a taxonomic group in a sample does not provide information on their activity and that DNA alone can hardly infer metabolic potential. Given its cold conditions but dynamic freeze-melt dynamics, future studies performed in lake ice would therefore greatly benefit from combining these two approaches, to adequately capture living microbial communities.

Perennial lake ice, although dominated by black ice, does exhibit some internal structures. Indeed, black ice forms through accretion on the bottom, while surface ice is ablated by melt or wind (Dugan et al., 2013). Conditions in thick, multi-annual lake ice also differ throughout its thickness – more labile, lake originating organic compounds are expected to be found in bottom ice, while surface ice receives more light. We therefore expected lake ice strata to be a driver of microbial community composition. While analysis of variance based on ice layers were not significant, the PCoA on DNA sample shows clustering of sample types with middle and top samples being very close to each other. This may be due to some connection between ice layers, which may be possible during the polar summer, when the ice cover is no longer landlocked, and wind-driven movement and solar radiation may lead to the formation of vertical cracks throughout the ice. Bottom samples were not as closely grouped amongst themselves indicating that cores A & B differed greatly in terms of community assemblage (**Figure 4**). This could be explained by the nature of bottom ice; as life requires liquid water and the bottom ice layer is a solid, monolithic one with few pockets of water compared to the more broken or fragmented upper layers of ice, this may preclude the establishment of stable communities. The microbial composition of bottom ice may be more heterogenous, driven for example by the inclusion of aggregates (Santibáñez et al., 2019) which may drive samples away from each other taxonomically speaking. Bottom ice variations could also be explained by the inclusion of water-column biofilm-forming microbes to the bottom of the ice core. No clear pattern can be identified in the RNA PCoA indicating that variables other than ice strata likely drive community composition.

Water samples from under the ice showed more variability between each other – this is surprising given that WHL is stratified and both samples came from surface waters. However, the presence of internal convective currents (Bégin et al., 2021b) may cause some disturbances which could increase the taxonomic distance between our samples. More likely, however, the holes bored into the ice cover of WHL may have led to water movement that may blur actual lacustrine communities. Indeed, when holes are bored, they immediately fill with water from directly under the ice, so water collected at 0 m after holes are made may not be truly representative of the under-ice environment. Still, apart from the heterogeneity of water samples, we expected the water column to have communities distinct from those found in lake ice, given the contrasts in living conditions. While ice and water communities were not significantly different based on analysis of variance, visual inspection of the PCoA show that water samples do appear to cluster away from ice. This suggests that there could be a difference between the ice dwelling communities and the underlying water communities, and greater statistical power (attained through a greater number of ice cores) would be necessary to confirm this hypothesis.

Similarly to taxonomic composition, given the distinct environmental conditions of lake ice, we expected genes associated with psychrophily to also exhibit vertical distribution patterns. The distance-based dendrogram based on psychrophilic gene partitioning showed greater dissimilarities between top ice and bottom/water samples, with one middle sample being clustered quite far with the bottom layers. This result was expected since environmental pressures differ between the upper and lower layers of ice: in a polar desert like WHI, ice accumulates mainly from the freezing of the water

column and contains more particles (like organic matter) originating from the lake, while surface ice is mostly driven by atmospheric deposition, and is more irradiated. It was expected then that the bottom layer of ice would be similar to the underlying water, that ice from the middle of cores would have had time to differentiate itself (through accumulation of ice to the bottom), and that surface ice would be quite different.

DNA from one core (replicate A), specifically from its top and middle layers, provided the vast majority of contigs that allowed for the assembly of our MAGs. This suggests greater diversity and representation of these organisms in the upper layers of lake ice, as more read fragments were available to produce high-coverage, high-completion full genomes. It is then consistent that MAGs from genera found in higher abundances inside the ice (*Polaromonas*, *Paucibacter*, *Hymenobacter*, *Sphingomonas* and *Massilia*) (**Figure 3**) were mostly built from ice contigs, and that *Flavobacterium*, found in water, was assembled mainly from contigs recruited in water. The taxa for which we identified MAGs were all previously detected in various other cold habitats including the water column of Ward Hunt Lake (Comte et al., 2018) and in lake ice from another High Arctic lake in the region (Girard et al., 2023), as well as in a subarctic ice-covered lake (Zakharova et al., 2022) and other frozen habitats, as shown by a recent overview of the microbial communities of the global cryosphere (Bourquin et al., 2022).

The psychrophilic genes we detected in the shotgun metagenomic reads were not all detected on the MAGs that were assembled. This is expected, since metagenomic sequencing produces reads on a huge diversity of organisms, and low-occurrence or rare genes may not be captured at the depth we sequenced, leading to

incomplete MAGs and the potential absence of some psychrophilic markers. Of the MAGs that were assembled, they did not all share the same psychrophilic genes: this was expected since of the eleven MAGs we reported on, we found multiple taxa from distinct genera (ie. *Polaromonas*, *Hymenobacter*, etc.), which are expected to have more divergent genomic make-ups than organisms belonging to the same genus. However, even MAGs belonging to the same genus did not have the same genes: for example, one *Polaromonas* had cold-shock *cspA* and solute uptake *proV*, *proW* and *proX* while the other did not. Similarly, one *Hymenobacter* had solute uptake *otsA* and *otsB* and oxidative stress *katE* and *sodA* that the other was missing. Possible explanations for these discrepancies are that bacteria, and psychrophilic bacteria in particular, have particularly big accessory genomes that can vary greatly between environments (Cazares et al., 2014; Wyka et al., 2022), and some genes that we detected may be on plasmids which are not part of the circular genomes that make up our MAGs. Furthermore, spatial niche heterogeneity (such as water bubbles or cracks in ice) may have allowed fine-scale niche adaptation, supporting the presence of different psychrophilic traits within a given genus.

Those divergences could also be attributed to bias in the sequencing or assembly methods, as although these MAGs were of high completion, they were not 100% complete. Portions of the genome missing from the covered MAGs may encode psychrophily genes that could therefore not be associated to these organisms, or as mentioned above, these genes could be on mobile genetic elements (like plasmids) which are not part of the MAGs, given it is particularly difficult to assemble plasmids and to link them to a specific taxonomic group. Finally, it is also possible that the

differences in psychrophilic traits between MAGs of the same genus may be due to genetic variability between strains since the identification is at the genus. In the case of *Hymenobacter*, since we know it's a low activity genus, one of the MAGs could be of a dormant ecotype (9% of presence vs. 1.5% of activity).

Together these data show a first portrait of life and activity inside multiannual lake ice. Although our sample set is limited by low statistical power, we did observe differences between the communities of ice strata and its underlying water. *Polaromonas*, a known psychrophile, is our most abundant genus in ice and one the more abundant one in the water column. We've also found that psychrophilic adaptations are mainly distributed in upper ice layers, while lower portions of ice are more similar to the water column and may favour distinct adaptations that were not captured in our analysis. The combination of the cDNA amplicon sequencing with the metagenomic sequencing of the same samples allowed us to infer the functional potential of our eleven high completion MAGs, showing that diverse organisms similar to what has been found elsewhere in the cryosphere can thrive in lake ice. As warming progresses in the Arctic, this study presents a baseline of lake ice microbial communities, and it will be interesting to compare these community results with future observations of lake ice with either longer or shorter cover durations. Furthermore, these data present a snapshot of potential organisms that may be released into the WHL water column in the event of complete ice-off, which may transform the lacustrine assemblages of this lake at the forefront of climate change.

CHAPITRE 3: Tests préliminaires de décontamination de carottes de glace

3.1 INTRODUCTION

Avant de débiter l'échantillonnage de la glace, nous nous sommes questionnés quant à la potentielle contamination des échantillons lors des prélèvements. La contamination dans ce contexte fait référence au transfert de microorganismes exogènes à la glace (provenant par exemple d'environnements avoisinants ou de matériel mal décontaminé) à l'échantillon pendant sa collecte, faussant notre lecture de la communauté qui y vit réellement. La biomasse attendue dans la glace étant très faible (Miteva, 2008), la contamination de nos échantillons par des microorganismes de la colonne d'eau, de la neige, de l'air à la surface de la carotte ou de la personne réalisant l'échantillonnage pourrait biaiser les résultats obtenus lors de l'analyse. Ceci pourrait être une barrière majeure, étant donné que ce projet avait pour objectif d'identifier des fonctions biologiques d'organismes spécifiquement de la glace (et non de la neige environnante), et que nos approches de séquençage nécessitent une quantité minimale d'acides nucléiques pour tracer un portrait réaliste des communautés microbiennes.

La contrainte de la quantité d'ADN disponible pour le séquençage est aussi ce qui rend la décontamination particulièrement délicate, puisque selon l'approche elle implique une perte appréciable de matériel et donc de biomasse pour les analyses subséquentes. Avant d'échantillonner le lac Ward Hunt, nous avons donc testé et adapté des protocoles de décontamination des échantillons inspirés de la littérature, pour s'assurer d'être en mesure de distinguer les organismes psychrophiles de la glace des

organismes exogènes (venant de l'environnement ou de l'échantillonnage), sans sacrifier une portion trop importante de nos échantillons.

Dans la littérature, diverses méthodes ont été proposées pour décontaminer de grandes carottes de glace de glaciers (Christner et al., 2005; Zhong et al., 2021), mais elles sont peu appropriées pour nos carottes de glace de lac plus petites, qui ne font que 9 cm de diamètre. Deux de ces protocoles proposent le retrait successif de trois couches de glace sur la circonférence de la carotte, d'abord en grattant la couche extérieure, en décontaminant ensuite à l'éthanol 95% puis en rinçant à l'eau stérile (Christner et al., 2005; Zhong et al., 2021). Ces propositions induisent une perte de 3cm de diamètre aux carottes, ce qui implique une perte d'environ 56% de volume dans le cas de nos carottes de glace de lac et risque de rendre le séquençage difficile, voire impossible, sachant que la glace abrite proportionnellement peu d'organismes en comparaison avec d'autres environnements.

Des tests inspirés de méthodes plus conservatrices, toujours conçues pour de la glace de glaciers (Rogers et al., 2004), ont donc été réalisés pour déterminer leur applicabilité à des carottes de glace de lac de petite taille. Il est à noter que les méthodes proposées dans la littérature avaient toutes fait l'objet de tests en laboratoire sur de la glace artificielle, alors que nos tests ont été réalisés sur de la glace de lac réelle, prélevée à l'hiver dans la forêt boréale. Quatre traitements combinant deux solutions stérilisantes (éthanol 95% et hypochlorite de sodium 5% (NaOCl)) et deux modes d'application (vaporisation et immersion dans un bain de solution) ont été testés. Un bain et une vaporisation d'eau Milli-Q stérile ont servi de contrôles négatifs. L'efficacité des traitements, qui visent à retirer les couches de glace les plus extérieures

de la carotte (qui sont les plus sujettes à être contaminées) sera déterminée par la quantité d'ADN mesurée après extraction; moins un échantillon contiendra d'acides nucléiques (tout en respectant le seuil minimal pour nos applications subséquentes), plus sa décontamination aura été efficace. Une augmentation de la quantité d'acides nucléiques par rapport aux contrôles négatifs suggérera que la décontamination, même si elle a éliminé les couches extérieures de la carotte, a en fait introduit de l'ADN exogène, et a donc contaminé l'échantillon. En effet, la contamination augmente de plusieurs ordres de grandeur la quantité d'acides nucléiques détectable à la surface des carottes de glace (Zhong et al., 2021).

Des limites logistiques se sont imposées dans l'identification des méthodes à tester. Dans l'article de Rogers et al., le bain d'hypochlorite de sodium 5% est privilégié. Toutefois, notre approche vise une application sur le terrain en région isolée (sans accès à de l'infrastructure de laboratoire), qui en raison de l'enjeu du transport de matières dangereuses et de gestion des déchets chimiques, limite notre capacité à réaliser des bains de trempage. La vaporisation d'hypochlorite de sodium serait idéale, c'est pourquoi nous avons voulu vérifier son efficacité. Aussi, le bain comporte un risque de dilution de la solution dans la mesure où il est impossible de renouveler la solution pour chaque carotte prélevée sur le terrain. Nous avons émis l'hypothèse que l'immersion sera la méthode la plus efficace en raison de l'homogénéité de son application, bien que la vaporisation soit plus pratique en contexte de campagne d'échantillonnage. Cependant, l'un des risques encourus par l'usage répété d'un bain est qu'il perde en efficacité (par dilution des agents actifs) ou devienne lui-même contaminé. L'effet de dilution sur l'efficacité des bains de solution sera testé en

fonction des résultats obtenus. Aussi, la nature physique de la glace de lac n'est pas la même que celle de la glace artificielle employée par Rogers; des canaux et imperfections ayant été façonnées par divers épisodes de fonte et de gèle de l'eau et de la neige au cours de l'hiver, la glace de lac peut être très hétérogène, sa glace blanche en surface est souvent friable et opaque alors que sa glace noire à l'interface de la colonne d'eau est souvent très solide et transparente. Cette différence de matrice nous a amenés à préalablement tester le bain d'éthanol dans le but de déterminer si la carotte de glace de lac peut être immergée sans que le décontaminant ne pénètre le centre de la carotte par ses pores et canaux.

3.2 MATÉRIEL ET MÉTHODES

Une vingtaine de carottes de glace ont été récoltés pour ces tests en mars 2022 grâce à un carottier Mark II (Kovacs Enterprise, Roseburg, OR, USA) sur le lac Simoncouche (48.244365, -71.252668) dans la Réserve faunique des Laurentides, puis conservée à -15°C dans des Whirl-Paks© (Filtration Group, Pleasant Prairie, WI, USA) stériles jusqu'aux tests de décontamination. Afin de limiter la contamination, toutes les manipulations de carottes de glace ont été effectuées en portant des gants stériles et nettoyés à l'isopropanol 70% sur le terrain. Pour éviter tout choc thermique, les solutions et les carottes ont été placées à une température de 0°C pendant au moins une heure avant les traitements. Afin de s'assurer que la décontamination ne touche que la surface extérieure de la glace (et non pas la glace interne, qui sert à nos analyses subséquentes), une carotte de glace a été immergée 60 secondes dans une solution d'éthanol 40% colorée à l'aide de colorant alimentaire vert (**Figure 7**).

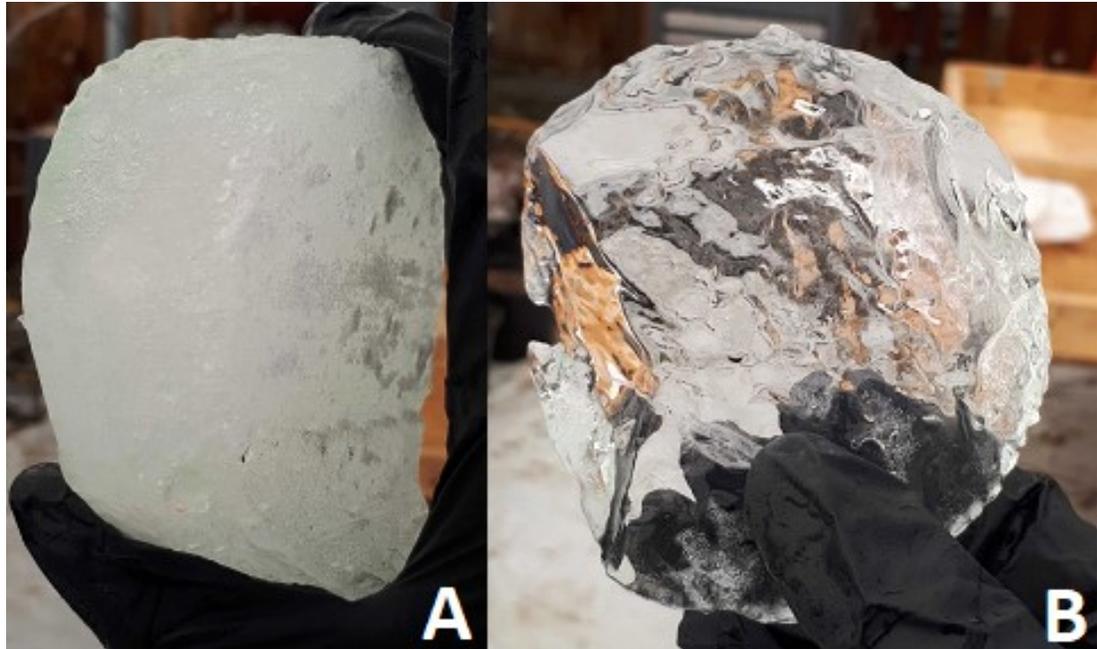


Figure 7. Morceaux de carotte de glace de lac après 60 secondes d’immersion dans une solution 40% éthanol et colorant alimentaire vert. **A.** Glace blanche de surface et **B.** glace noire à l’interface de la colonne d’eau, après immersion. Aucune trace de coloration n’est visible dans les carottes, indiquant que les solutions de décontamination n’ont pas pénétré l’intérieur des carottes.

Au terme de nos tests préliminaires, aucune couleur n’était visible dans la carotte, indiquant que la solution n’a pas pénétré l’intérieur des carottes. Ceci a confirmé que nous pouvions tester les protocoles de décontamination sans nous inquiéter de l’intégrité du cœur des carottes, qui sert ensuite aux analyses subséquentes. Nous avons donc procédé avec les tests de décontamination. Six traitements ont été effectués en triplicata, soit (1) l’immersion dans un bain d’éthanol 95% pendant 60 secondes, (2) la vaporisation avec 100 ml d’éthanol 95%, (3) l’immersion dans un bain d’hypochlorite de sodium 5% pendant 10 secondes, (4) la vaporisation avec 100 ml d’hypochlorite de sodium 5%, (5) l’immersion dans un bain d’eau stérile pendant 60 secondes (contrôle négatif) et (6) la vaporisation avec 100 ml d’eau stérile (contrôle négatif). Ensuite, pour les six traitements, les carottes de glace ont été rincées deux fois par vaporisation avec

100 ml d'eau stérile (Milli-Q) puis placées dans des sacs stériles. Les carottes ont par la suite été fondues au réfrigérateur à 4°C avant d'être filtrées pour concentrer les microorganismes sur des filtres Sterivex™ (0.22µm) en vue de l'extraction des acides nucléiques, tel que décrit dans le Chapitre 2. L'ADN résultant a été quantifié à l'aide du Qubit 3.0 Fluorometer (Thermo Fisher Scientific) et de la trousse dsDNA HS selon le protocole du fabricant. Des PERMANOVA visant à comparer l'effet des traitements sur le rendement en acides nucléiques ont été réalisées telles que décrites dans le **Chapitre 2.2.7**.

3.3 RÉSULTATS ET DISCUSSION

Tableau 2. Résultats des décontaminations pour les solutions contrôle, hypochlorite de sodium 5% et éthanol 95% par bain et par vaporisation. La limite de détection pour la quantification de l'ADN selon la trousse utilisée est de 0.005 ng/μl, pour un minimum de 0.1 ng d'ADN dans l'échantillon. Les valeurs non détectables (en deçà de ce seuil) ont été indiquées comme telles dans le tableau.

Carotte	Traitement	Solution	Mode d'application	Réplikat	Concentration d'ADN (ng/μl)
EB1	1	Éthanol	Bain	A	0.085
EB2	1	Éthanol	Bain	B	0.0888
EB3	1	Éthanol	Bain	C	0.151
EB4	1	Éthanol	Bain	D	0.0516
EV1	2	Éthanol	Vaporisation	A	Indétectable
EV2	2	Éthanol	Vaporisation	B	0.044
EV3	2	Éthanol	Vaporisation	C	0.013
HB1	3	Hypochlorite	Bain	A	Indétectable
HB2	3	Hypochlorite	Bain	B	Indétectable
HB3	3	Hypochlorite	Bain	C	Indétectable
HV1	4	Hypochlorite	Vaporisation	A	0.0338
HV2	4	Hypochlorite	Vaporisation	B	Indétectable
HV3	4	Hypochlorite	Vaporisation	C	Indétectable
CB1	5	Contrôle	Bain	A	Indétectable
CB2	5	Contrôle	Bain	B	0.0128
CB3	5	Contrôle	Bain	C	0.0968
CV1	6	Contrôle	Vaporisation	A	0.111
CV2	6	Contrôle	Vaporisation	B	0.011
CV3	6	Contrôle	Vaporisation	C	0.133

Tous les rendements en ADN obtenus sont plus faibles que le minimum recommandé pour un séquençage efficace de 1 ng/μl, quel que soit le traitement (**Tableau 2**). Sept valeurs sur 19 sont sous le minimum détectable par l'appareil, dont cinq pour le traitement à l'hypochlorite de sodium. Les deux autres sont obtenus par l'éthanol et un contrôle. Les valeurs le plus élevées d'ADN (>0.1 ng/μl) sont obtenues par la vaporisation d'eau stérile (x2) et le bain d'éthanol (x1). Aucune des trois

variables explicatives (solution, mode d'application, réplicat) n'explique significativement le rendement en ADN ($p = 0.8213$, $p = 0.5751$, $p = 0.9446$). En attribuant artificiellement la valeur du seuil de détection aux échantillons sous la limite de l'instrument, aucune des mêmes trois variables explicatives n'explique significativement le rendement ($p = 0.07504$, $p = 0.77317$, $p = 0.75668$), bien que la solution s'en rapproche.

3.4 CONCLUSION

À première vue, il semble que l'hypochlorite de sodium soit le traitement qui ait été le plus efficace, puisque les valeurs de concentration de l'ADN extrait de ces échantillons sont sous le minimum détectable (**Tableau 2**). Les valeurs mesurées pour les autres traitements sont particulièrement faibles puisqu'il est généralement recommandé d'avoir des concentrations de >10 ng/ μ L, et un minimum de >1 ng/ μ L pour un séquençage métagénomique efficace, selon la plateforme qui a réalisé nos réactions de séquençage (*Integrated Microbiome Resource*, Dalhousie University). Comme il est attendu que la contamination augmente de plusieurs ordres de grandeur la quantité d'ADN à la surface des carottes (Zhong et al., 2021), nous avons conclu que la contamination, si elle avait lieu, était minime. Nos méthodes de collecte sur le terrain, impliquant des gants stériles, l'utilisation de surfaces autoclavées pour le travail, l'aseptisation du matériel avec de l'isopropanol ainsi qu'une manipulation minimale des carottes semblent donc avoir protégé la glace de contaminants exogènes, soit de la neige sur le lac ainsi que de la procédure d'échantillonnage en elle-même. Ainsi, l'hypochlorite de sodium, qui semble plus efficace, s'attaque probablement en majorité

à des organismes endogènes qui sont naturellement peu abondants dans la glace, comme le démontrent les résultats des immersions en bain d'eau stérile (**Tableau 2**).

Au vu de ces résultats, et en raison des complications techniques à transporter de grandes quantités de produits chimiques dangereux dans un environnement éloigné et isolé, nous avons conclu que la décontamination, lorsque l'échantillonnage est effectué en prenant toutes les précautions nécessaires à la prévention de la contamination, n'était pas nécessaire. Des analyses plus poussées, comme une contamination intentionnelle des carottes avec une solution d'organismes connus ou de la neige fondue, ou bien le séquençage de l'ADN extrait, auraient pu être entreprises afin de nous assurer que l'ADN décompté n'était pas d'origine allochtone ou même anthropique. Malgré leurs limites, ces tests préliminaires nous ont permis de déployer un design d'échantillonnage terrain soutenu par des résultats et cohérent avec les contraintes matérielles qu'implique une campagne dans le Haut Arctique. Les contaminants réels et potentiels de nos échantillons ont été triés suite au séquençage lors des analyses bio-informatiques (voir **Chapitre 2.2.4**). Les résultats obtenus au cours de ce projet sont cohérents avec ce qui était attendu et tendent à confirmer que notre échantillonnage prudent a suffi à produire un portrait juste des communautés microbiennes de la glace.

CHAPITRE 4: Conclusion générale

Ce projet nous a permis de faire la lumière sur les communautés microbiennes de la glace du lac Ward Hunt et plus largement sur l'activité potentielle de ces microorganismes. Les adaptations psychrophiles peuvent nous renseigner sur les limites du vivant et encore davantage dans le contexte de glace multiannuelle. Bien que la pérennité du couvert de glace du lac Ward Hunt ait été fragilisée, peu de couverts de glace d'eau douce de six ans peuvent être échantillonnés dans l'hémisphère nord. La position de l'île Ward Hunt dans le dernier refuge de glace (LIA, *Last Ice Area*) risque également de protéger ce site des grandes variations vécues par d'autres lacs du Haut Arctique. En effet, les gyres et courants marins dominants dans l'Océan Arctique poussent la glace de mer contre le nord du Groenland et de l'île d'Ellesmere, où l'on retrouve la glace de mer la plus vieille et la plus épaisse de l'Arctique (Newton et al., 2021). Il est attendu que la région de la LIA soit la dernière à maintenir de la glace de mer estivale d'ici 2050, conférant à la région un statut de refuge pour les espèces dépendantes de la glace. L'effet refroidissant de cette bande de glace de mer épaisse contribue significativement au maintien de conditions froides le long de la marge terrestre de la LIA (Girard et al., 2023) et la position de l'île Ward Hunt dans cette région tampon risque de la protéger (pour l'instant) de trop grandes variations. Si le couvert de glace du lac est vulnérable, nous nous attendons donc à ce qu'il persiste plus longtemps que pour d'autres lacs qui ne sont pas dans la LIA. Il est donc important de documenter la composition biologique de la glace de ce lac pendant qu'elle est toujours multiannuelle.

La glace du lac Ward Hunt était dominée par les Pseudomonata, qui représentaient près de la moitié de toutes les séquences métagénomiques dans la glace, mais seulement le tiers dans l'eau sous glace. Les Bacteroidota étaient le deuxième phylum le plus répandu, avec une abondance plus élevée dans l'eau et dans la glace supérieure que dans le centre de la carotte de glace. Les Pseudomonata étaient également le phylum le plus représenté dans le séquençage de l'ADN par amplicon, à l'instar du séquençage *shotgun* métagénomique. Bacteroidota et Actinomycetota constituaient la majeure partie du reste de la communauté de ce même séquençage. Bacteroidota était surtout présent en surface et dans l'eau alors que Actinomycetota était plus uniformément distribué dans les échantillons de glace, mais moins présent dans l'eau de surface. Les phylums dominants étaient donc généralement cohérents entre les données de séquençage métagénomique et amplicon (**Figures 2A et B**), et comparables à d'autres habitats de la cryosphère. L'annotation taxonomique des séquences d'ARNr 16S au niveau du phylum était aussi dominée par Pseudomonadota, dans des proportions plus élevées que celles trouvées dans l'ADNr, au contraire d'Actinomycetota qui conserve une distribution similaire et de Bacteroidota qui est peu actif (**Figures 2B et C**). Ces résultats illustrent bien qu'il serait périlleux d'inférer l'activité d'un organisme sur la seule base de sa présence, et démontrent l'importance de l'emploi d'approches complémentaires.

Polaromonas, un psychrophile ubiquitaire bien connu (Ciok, et al. 2018), était globalement le genre le plus abondant dans nos échantillons. Cependant, c'est *Flavobacterium* qui était le genre dominant dans nos échantillons d'eau. Les assignations taxonomiques des séquences d'ADNc ont montré des tendances similaires

aux séquençages d'ADNr (**Figures 3B et C**). Onze génomes assemblés à partir de métagénomes (MAGs) ont été produits grâce au séquençage métagénomique. Ces MAGs sont des reconstructions imparfaites de génomes à partir de toutes les séquences d'un métagénome. On y retrouve certains des membres les plus présents et actifs dans nos échantillons comme *Polaromonas* (2x), *Hymenobacter* (2x), *Flavobacterium* et *Paucibacter*. La plupart des MAGs ont été construits à partir de contigs des couches supérieures et du centre de la glace (**Figure 6A**). *Flavobacterium* était cependant une exception, car ce MAG était principalement reconstruit à partir de contigs assemblés dans l'eau. Ces résultats sont donc cohérents avec les annotations taxonomiques précédentes (Bourquin et al., 2022).

La vie dans la cryosphère requiert plusieurs adaptations cellulaires, encodées par des gènes permettant par exemple l'absorption accrue de solutés, la protection contre les réactions oxydatives toxiques et la protection contre la dénaturation par le froid. Les gènes encodant ces fonctions étaient généralement proportionnellement moins abondants dans l'eau sous la glace, ainsi que dans la couche inférieure de la glace de lac, que dans les couches supérieures de la glace (**Figure 5**). Ce résultat était attendu puisque les pressions environnementales diffèrent entre les couches supérieures et inférieures de glace, et que la glace du bas du couvert est formée par le gel d'eau de lac. Ces adaptations identifiées au sein de nos échantillons ne se retrouvent pas dans tous nos MAGs. Sept des treize gènes d'adaptation impliqués dans l'absorption des solutés sont absents de tous nos MAGs. C'était prévisible puisque seuls onze d'entre eux ont été construits. La variabilité génétique entre les genres est aussi importante puisque les MAGs de mêmes genres ne partagent pas tous les mêmes adaptations. Ce

projet marque la première reconstruction de génomes de microorganismes provenant de la glace de lac, en plus de démontrer qu'ils pourraient y être actifs. Leur isolement en culture ayant pour objectif de tester les limites de leurs adaptations pourrait marquer une prochaine étape dans l'exploration du potentiel fonctionnel des psychrophiles.

De futurs projets pourraient s'intéresser à la comparaison entre les lacs en fonction de la durée du couvert de glace. Cependant, nos résultats montrent que la glace de lac, plus largement, comporte une très faible biomasse et que la glace saisonnière, par sa nature transitoire, abrite un plus faible nombre de microorganismes (**Annexe 2 & Tableau S4**). Cette faible biomasse diminue les chances d'un séquençage représentatif de la communauté entière. Pour notre séquençage métagénomique, l'un de nos huit échantillons d'ADN a dû être rejeté en raison d'un mauvais rendement au séquençage alors qu'une majorité de nos échantillons d'ARN a nécessité une filtration importante de contamination. Aucune contamination n'a été détectée dans les données issues du séquençage métagénomique ni du séquençage par amplicon de l'ADN, suggérant que la contamination de l'ARN est survenue soit pendant l'extraction des acides nucléiques, ou qu'elle n'était détectable qu'en raison de la quantité d'ARN extrêmement faible présente dans les échantillons.

Si notre étude présente les communautés bactériennes de la glace de lac, il aurait également été intéressant de s'intéresser aux eucaryotes et aux virus, qui font partie des assemblages microbiens de cet habitat, et qui peuvent être détectés dans des métagénomes. Toutefois, compte tenu de l'utilisation des amorces spécifiques aux bactéries que nous avons employées dans l'approche par amplicon et de notre souhait de comparer les méthodes de séquençage, nous avons restreint notre analyse aux

bactéries seulement. Les données de séquençage découlant de ce mémoire pourront toutefois alimenter de futurs projets sur les eucaryotes et les virus de la glace de lac.

Ensemble, ces données brossent un premier portrait de la vie et de l'activité à l'intérieur de la glace de lac multiannuelle et démontrent que le microbiome de ce compartiment jusqu'à maintenant sous-étudié de la cryosphère est comparable à ceux d'autres milieux cryosphériques d'eau douce. La proportion importante d'organismes dont nous n'avons pas pu résoudre la taxonomie en raison de l'absence d'homologue dans les bases de données représente bien à quel point les communautés de la glace sont méconnues et doivent être explorées davantage. Il serait aussi intéressant de déterminer si la composition de la glace de lac multiannuelle en Arctique est comparable à celle de lacs à couvert saisonnier, comme en Boréale. Des tests de rendement d'ADN ont révélé que la glace en Arctique contient significativement plus d'ADN que celle en zone boréale (**Annexe 2**). Toutefois, il serait intéressant de déterminer, via le séquençage de ce matériel génomique, si la composition des communautés est la même en Boréale qu'en Arctique, et si la proportion de taxons inconnus diminue dans la région boréale. On pourrait ainsi déterminer si la fraction « *unassigned* » est spécifique à l'Arctique (dont les communautés sont généralement sous-étudiées), ou à la glace de lac, qui est un habitat encore plus méconnu.

Bien que notre ensemble d'échantillons offre une faible puissance statistique, il semble y avoir des différences entre les communautés des strates de glace et l'eau sous-jacente. *Polaromonas*, un psychrophile connu, est notre genre le plus abondant dans la glace et l'un des plus abondants dans la colonne d'eau. Nous avons également constaté que les adaptations psychrophiles sont principalement distribuées dans les couches de

glace supérieures. La combinaison du séquençage par amplicon de l'ADNc avec le séquençage métagénomique des mêmes échantillons nous permet de déduire le potentiel fonctionnel de nos onze MAGs à haut degré de complétude. Il serait intéressant de comparer ces résultats de communautés avec de futurs échantillonnages de la glace du lac Ward Hunt à des moments où le couvert de glace est d'une durée plus ou moins longue, ou encore avec des lacs boréaux ou subarctiques dont le couvert est saisonnier et de nature différente.

BIBLIOGRAPHIE

- Alneberg, J., Bjarnason, B.S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U.Z., Lahti, L., Loman, N.J., Andersson, A.F., Quince, C., 2014. Binning metagenomic contigs by coverage and composition. *Nat Methods* 11, 1144–1146. <https://doi.org/10.1038/nmeth.3103>
- Anesio, A.M., Bellas, C.M., 2011. Are low temperature habitats hot spots of microbial evolution driven by viruses? *Trends Microbiol* 19, 52–57. <https://doi.org/10.1016/j.tim.2010.11.002>
- Anesio, A.M., Laybourn-Parry, J., 2012. Glaciers and ice sheets as a biome. *Trends Ecol Evol* 27, 219–225. <https://doi.org/10.1016/j.tree.2011.09.012>
- Antoniades, D., Crawley, C., Douglas, M.S.V., Pienitz, R., Andersen, D., Doran, P.T., Hawes, I., Pollard, W., Vincent, W.F., 2007. Abrupt environmental change in Canada's northernmost lake inferred from fossil diatom and pigment stratigraphy. *Geophys Res Lett* 34, 2007GL030947. <https://doi.org/10.1029/2007GL030947>
- Bégin, P.N., Tanabe, Y., Kumagai, M., Culley, A.I., Paquette, M., Sarrazin, D., Uchida, M., Vincent, W.F., 2021a. Extreme warming and regime shift toward amplified variability in a far northern lake. *Limnol Oceanogr* 66. <https://doi.org/10.1002/lno.11546>
- Bégin, P.N., Tanabe, Y., Rautio, M., Wauthy, M., Laurion, I., Uchida, M., Culley, A.I., Vincent, W.F., 2021b. Water column gradients beneath the summer ice of a High Arctic freshwater lake as indicators of sensitivity to climate change. *Sci Rep* 11, 2868. <https://doi.org/10.1038/s41598-021-82234-z>
- Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Bondarenko, N.A., Belykh, O.I., Golobokova, L.P., Artemyeva, O.V., Logacheva, N.F., Tikhonova, I.V., Lipko, I.A., Kostornova, T.Ya., Parfenova, V.V., Khodzher, T.V., Ahn, T.-S., Zo, Y.-G., 2012. Stratified distribution of nutrients and extremophile biota within freshwater ice covering the surface of Lake Baikal. *J Microbiol.* 50, 8–16. <https://doi.org/10.1007/s12275-012-1251-1>
- Bourquin, M., Busi, S.B., Fodelianakis, S., Peter, H., Washburne, A., Kohler, T.J., Ezzat, L., Michoud, G., Wilmes, P., Battin, T.J., 2022. The microbiome of cryospheric ecosystems. *Nat Commun* 13, 3087. <https://doi.org/10.1038/s41467-022-30816-4>
- Bowman, J.P., 2017. Genomics of Psychrophilic Bacteria and Archaea, in: Margesin, R. (Ed.), *Psychrophiles: From Biodiversity to Biotechnology*. Springer International Publishing, Cham, pp. 345–387. https://doi.org/10.1007/978-3-319-57057-0_15
- Bradley, J.A., Trivedi, C.B., Winkel, M., Mourot, R., Lutz, S., Larose, C., Keuschnig, C., Doting, E., Halbach, L., Zervas, A., Anesio, A.M., Benning, L.G., 2022. Active and dormant microorganisms on glacier surfaces. *Geobiology* 21, 244–261. <https://doi.org/10.1111/gbi.12535>

- Busi, S.B., Bourquin, M., Fodelianakis, S., Michoud, G., Kohler, T.J., Peter, H., Pramateftaki, P., Styllas, M., Tolosano, M., De Staercke, V., Schön, M., de Nies, L., Marasco, R., Daffonchio, D., Ezzat, L., Wilmes, P., Battin, T.J., 2022. Genomic and metabolic adaptations of biofilms to ecological windows of opportunity in glacier-fed streams. *Nat Commun* 13, 2168. <https://doi.org/10.1038/s41467-022-29914-0>
- Butler, T.M., Wilhelm, A.-C., Dwyer, A.C., Webb, P.N., Baldwin, A.L., Techtmann, S.M., 2019. Microbial community dynamics during lake ice freezing. *Sci Rep* 9, 6231. <https://doi.org/10.1038/s41598-019-42609-9>
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13, 581–583. <https://doi.org/10.1038/nmeth.3869>
- Campbell, J.L., Mitchell, M.J., Groffman, P.M., Christenson, L.M., Hardy, J.P., 2005. Winter in northeastern North America: a critical period for ecological processes. *Front Ecol Environ* 3, 314–322. [https://doi.org/doi.org/10.1890/1540-9295\(2005\)003\[0314:winnaa\]2.0.co;2](https://doi.org/doi.org/10.1890/1540-9295(2005)003[0314:winnaa]2.0.co;2)
- Cavaco, M.A., St. Louis, V.L., Engel, K., St. Pierre, K.A., Schiff, S.L., Stibal, M., Neufeld, J.D., 2019. Freshwater microbial community diversity in a rapidly changing High Arctic watershed. *FEMS Microbiol Ecol* 95, fiz161. <https://doi.org/10.1093/femsec/fiz161>
- Cavaliere, E., Fournier, I.B., Hazuková, V., Rue, G.P., Sadro, S., Berger, S.A., Cotner, J.B., Dugan, H.A., Hampton, S.E., Lottig, N.R., McMeans, B.C., Ozersky, T., Powers, S.M., Rautio, M., O'Reilly, C.M., 2021. The lake ice continuum concept: Influence of winter conditions on energy and ecosystem dynamics. *J Geophys Res Biogeosci* 126, 1–20. <https://doi.org/10.1029/2020JG006165>
- Cazares, A., Mendoza-Hernández, G., Guarneros, G., 2014. Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. *BMC Genomics* 15, 1146. <https://doi.org/10.1186/1471-2164-15-1146>
- Christner, B.C., Mikucki, J.A., Foreman, C.M., Denson, J., Prisco, J.C., 2005. Glacial ice cores: A model system for developing extraterrestrial decontamination protocols. *Icarus* 174, 572–584. <https://doi.org/10.1016/j.icarus.2004.10.027>
- Ciok, A., Budzik, K., Zdanowski, M.K., Gawor, J., Grzesiak, J., Decewicz, P., Gromadka, R., Bartosik, D., Dziewit, L., 2018. Plasmids of psychrotolerant *Polaromonas* spp. isolated from Arctic and Antarctic glaciers – diversity and role in adaptation to polar environments. *Front Microbiol* 9, 1285. <https://doi.org/10.3389/fmicb.2018.01285>
- Collins, T., Margesin, R., 2019. Psychrophilic lifestyles: Mechanisms of adaptation and biotechnological tools. *Appl Microbiol Biotechnol* 103, 2857–2871. <https://doi.org/10.1007/s00253-019-09659-5>
- Comte, J., Culley, A.I., Lovejoy, C., Vincent, W.F., 2018. Microbial connectivity and sorting in a High Arctic watershed. *ISME J* 12, 2988–3000. <https://doi.org/10.1038/s41396-018-0236-4>
- Cruaud, P., Vigneron, A., Fradette, M.-S., Charette, S.J., Rodriguez, M.J., Dorea, C.C., Culley, A.I., 2017. Open the Sterivex™ casing: An easy and effective way to improve DNA extraction yields: DNA extraction from Sterivex™

- filters. *Limnol Oceanogr Methods* 15, 1015–1020.
<https://doi.org/10.1002/lom3.10221>
- Darcy, J.L., Lynch, R.C., King, A.J., Robeson, M.S., Schmidt, S.K., 2011. Global distribution of *Polaromonas* phylotypes - evidence for a highly successful dispersal capacity. *PLoS ONE* 6, e23742.
<https://doi.org/10.1371/journal.pone.0023742>
- Dinsdale, E.A., Edwards, R.A., Hall, D., Angly, F., Breitbart, M., Brulc, J.M., Furlan, M., Desnues, C., Haynes, M., Li, L., McDaniel, L., Moran, M.A., Nelson, K.E., Nilsson, C., Olson, R., Paul, J., Brito, B.R., Ruan, Y., Swan, B.K., Stevens, R., Valentine, D.L., Thurber, R.V., Wegley, L., White, B.A., Rohwer, F., 2008. Functional metagenomic profiling of nine biomes. *Nature* 452, 629–632. <https://doi.org/10.1038/nature06810>
- Dugan, H.A., Obryk, M.K., Doran, P.T., 2013. Lake ice ablation rates from permanently ice-covered Antarctic lakes. *J Glaciol* 59, 491–498.
<https://doi.org/10.3189/2013JoG12J080>
- Eddy, S.R., 2018. HMMER 3.2.1.
- Edgar, R.C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32, 1792–1797.
<https://doi.org/10.1093/nar/gkh340>
- Edwards, A., Pachebat, J.A., Swain, M., Hegarty, M., Hodson, A.J., Irvine-Fynn, T.D.L., Rassner, S.M.E., Sattler, B., 2013. A metagenomic snapshot of taxonomic and functional diversity in an alpine glacier cryoconite ecosystem. *Environ Res Lett* 8, 035003. <https://doi.org/10.1088/1748-9326/8/3/035003>
- Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., Delmont, T.O., 2015. Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ* 3, e1319. <https://doi.org/10.7717/peerj.1319>
- Eren, A.M., Kiefl, E., Shaiber, A., Veseli, I., Miller, S.E., Schechter, M.S., Fink, I., Pan, J.N., Yousef, M., Fogarty, E.C., Trigodet, F., Watson, A.R., Esen, Ö.C., Moore, R.M., Clayssen, Q., Lee, M.D., Kivenson, V., Graham, E.D., Merrill, B.D., Karkman, A., Blankenberg, D., Eppley, J.M., Sjödin, A., Scott, J.J., Vázquez-Campos, X., McKay, L.J., McDaniel, E.A., Stevens, S.L.R., Anderson, R.E., Fuessel, J., Fernandez-Guerra, A., Maignien, L., Delmont, T.O., Willis, A.D., 2020. Community-led, integrated, reproducible multi-omics with anvi'o. *Nat Microbiol* 6, 3–6. <https://doi.org/10.1038/s41564-020-00834-3>
- Feller, G., 2003. Molecular adaptations to cold in psychrophilic enzymes. *Cell Mol Life Sci* 60, 648–662. <https://doi.org/10.1007/s00018-003-2155-3>
- Feller, G., Gerday, C., 2003. Psychrophilic enzymes: Hot topics in cold adaptation. *Nat Rev Microbiol* 1, 200–208. <https://doi.org/doi:10.1038/nrmicro773>
- Gawor, J., Grzesiak, J., Sasin-Kurowska, J., Borsuk, P., Gromadka, R., Górniak, D., Świątecki, A., Aleksandrak-Piekarczyk, T., Zdanowski, M.K., 2016. Evidence of adaptation, niche separation and microevolution within the genus *Polaromonas* on Arctic and Antarctic glacial surfaces. *Extremophiles* 20, 403–413. <https://doi.org/10.1007/s00792-016-0831-0>
- Girard, C., Vincent, W.F., Culley, A.I., 2023. Arctic bacterial diversity and connectivity in the coastal margin of the Last Ice Area. *ISME Commun* 3, 105. <https://doi.org/10.1038/s43705-023-00313-w>

- Gouy, M., Tannier, E., Comte, N., Parsons, D.P., 2020. Seaview version 5: A multiplatform software for multiple sequence alignment, molecular phylogenetic analyses, and tree reconciliation. *Mol Biol Evol* 27, 221–224. <https://doi.org/10.1093/molbev/msp259>
- Gruber-Vodicka, H.R., Seah, B.K.B., Pruesse, E., 2020. phyloFlash: Rapid small-subunit rRNA profiling and targeted assembly from metagenomes. *mSystems* 5, e00920-20. <https://doi.org/10.1128/mSystems.00920-20>
- Halbach, L., Chevrollier, L.-A., Doting, E.L., Cook, J.M., Jensen, M.B., Benning, L.G., Bradley, J.A., Hansen, M., Lund-Hansen, L.C., Markager, S., Sorrell, B.K., Tranter, M., Trivedi, C.B., Winkel, M., Anesio, A.M., 2022. Pigment signatures of algal communities and their implications for glacier surface darkening. *Sci Rep* 12, 17643. <https://doi.org/10.1038/s41598-022-22271-4>
- Hampton, S.E., Galloway, A.W.E., Powers, S.M., Ozersky, T., Woo, K.H., Batt, R.D., Labou, S.G., O'Reilly, C.M., Sharma, S., Lottig, N.R., Stanley, E.H., North, R.L., Stockwell, J.D., Adrian, R., Weyhenmeyer, G.A., Arvola, L., Baulch, H.M., Bertani, I., Bowman, L.L., Carey, C.C., Catalan, J., Colom-Montero, W., Domine, L.M., Filip, M., Granados, I., Gries, C., Grossart, H., Haberman, J., Haldna, M., Hayden, B., Higgins, S.N., Jolley, J.C., Kahilainen, K.K., Kaup, E., Kehoe, M.J., MacIntyre, S., Mackay, A.W., Mariash, H.L., McKay, R.M., Nixdorf, B., Nöges, P., Nöges, T., Palmer, M., Pierson, D.C., Post, D.M., Pruett, M.J., Rautio, M., Read, J.S., Roberts, S.L., Rücker, J., Sadro, S., Silow, E.A., Smith, D.E., Sterner, R.W., Swann, G.E.A., Timofeyev, M.A., Toro, M., Twiss, M.R., Vogt, R.J., Watson, S.B., Whiteford, E.J., Xenopoulos, M.A., 2017. Ecology under lake ice. *Ecol Lett* 20, 98–111. <https://doi.org/10.1111/ele.12699>
- Hassan, N., Anesio, A.M., Rafiq, M., Holtvoeth, J., Bull, I., Haleem, A., Shah, A.A., Hasan, F., 2020a. Temperature driven membrane lipid adaptation in glacial psychrophilic bacteria. *Front Microbiol* 11, 824. <https://doi.org/10.3389/fmicb.2020.00824>
- Hassan, N., Anesio, A.M., Rafiq, M., Holtvoeth, J., Bull, I., Williamson, C.J., Hasan, F., 2020b. Cell membrane fatty acid and pigment composition of the psychrotolerant cyanobacterium *Nodularia spumigena* CHS1 isolated from Hopar glacier, Pakistan. *Extremophiles* 24, 135–145. <https://doi.org/10.1007/s00792-019-01141-4>
- Hattersley-Smith, G., 1963. The Ward Hunt Ice Shelf: Recent Changes of the Ice Front. *J Glaciol* 4, 415–424. <https://doi.org/10.3189/S0022143000027830>
- Hattersley-Smith, G., Crary, A.P., Christie, R.L., 1955. Northern Ellesmere Island, 1953 and 1954. *Arctic* 8, 2–36. <https://doi.org/10.14430/arctic3802>
- Herlemann, D.P.R., Labrenz, M., Jürgens, K., Bertilsson, S., Waniek, J.J., Andersson, A.F., 2011. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J* 5, 1571–1579. <https://doi.org/10.1038/ismej.2011.41>
- Huber, P., Cornejo-Castillo, F.M., Ferrera, I., Sánchez, P., Logares, R., Metz, S., Balagué, V., Acinas, S.G., Gasol, J.M., Unrein, F., 2019. Primer design for an accurate view of picocyanobacterial community structure by using high-throughput sequencing. *Appl Environ Microbiol* 85, e02659-18. <https://doi.org/10.1128/AEM.02659-18>

- Imbeau, E., Vincent, W.F., Wauthy, M., Cusson, M., Rautio, M., 2021. Hidden stores of organic matter in northern lake ice: Selective retention of terrestrial particles, phytoplankton and labile carbon. *J Geophys Res Biogeosci* 126. <https://doi.org/10.1029/2020JG006233>
- Jungblut, A.D., Lovejoy, C., Vincent, W.F., 2010. Global distribution of cyanobacterial ecotypes in the cold biosphere. *ISME J* 4, 191–202. <https://doi.org/10.1038/ismej.2009.113>
- Kämpfer, P., Rainey, F.A., Andersson, M.A., Nurmiäho Lassila, E.L., Ulrych, U., Busse, H.J., Weiss, N., Mikkola, R., Salkinoja-Salonen, M., 2000. *Frigoribacterium faeni* gen. nov., sp. nov., a novel psychrophilic genus of the family *Microbacteriaceae*. *Int J Syst Evol Microbiol* 50, 355–363. <https://doi.org/10.1099/00207713-50-1-355>
- Kans, J., 2013. Entrez Direct: E-utilities on the Unix Command Line.
- Katlein, C., Arndt, S., Nicolaus, M., Perovich, D.K., Jakuba, M.V., Suman, S., Elliott, S., Whitcomb, L.L., McFarland, C.J., Gerdes, R., Boetius, A., German, C.R., 2015. Influence of ice thickness and surface properties on light transmission through Arctic sea ice. *JGR Oceans* 120, 5932–5944. <https://doi.org/10.1002/2015JC010914>
- Kim, E.H., Jeong, H.-J., Lee, Y.K., Moon, E.Y., Cho, J.-C., Lee, H.K., Hong, S.G., 2011. *Actimicrobium antarcticum* gen. nov., sp. nov., of the family *Oxalobacteraceae*, isolated from Antarctic coastal seawater. *Curr Microbiol* 63, 213–217. <https://doi.org/10.1007/s00284-011-9962-9>
- Kivilä, E.H., Prėskienis, V., Gaudreault, N., Girard, C., Rautio, M., 2023. Variability in lake bacterial growth and primary production under lake ice: Evidence from early winter to spring melt. *Limnol Oceanogr* Ino.12447. <https://doi.org/10.1002/lno.12447>
- Klassen, J.L., Foght, J.M., 2011. Characterization of *Hymenobacter* isolates from Victoria Upper Glacier, Antarctica reveals five new species and substantial non-vertical evolution within this genus. *Extremophiles* 15, 45–57. <https://doi.org/10.1007/s00792-010-0336-1>
- Korver, M.C., Lehner, B., Cardille, J.A., Carrea, L., 2024. Surface water temperature observations and ice phenology estimations for 1.4 million lakes globally. *Remote Sens Environ* 308, 114164. <https://doi.org/10.1016/j.rse.2024.114164>
- Lehnerr, I., St. Louis, V.L., Sharp, M., Gardner, A.S., Smol, J.P., Schiff, S.L., Muir, D.C.G., Mortimer, C.A., Michelutti, N., Tarnocai, C., St. Pierre, K.A., Emmerton, C.A., Wiklund, J.A., Köck, G., Lamoureux, S.F., Talbot, C.H., 2018. The world’s largest High Arctic lake responds rapidly to climate warming. *Nat Commun* 9, 1290. <https://doi.org/10.1038/s41467-018-03685-z>
- Leppäranta, M., 2010. Modelling the Formation and Decay of Lake Ice, in: George, G. (Ed.), *The Impact of Climate Change on European Lakes*. Springer Netherlands, Dordrecht, pp. 63–83. https://doi.org/10.1007/978-90-481-2945-4_5
- Li, D., Luo, R., Liu, C.-M., Leung, C.-M., Ting, H.-F., Sadakane, K., Yamashita, H., Lam, T.-W., 2016. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102, 3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>

- Liu, Y., Ji, M., Yu, T., Zaugg, J., Anesio, A.M., Zhang, Z., Hu, S., Hugenholtz, P., Liu, K., Liu, P., Chen, Y., Luo, Y., Yao, T., 2022. A genome and gene catalog of glacier microbiomes. *Nat Biotechnol* 40, 1341–1348. <https://doi.org/10.1038/s41587-022-01367-2>
- Magnuson, J.J., Robertson, D.M., Benson, B.J., Wynne, R.H., Livingstone, D.M., Arai, T., Assel, R.A., Barry, R.G., Card, V., Kuusisto, E., Granin, N.G., Prowse, T.D., Stewart, K.M., Vuglinski, V.S., 2000. Historical trends in lake and river ice cover in the northern hemisphere. *Science* 289, 1743–1746. <https://doi.org/10.1126/science.289.5485.1743>
- McGinnis, S., Madden, T.L., 2004. BLAST: At the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* 32, W20–W25. <https://doi.org/10.1093/nar/gkh435>
- McMurdie, P.J., Holmes, S., 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* 8, e61217. <https://doi.org/10.1371/journal.pone.0061217>
- Menzel, P., Ng, K.L., Krogh, A., 2016. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat Commun* 7, 11257. <https://doi.org/10.1038/ncomms11257>
- Michelutti, N., Douglas, M.S.V., Antoniadou, D., Lehnherr, I., St. Louis, V.L., St. Pierre, K., Muir, D.C.G., Brunskill, G., Smol, J.P., 2020. Contrasting the ecological effects of decreasing ice cover versus accelerated glacial melt on the High Arctic’s largest lake. *Proc R Soc B* 287. <https://doi.org/10.1098/rspb.2020.1185>
- Miteva, V., 2008. Bacteria in Snow and Glacier Ice, in: Margesin, R., Schinner, F., Marx, J.-C., Gerday, C. (Eds.), *Psychrophiles: From Biodiversity to Biotechnology*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 31–50. https://doi.org/10.1007/978-3-540-74335-4_3
- Mohit, V., Culley, A., Lovejoy, C., Bouchard, F., Vincent, W.F., 2017. Hidden biofilms in a far northern lake and implications for the changing Arctic. *npj Biofilms Microbiomes* 3, 17. <https://doi.org/10.1038/s41522-017-0024-3>
- Nam, G.G., Joung, Y., Park, M., Kim, S., Jeon, H.T., Cho, J.-C., 2017. *Flavobacterium soyangense* sp. nov., a psychrotolerant bacterium, isolated from an oligotrophic freshwater lake. *Int J Syst Evol Microbiol* 67, 2440–2445. <https://doi.org/10.1099/ijsem.0.001987>
- NEIGE, 2023. Dynamics of ice cover over a far northern lake : Direct observations of Ward Hunt Lake, Canadian High Arctic, by automated camera., v. 1.2.0 (2016-2022). <https://doi.org/10.5885/45648CE-1A9AB63DFF91440B>
- Newton, R., Pfirman, S., Tremblay, L.B., DeRepentigny, P., 2021. Defining the “ice shed” of the Arctic Ocean’s Last Ice Area and its future evolution. *Earth’s Future* 9. <https://doi.org/10.1029/2021EF001988>
- Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O’Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlenn, D., Ouellette, M.-H., Ribeiro Cunha, E., Smith, T., Stier, A., Ter Braak, C.J.F., Weedon, J.,

2001. vegan: Community Ecology Package.
<https://doi.org/10.32614/CRAN.package.vegan>
- Paquette, M., Fortier, D., Mueller, D.R., Sarrazin, D., Vincent, W.F., 2015. Rapid disappearance of perennial ice on Canada's most northern lake. *Geophys Res Lett* 42, 1433–1440. <https://doi.org/10.1002/2014GL062960>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 41, D590–D596. <https://doi.org/10.1093/nar/gks1219>
- Quesada, A., Vincent, W.F., 2012. Cyanobacteria in the Cryosphere: Snow, Ice and Extreme Cold, in: *Ecology of Cyanobacteria II*. Springer, Dordrecht, pp. 387–399.
- Raymond-Bouchard, I., Whyte, L.G., 2017. From Transcriptomes to Metatranscriptomes: Cold Adaptation and Active Metabolisms of Psychrophiles from Cold Environments, in: Margesin, R. (Ed.), *Psychrophiles: From Biodiversity to Biotechnology*. Springer International Publishing, Cham, pp. 437–457. https://doi.org/10.1007/978-3-319-57057-0_18
- Robinson, A.L., Ariano, S.S., Brown, L.C., 2021. The influence of snow and ice albedo towards improved lake ice simulations. *Hydrology* 8, 11. <https://doi.org/10.3390/hydrology8010011>
- Rogers, S.O., Theraisnathan, V., Ma, L.J., Zhao, Y., Zhang, G., Shin, S.-G., Castello, J.D., Starmer, W.T., 2004. Comparisons of protocols for decontamination of environmental ice samples for biological and molecular examinations. *Appl Environ Microbiol* 70, 2540–2544. <https://doi.org/10.1128/AEM.70.4.2540-2544.2004>
- Ruuskanen, M.O., Colby, G., St.Pierre, K.A., St.Louis, V.L., Aris-Brosou, S., Poulain, A.J., 2020. Microbial genomes retrieved from High Arctic lake sediments encode for adaptation to cold and oligotrophic environments. *Limnol Oceanogr* 65. <https://doi.org/10.1002/lno.11334>
- Santibáñez, P.A., Michaud, A.B., Vick-Majors, T.J., D'Andrilli, J., Chiuchiolo, A., Hand, K.P., Priscu, J.C., 2019. Differential incorporation of bacteria, organic matter, and inorganic ions into lake ice during ice formation. *J Geophys Res Biogeosci* 124, 585–600. <https://doi.org/10.1029/2018JG004825>
- Sharma, S., Blagrove, K., Magnuson, J.J., O'Reilly, C.M., Oliver, S., Batt, R.D., Magee, M.R., Straile, D., Weyhenmeyer, G.A., Winslow, L., Woolway, R.I., 2019. Widespread loss of lake ice around the Northern Hemisphere in a warming world. *Nat Clim Chang* 9, 227–231. <https://doi.org/10.1038/s41558-018-0393-5>
- Tran, P., Ramachandran, A., Khawasik, O., Beisner, B.E., Rautio, M., Huot, Y., Walsh, D.A., 2018. Microbial life under ice: Metagenome diversity and *in situ* activity of Verrucomicrobia in seasonally ice-covered Lakes. *Environ Microbiol* 20, 2568–2584. <https://doi.org/10.1111/1462-2920.14283>
- Varliero, G., Holland, A., Barker, G.L.A., Yallop, M.L., Fountain, A.G., Anesio, A.M., 2021. Glacier clear ice bands indicate englacial channel microbial distribution. *J Glaciol* 67, 811–823. <https://doi.org/10.1017/jog.2021.30>

- Veillette, J., Martineau, M.-J., Antoniadou, D., Sarrazin, D., Vincent, W.F., 2010. Effects of loss of perennial lake ice on mixing and phytoplankton dynamics: insights from High Arctic Canada. *Ann Glaciol* 51, 56–70. <https://doi.org/10.3189/172756411795931921>
- Verpoorter, C., Kutser, T., Seekell, D.A., Tranvik, L.J., 2014. A global inventory of lakes based on high-resolution satellite imagery. *Geophys Res Lett* 41, 6396–6402. <https://doi.org/10.1002/2014GL060641>
- Vigneron, A., Cruaud, P., Lovejoy, C., Vincent, W.F., 2022. Genomic evidence of functional diversity in DPANN archaea, from oxic species to anoxic vampiristic consortia. *ISME Commun* 2, 4. <https://doi.org/10.1038/s43705-022-00088-6>
- Vincent, W.F., 2010. Microbial ecosystem responses to rapid climate change in the Arctic. *ISME J* 4, 1087–1090. <https://doi.org/10.1038/ismej.2010.108>
- Vincent, W.F., Fortier, D., Lévesque, E., Boulanger-Lapointe, N., Tremblay, B., Sarrazin, D., Antoniadou, D., Mueller, D.R., 2011. Extreme ecosystems and geosystems in the Canadian High Arctic: Ward Hunt Island and vicinity. *Écoscience* 18, 236–261. <https://doi.org/10.2980/18-3-3448>
- Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., Venables, B., 2022. *gplots* 3.1.3.
- Wyka, S., Mondo, S., Liu, M., Nalam, V., Broders, K., 2022. A large accessory genome and high recombination rates may influence global distribution and broad host range of the fungal plant pathogen *Claviceps purpurea*. *PLoS ONE* 17, e0263496. <https://doi.org/10.1371/journal.pone.0263496>
- Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., Whitman, W.B., Euzéby, J., Amann, R., Rosselló-Móra, R., 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol* 12, 635–645. <https://doi.org/10.1038/nrmicro3330>
- Zakharova, Y., Bashenkhaeva, M., Galachyants, Y., Petrova, D., Tomberg, I., Marchenkov, A., Kopyrina, L., Likhoshway, Y., 2022. Variability of microbial communities in two long-term ice-covered freshwater lakes in the subarctic region of Yakutia, Russia. *Microb Ecol* 84, 958–973. <https://doi.org/10.1007/s00248-021-01912-7>
- Zhong, Z.-P., Tian, F., Roux, S., Gazitúa, M.C., Solonenko, N.E., Li, Y.-F., Davis, M.E., Van Etten, J.L., Mosley-Thompson, E., Rich, V.I., Sullivan, M.B., Thompson, L.G., 2021. Glacier ice archives nearly 15,000-year-old microbes and phages. *Microbiome* 9, 160. <https://doi.org/10.1186/s40168-021-01106-w>
- Zhou, Y., Hiller, C., Andersson, S., Jakobsson, E., Zhou, L., Hawkes, J.A., Kothawala, D.N., Tranvik, L.J., 2023. Selective exclusion of aromatic organic carbon during lake ice formation. *Geophys Res Lett* 50, e2022GL101414. <https://doi.org/10.1029/2022GL101414>

ANNEXE 1: Supplementary material (Chapitre 2)

Table S1. Mean length of sequenced reads, the number of forward (R1) and reverse (R2) reads per sample and the number of contigs obtained after assembly using Megahit. The co-assembly uses the total reads from all eight samples.

Samples	Mean length	Reads R1	Reads R2	Contigs
Top A	658.5	2 167 670	2 266 259	181 024
Top B	542.6	303 296	319 063	39 483
Middle A	639.8	2 703 614	2 887 242	187 768
Middle C	527.9	640 815	679 784	66 588
Bottom B	528	622 485	665 354	65 222
Bottom C	489.1	36 300	38 521	4 997
Water A	619.3	1 716 844	1 793 168	174 499
Water B	550.6	2 467 308	2 570 649	213 492
Co-assembly	643.6			757 362

Table S2. Total number of raw reads per amplicon sequencing samples and the resulting number of reads before and after DADA2 processing.

Sample	Sequencing	Input	Final
Top A	RNA	55313	26424
Top B	RNA	66736	36292
Top C	RNA	77256	52181
Middle A	RNA	60022	24677
Middle B	RNA	83664	5879
Middle C	RNA	68415	28475
Bottom A	RNA	67147	44843
Bottom B	RNA	74375	52754
Bottom C	RNA	68776	43958
Water A	RNA	54616	28835
Water B	RNA	65989	41807
Water C	RNA	58232	10683
Top A	DNA	71698	7082
Top B	DNA	61827	6948
Top C	DNA	76978	10006
Middle A	DNA	62111	6045
Middle B	DNA	62018	8019
Middle C	DNA	66062	8776
Bottom A	DNA	49496	11599
Bottom B	DNA	55144	8982
Water A	DNA	50584	8624
Water B	DNA	63457	9142
Water C	DNA	70455	10553

Table S3. Psychrophilic genes, consensus sequence lengths, HMMER expected matches and whether BLASTx correctly identified our best HMMER hits or not. “NoMatch” means HMMER did not find the gene in our contigs.

Gene	Consensus sequence length	HMMER- match	BLASTx	Retained for downstream analyses
Synthesis of polyunsaturated fatty acids				
<i>pfaA</i>	454	7.10E-05	YES	NO
<i>pfaB</i>	750	NoMatch	NA	NO
<i>pfaC</i>	879	2.30E-13	NO	NO
<i>pfaD</i>	1560	NoMatch	NA	NO
<i>pfaE</i>	673	NoMatch	NA	NO
Solute uptake system				
<i>proV</i>	821	2.30E-110	YES	YES
<i>proW</i>	923	1.70E-104	YES	YES
<i>proX</i>	914	2.80E-207	YES	YES
<i>opuA</i>	1136	1.80E-74	YES	YES
<i>opuE</i>	1481	1.50E-72	NO	NO
<i>opuB</i>	NA	NA	NA	NA
<i>opuC</i>	NA	NA	NA	NA
<i>opuD</i>	1136	2.00E-79	YES	YES
<i>ectP</i>	2721	1.50E-51	YES	YES
<i>lpoP</i>	791	NoMatch	NA	NA
<i>betP</i>	1290	3.00E-129	YES	YES
<i>betA</i>	1650	2.10E-179	YES	YES
<i>betB</i>	587	1.40E-56	YES	YES
<i>gbsA</i>	1475	5.10E-64	YES	YES
<i>gbsB</i>	1211	4.00E-18	YES	YES
<i>otsA</i>	874	5.10E-65	YES	YES
<i>otsB</i>	871	9.20E-77	YES	YES
Protection against toxic oxidative reaction products				
<i>katE</i>	853	8.80E-189	YES	YES
<i>katG</i>	2409	9.10E-232	YES	YES
<i>katN</i>	887	3.30E-39	NO	NO
<i>btuE</i>	561	3.10E-39	YES	YES
<i>efeB</i>	1357	2.00E-38	YES	YES
<i>sodN</i>	398	8.40E-05	NO	NO
<i>sodA</i>	747	4.40E-82	YES	YES
<i>sodB</i>	780	3.60E-113	YES	YES
<i>sodC</i>	581	1.00E-77	YES	YES
Provide flexibility to tRNA and stabilise ribosomes against cold denaturation				
<i>cshA</i>	872	1.50E-12	NO	NO
<i>cshB</i>	NA	NA	NA	NA

<i>rhlB</i>	1283	2.70E-14	NO	NO
<i>srmB</i>	1099	1.70E-104	YES	YES
<i>cspA</i>	1080	1.90E-17	YES	YES

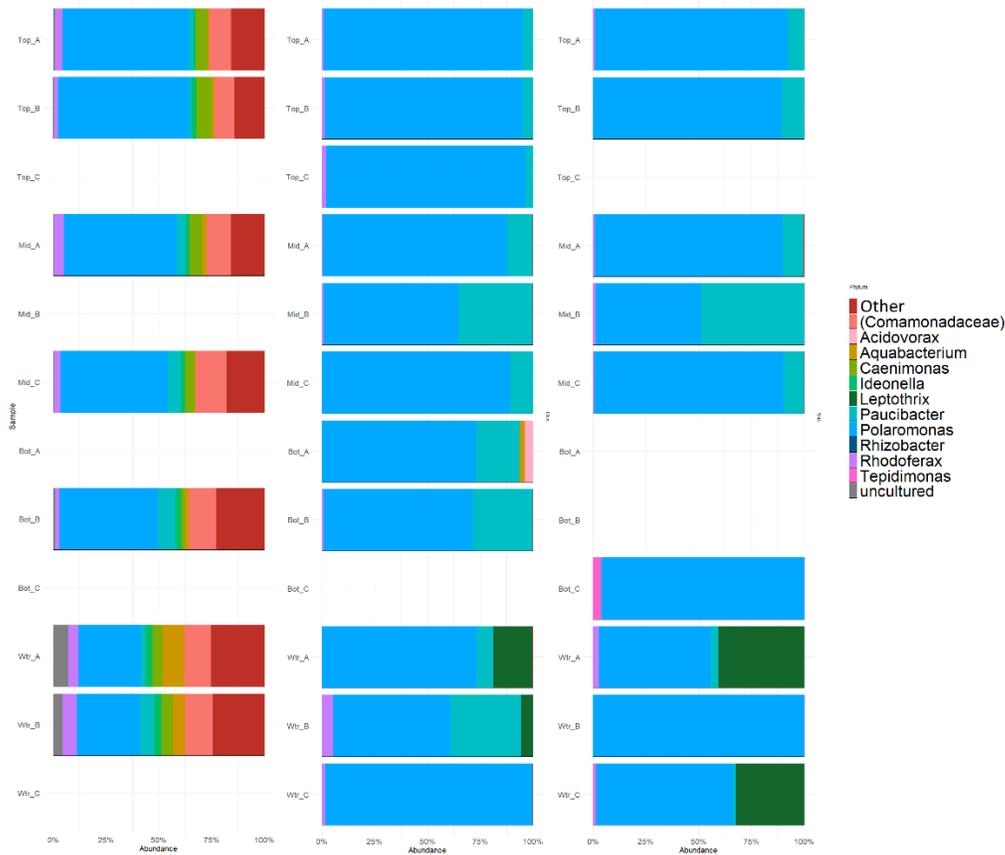


Figure S1. Relative abundance of *Commamonadaceae* in ice and water below ice from triplicate ice cores for **A.** shotgun metagenomic sequencing, **B.** amplicon sequencing of the 16S rDNA gene and **C.** amplicon sequencing of 16S rRNA transcripts (cDNA). Gaps show samples for which sequencing yielded too few reads (**Table S1**). Genera representing less than 0.01% (shotgun sequencing) and 0.01% (amplicon sequencing) of relative abundance were combined into “Others”.

ANNEXE 2: Quantification de l'ADN dans la glace de lacs boréaux et arctique

En préparation à la campagne arctique de l'été 2022, une autre campagne d'échantillonnage hivernal préparatoire a eu lieu en février 2022 sur les lacs boréaux Hertel (45.54 -73.15, Mont-Saint-Hilaire), Croche (45.99 -74.00, Saint-Hippolyte) et Simoncouche (48.23 -71.25, Réserve faunique des Laurentides). Ces échantillons ont été maintenus au congélateur à -20°C jusqu'à ce qu'ils puissent être traités comme indiqué dans le **Chapitre 2.2.2**. Une PERMANOVA a permis de déterminer si la géographie (boréale vs arctique) avait un effet sur les quantités d'ADN extraites des lacs boréaux par rapport au lac Ward Hunt (**Tableau S4**). Nous constatons que les quantités d'ADN sont significativement plus faibles dans les échantillons de glace boréale que dans la glace arctique du lac Ward Hunt ($p = 0.02556$).

Tableau S4. Quantification de l'ADN dans les échantillons après extraction pour trois lacs boréaux et le lac Ward Hunt en Arctique. La limite de détection pour la quantification de l'ADN selon la trousse utilisée est de 0.005 ng/µl, pour un minimum de 0.1 ng d'ADN dans l'échantillon. Les valeurs non détectables (en deçà de ce seuil) ont été indiquées comme telles dans le tableau.

ID	Lac	Milieu	Concentration (ng/µl)
SIM-A-Top	Simoncouche	Boréal	0.046
SIM-A-Middle	Simoncouche	Boréal	0.026
SIM-B-Top	Simoncouche	Boréal	0.02
SIM-B-Middle	Simoncouche	Boréal	Indétectable
SIM-C-Top	Simoncouche	Boréal	0.0292
SIM-C-Middle	Simoncouche	Boréal	0.0296
CHR-A-Top	Croche	Boréal	Indétectable
CRH-A-Middle	Croche	Boréal	0.0182
CRH-A-Bottom	Croche	Boréal	Indétectable
CRH-B-Top	Croche	Boréal	0.0228
CHR-B-Middle	Croche	Boréal	0.262
CRH-B-Middle	Croche	Boréal	0.142
CRH-C-Top	Croche	Boréal	Indétectable
CHR-C-Middle	Croche	Boréal	0.0522
CHR-C-Bottom	Croche	Boréal	Indétectable
HTL-A-Top	Hertel	Boréal	0.0712
HTL-A-Middle	Hertel	Boréal	Indétectable
HTL-A-Bottom	Hertel	Boréal	Indétectable
HTL-B-Top	Hertel	Boréal	0.0114
HTL-B-Middle	Hertel	Boréal	0.102
HTL-B-Bottom	Hertel	Boréal	0.0442
HTL-C-Top	Hertel	Boréal	0.0176
HTL-C_Middle	Hertel	Boréal	Indétectable
HTL-C-Bottom	Hertel	Boréal	Indétectable
WHL-A-Top	Ward Hunt	Arctique	0.872
WHL-A-Middle	Ward Hunt	Arctique	1.68
WHL-A-Bottom	Ward Hunt	Arctique	Indétectable
WHL-B-Top	Ward Hunt	Arctique	0.68
WHL-B-Middle	Ward Hunt	Arctique	0.0758
WHL-B-Bottom	Ward Hunt	Arctique	0.0798
WHL-C-Top	Ward Hunt	Arctique	0.0658
WHL-C-Middle	Ward Hunt	Arctique	0.0864
WHL-C-Bottom	Ward Hunt	Arctique	0.013