THESIS

Presented at:

ÉCOLE CENTRALE DE LYON

For the title of:

Doctor

of doctoral school: Électronique, électrotechnique, automatique

UNIVERSITÉ DE LYON

Speciality: Bioengineering

By

Rose Layton

Microbial community structure and function across the seawater-sea ice-frost flower-snow-atmosphere continuum

Pr. Lyle White: Canada Research Chair in Polar Microbiology, Professor - Department of Natural Resources Sciences– McGill University	**Reviewer**
Pr. Jill Mikucki: Assistant Professor of Microbiology – Department of Microbiology- University of Tennessee	**Reviewer**
Pr Mohamed Jebbar : Enseignant-chercheur - Laboratoire de Microbiologie des Environnements - Université de Brest	**External member of the jury**
Dr Jeff Bowman: Assistant Professor- Integrative Oceanography Division- Scripps Institution of Oceanography	**External member of the jury**
Dr. Pascal Simonet: professeur à Groupe de Génomique Microbienne Environnementale - Ecole Centrale de Lyon - Lyon (France)	**President of the jury**
Dr. Catherine Larose: in charge of research (CNRS) at Environmental Microbial Genomics Group - Ecole Centrale de Lyon - Lyon (France)	**Thesis advisor**
Pr. Timothy M. Vogel : professor at Environmental Microbial Genomics Group - Ecole Centrale de Lyon - Lyon (France)	**Co-supervisor**

THÈSE

Présentée devant :

ÉCOLE CENTRALE DE LYON Pour obtenir le grade de :

DOCTEUR

de l'école doctorale Électronique, électrotechnique, automatique

UNIVERSITÉ DE LYON

Spécialité : Ingénierie du vivant

par

Rose Layton

Structure et fonction de la communauté microbienne à travers le continuum eau de mer-glace de glace-fleur-neige-atmosphère

Pr. Lyle White: Canada Research Chair in Polar Microbiology, Professor - Department of Natural Resources Sciences– McGill University	**Reviewer**
Pr. Jill Mikucki: Assistant Professor of Microbiology – Department of Microbiology- University of Tennessee	**Reviewer**
Pr Mohamed Jebbar : Enseignant-chercheur - Laboratoire de Microbiologie des Environnements - Université de Brest	**External member of the jury**
Dr Jeff Bowman: Assistant Professor- Integrative Oceanography Division- Scripps Institution of Oceanography	**External member of the jury**
Dr. Pascal Simonet: professeur à Groupe de Génomique Microbienne Environnementale - Ecole Centrale de Lyon - Lyon (France)	**President of the jury**
Dr. Catherine Larose: in charge of research (CNRS) at Environmental Microbial Genomics Group - Ecole Centrale de Lyon - Lyon (France)	**Thesis advisor**
Pr. Timothy M. Vogel : professor at Environmental Microbial Genomics Group - Ecole Centrale de Lyon - Lyon (France)	**Co-supervisor**

Acknowledgements

The people who made this thesis possible are numerous and I express my sincere gratitude to everyone who has contributed scientifically or otherwise. It is very difficult to express my thanks to my supervisors but in as few words as possible I'd like to thank Catherine for being the most fantastic and optimistic supervisor, mentor, unpaid therapist and friend. I am certain that without your support, this would not have been possible. I'd would also like to thank Tim, whose perpetual supply of ideas is equal parts invaluable and maddening but softened by the supply of good wine.

I would like to acknowledge all the members of Laboratoire Ampere; especially Laure for ensuring the smooth operation of the lab, Christoph for his instrumental knowledge and support in sequencing and coffee-making, Pascal for his endless bibliography and Edith and Marie Christine for limiting my involvement in bureaucracy! I have to thank all members of my office; Benoit, Mia, Romie, Adrien and Conchi. A special thanks to Benoit for displaying so much patience for my bioinformatics deficit; it has surely improved. Thanks to Adrien and Conchi (the queen!) who supplied beer and a lot of laughs and Mia and Romie for their willingness to help. Friday night beers with Graeme, Christina and Linda have kept my mind sharp – so cheers to you guys.

I have to say a massive thank you to all the people I collaborated with. I was lucky enough to start my PhD with a network of fabulous people under the umbrella of MicroArctic and you all have been inspiring in your own unique ways! I have to quickly mention Alex Holland for an unforgettable field season in Ny Alesund; it really would not have been the same without you… and of course Lucie Malard for always wanting to adventure when no-one else would! A special thanks to Liane Benning as I was originally connected with this group because of you! I have to thank everyone at IPEV for facilitating my trip to Svalbard and the community at Ny Alesund for being so welcoming. I'd also like to thank my other collaborators including James France and Jan Kaiser for letting us microbiologists into the clean world of your sea ice chamber and those in Grenoble for running the chemistry.

My friends have been a massive support system and kept my head above water especially during a difficult third year so thank you guys. I have a special thanks to all my climbing

buddies by helping me live this escape. I can’t possibly list you all for fear of missing someone out but you’re all amazing!

Finally, I could not ask for a more supportive family. So, for possibly the one and only time in your lifetime, here is a sincere thank you with written evidence for my brother; you are more important than you know. Nothing parallels the support and love I feel from my parents and for that I will be forever indebted to you; so thank you both and I owe this to you. Last but not least, I want to thank dogs, wine and rock climbing for making the world a better place.

Résumé

La glace de mer est un écosystème important et étendu (34 million km^2) et joue divers rôles importants sur le plan écologique, allant d'un habitat pour les micro- et macro-organismes à l'impact sur le budget énergétique global. La glace de mer subit des changements drastiques en raison des changements climatiques. Ces changements ont une influence sur le type de glace dominant des régions polaires (de MYI à FYI) et, par conséquent, sur l'expansion des environnements associés tels que la neige saline et les fleurs de givre. Bien que ces écosystèmes soient connus pour abriter une grande variété de microorganismes, notre compréhension des conséquences des changements climatiques sur ces communautés est limitée. L'objectif principal de cette thèse était d'évaluer la colonisation microbienne, l'adaptation et la dynamique des communautés en relation avec la formation de la glace de mer et le développement des habitats à différentes échelles spatiales et temporelles. Pour y parvenir, des techniques de séquençage de nouvelle génération ont été utilisées pour interroger taxonomiquement et fonctionnellement des modèles de laboratoire et des modèles naturels de glace de mer, de neige et de fleurs de givre.

Le premier objectif portait sur la manière dont les populations eucaryotes et bactériennes s'établissaient et se répartissaient dans la glace de mer et les fleurs de givre. Pour éviter les interférences des conditions dominantes, par exemple de la géographie locale et les courants océaniques, une chambre de glace de mer expérimentale a été utilisée. La répartition taxonomique et fonctionnelle des bactéries en environnements distincts (fleur de givre, saumure, matrice de glace et eau de mer) a permis de soutenir l'hypothèse d'une sélection basée sur les niches et a été renforcée par l'apparition de bactéries de glace de mer archétypales, non identifiées dans l'eau de mer d'ensemencement. En revanche, la cohérence des communautés eucaryotes dans l'ensemble du profil a permis d'identifier l'influence de la stochasticité. Nos données suggèrent que l'enrichissement bactérien dans la glace de mer se produit à un rythme similaire à celui observé dans la nature, malgré une représentation minimale des diatomées. Ceci contraste avec une hypothèse largement diffusée qui propose que l'enrichissement bactérien est lié à l'attachement aux diatomées. Notre deuxième objectif était d'identifier les facteurs environnementaux qui ont influencé la sélection des communautés dans la glace de mer et la neige. Si l'assemblage des communautés microbiennes est déterminé par les conditions dominantes, les adaptations correspondantes

devraient se refléter dans les métagénomes. En effet, des signatures fonctionnelles telles que la motilité, la chimiotaxie, la réponse aux UV et le transfert de gènes ont toutes été trouvées en plus grande abondance dans la neige saline et certains horizons glaciaires par rapport à l'eau de mer, présentant ainsi des facteurs liés à la structuration des communautés.

Un dernier objectif était d'étudier le potentiel taxonomique et fonctionnel des fleurs de givre et de la neige saline. Dans l’étude en microcosme (fleurs de givre) et les expériences de terrain (neige saline), un enrichissement spécifique de certains taxons a été observé. Les *Colwellia* et les *Glaciecola* étaient très largement dominants dans les environnements de neige saline, indépendamment des facteurs temporels ou spatiaux. Les génomes assemblés métagénomiquement (MAGs) ont suggéré que les mécanismes de réparation liés aux UV et la capacité à utiliser la lumière pour la croissance par des pompes à protons de la protéorhodopsine pourraient constituer des éléments qui expliquent leur succès dans ces habitats.

Abstract

Sea ice is an important and extensive (34 million km^2) ecosystem and has a diverse array of ecologically important roles; from providing a habitat to micro- and macro- organisms to impacting the global energy budget. The sea ice landscape is experiencing drastic changes as a result of climate change. These changes can be seen in the shift of dominant ice type of Polar Regions (from MYI to FYI) and consequently the expansion of associated environments such as brine wetted snow and frost flowers. Although these ecosystems are known to host a diverse array of microorganisms, our understanding of the consequences of climate change on these communities is limited. The overarching goal of this thesis was to assess microbial colonisation, community adaptation and dynamics in relation to sea ice formation and the developing habitats at different spatial and temporal scales. In order to achieve this, next generation sequencing techniques were employed to taxonomically and functionally interrogate laboratory and natural sea ice, snow and frost flower models.

The first objective addressed how eukaryotic and bacterial populations established and partitioned in young sea ice and frost flowers. To avoid interference from prevailing conditions, local geography and ocean currents, a sea ice chamber microcosm setup was utilised. The taxonomic and functional partitioning of bacteria into distinct niches (frost flower, brine, ice matrix and seawater) supported the hypothesis of niche based selection and was bolstered by the appearance of archetypal sea ice bacteria, not identified in the seeding seawater. In contrast, the consistency of eukaryotic communities across the profile identified the influence of stochasticity. Our data suggested that bacterial enrichment in sea ice occurred at a similar rate as previously observed, despite minimal diatom representation. This stands in contrast to a widely circulated hypothesis that proposes bacterial enrichment is linked to diatom attachment. Our second objective was to identify the environmental factors that influenced community selection across sea ice and snow communities. If microbial community assembly is driven by the prevailing conditions, corresponding adaptations should be reflected in the metagenomes. Indeed, functional signatures including motility, chemotaxis, response to UV and genetic transfer were all found in greater abundance in the saline snow and some ice horizons relative to the seawater, presenting possible drivers of functional potential.

A final objective was to survey the taxonomic and functional potential of frost flowers and saline snow. In both microcosm (frost flowers) and field experiments (saline snow), a taxa-specific enrichment was observed. *Colwellia* and *Glaciecola* were overwhelmingly dominant in brine-fed snow environments, independent of temporal or spatial factors. Metagenomic assembled genomes (MAGs) suggested that UV repair mechanisms and the ability to gain a growth advantage from light using proteorhodopsin proton pumps may underscore their success in these habitats.

Table of Contents

Index of figures

Chapter 2

Chapter 3

Chapter 4

between clusters is indicated by height bar. Colours represent surface (purple) or basal snow from each sample site (blue=site A, green=site B, purple=site C and red=site D).

Figure 4.3: Stacked bar charts of the relative abundance of ASVs at class level. Samples are grouped by site and type

Figure 4.4: Venn diagram summarizing the number of shared and unique GO annotations. The left-most Venn diagram includes all ice core samples, the right-most compares site A and B ice core and brine samples.

Figure 4.5: Hierarchical cluster using ward clustering method and correlation as distance measure and 1000 bootstrap resampling based on relative abundance of all ASVs detected. Bootstrap probability (BP) values in percentage are indicated in green and AU pvaues are indicated in green. AU >= 95% are considered to be strongly supported by the data. Distance between clusters is indicated by height bar. Colours represent surface (purple) or basal snow from each sample site (red=site A, blue=site B, yellow=site C and green=site D).

Figure 4.6: Hierarchical cluster using ward clustering method and correlation as distance measure and 1000 bootstrap resampling based on relative abundance of all ASVs detected. Bootstrap probability (BP) values in percentage are indicated in green and AU pvalues are indicated in green. AU >= 95% are considered to be strongly supported by the data. Distance between clusters is indicated by height bar. Colours represent ice (purple), brine (green) or saline snow from each sample site (blue=site A, red=site B).

Figure 4.7: Venn diagram summarizing the number of shared and unique GO annotations of all basal snow horizons.

Chapter 5

Figure 5.1: Coverage estimates for (A) *Glaciecola* sp. and (B) *Colwellia* sp. MAGs

Figure 5.2: (A) Glaciecola and (B) *Colwellia* pangenome analysis of cultivar genomes and draft genomes recovered from saline snow metagenomes. Bars in the first 9 layers indicate the occurrence of a gene cluster in a given isolate genome. The next layers (blue) represent the number of genomes that contain the gene clusters, the number of genes in the gene cluster and the number of paralogs. The next layer, shows whether the gene clusters are single copy genes identified using the anvio function 'anvi-run-hmms'. The next 3 layers indicate the geometric, functional and combined homogeneity indices. The penultimate layer indicates whether a COG function could be assigned and the final layer corresponds to groups of gene clusters; red – gene clusters identified in only the saline snow *Colwellia* or *Glaciecola* MAG, purple

Figure 5.3: Organisation of contig "salinesnow__000000311699" encoding (from left to right) alcohol dehydrogenase, dimethylallyl diphosphate, beta-carotene 15,15'-dioxygenase, lycopene cyclase, phytoene synthase, phytoene, polyprenyl synthetase, bacteriorhodopsin and an ATPase. The heatmap displays the coverage values of this contig across all metagenomes

Index of supplementary figures

Chapter 2

Chapter 3

Chapter 4

Index of tables

Chapter 1

Table 1.1: Examples of sea ice sequencing-based studies

Chapter 4

Table 4.1: Differentially abundant ASVs (daASVs) and GO terms (daGOs) as identified by DESeq2 with a padj <0.05 and a log2foldChange > 1. Average proportional abundances of all enriched daASVs and daGOs are given in percentages. Numbers in parenthesis represent values for associated brine samples, when collected.

Table 4.2: Results of significance testing of r-squared transformed relative abundance data annotated at different levels using the "adonis" function from the R-package Vegan and grouping factor was either "site" or "sample type".

Table 4.3: Differentially abundant ASVs (daASVs) and GO terms (daGOs) as identified by DESeq2 with a padj <0.05 and a log2foldChange > 1. Average proportional abundances of all enriched daASVs and daGOs are given in percentages

Table 4.4: Results of significance testing of r-squared transformed relative abundance data annotated at different levels using the "adonis" function from the R-package Vegan and grouping factor was either "site" or "sample type"

Chapter 5

Table 5.1: Statistics of metagenomics assembled genomes resolved in this study

Index of supplementary tables

Chapter 2

Table S2.1: In-tank instrumentation installed and utilised during sea ice growth in the sea ice chamber

Table S2.2: Primers used in this study. Sequences in bold represent illumina adapter sequences.

Table S2.3A: ASV table 16S rRNA from Dada2 pipeline

Table S2.3B: ASV table 18S rRNA from Dada2 pipeline

Table S2.4A: Genera significantly (padj < 0.05) differentially abundant between source (seawater) and seed (all other) environments determined by DeSeq2 analysis

Table S2.4A: Genera significantly (padj < 0.05) differentially abundant between source (frost flowers) and seed (all other) environments determined by DeSeq2 analysis

Table S2.5: Relative abundance of reads annotated as Bacteria, Archaea or Eukarya in origin in sea ice chamber metagenomes

Chapter 3

Chapter 4

Chapter 5

Table S5.4: Functional annotations of gene calls and associated counts for cultivar Colwellia genomes and the saline snow Colwellia MAG

Table S5.5: Functional annotations of gene calls and associated counts for cultivar Glaciecola genomes and the saline snow Glaciecola MAG

Table S5.6: Annotations of gene calls associated with singleton gene clusters in the Colwellia MAG

Table S5.7: Annotations of gene calls associated with singleton gene clusters in the Glaciecola MAG

Chapitre 1 : Bibliographie - Écologie de la succession microbienne des glaces de mer

1.1 Introduction

Les microorganismes contiennent des membres de tous les domaines de la vie (par exemple, des bactéries, des champignons, des virus, des protozoaires, des algues). Ces micro-organismes forment des communautés avec des interactions complexes qui sont à l'origine de nombreux processus importants à l'échelle mondiale, tels que la biogéochimie (Bertrand et al., 2015), la pathogénie (Chow, Tang, & Mazmanian, 2011 ; Rogers, Hoffman, Carroll, & Bruce, 2013 ; Vayssier-Taussat et al, 2014), la symbiose (Beinart, 2019 ; Kamagata & Narihiro, 2016 ; Raina, Fernandez, Lambert, Stocker, & Seymour, 2019 ; Wein et al., 2019) et la bioremédiation (Paliwal, Puranik, & Purohit, 2012 ; Pandey, Chauhan, & Jain, 2009). Leur contribution à un si large éventail de processus est en partie due à leur capacité à coloniser n'importe quel habitat, même le plus inhospitalier (Bull & Goodfellow, 2019 ; Horikoshi & Bull, 2011). Bien que la colonisation par les microorganismes soit un processus continu, la colonisation primaire déterminera la manière dont les communautés microbiennes s'établiront dans de nouveaux environnements et régira la distribution, la composition et le développement ultérieur de la communauté. Il est clair que la façon dont les micro-organismes se trient dans les premiers stades de la formation d'une communauté est fondamentale pour l'écologie microbienne. Alors que les processus de colonisation ont été largement étudiés en culture (Ben-Jacob, Cohen, & Levine, 2000 ; Burrows, Elbert, Lawrence, & Pöschl, 2009 ; Dann, Clanahan, Paterson, & Mitchell, 2019 ; Grossart, Kiørboe, Tang, & Ploug, 2003 ; Hallatschek & Nelson, 2010 ; Kerr, Neuhauser, Bohannan, & Dean, 2006 ; Kerr, Riley, Feldman, & Bohannan, 2002 ; Lloyd & Allen, 2015 ; Venegas-Ortiz, Allen, & Evans, 2014), notre compréhension de leur fonctionnement dans les communautés naturelles est mal connue. La glace de mer est un excellent modèle pour élucider les processus de colonisation et l'adaptation des communautés microbiennes, étant donné sa formation rapide, son caractère saisonnier et ses multiples niches habitables avec des gradients physico-chimiques importants.

Les étapes de la colonisation peuvent être divisées en plusieurs grandes catégories. Dans un premier temps, les micro-organismes doivent être transférés passivement ou activement dans la nouvelle niche environnementale. Par conséquent, des facteurs tels que la diversité et la structure de la communauté source peuvent influencer la diversité et la structure de la communauté puits (Lindström & Langenheder, 2012 ; Zhou & Ning, 2017). Les adaptations innées de populations spécifiques (par exemple la chimiotaxie) peuvent également contribuer à leur capacité à créer de nouveaux habitats (Josenhans & Schweinitzer, 2010 ; Raina et al., 2019). Une fois présents, les pionniers microbiens seront soumis à de nouveaux paramètres physico-chimiques. Les pionniers qui réussissent doivent surmonter des pressions de sélection potentiellement nouvelles, telles que des gradients différents de température, de salinité, de lumière et de nutriments (Lindström & Langenheder, 2012). Cela conduit à l'exclusion environnementale de certaines espèces qui sont incapables de survivre une fois présentes (par exemple, lysées en raison de différences de pressions osmotiques) (Zhou & Ning, 2017). En outre, de nouvelles interactions intra-espèces (concurrence, coopération et mutualisme) peuvent favoriser l'établissement précoce de la structure et de la diversité des communautés. À long terme, la concurrence peut entraîner soit (1) une exclusion compétitive, où une population domine et l'autre est menacée d'extinction, (2) un cloisonnement métabolique des niches, où les concurrents peuvent coexister parce qu'ils se spécialisent dans différents types de ressources, soit (3) un cloisonnement physique des niches, où les concurrents se dispersent dans différentes niches territoriales (par exemple dans des environnements semi-solides) (Bauer, Kainz, Carmona-Gutierrez et Madeo, 2018).

Les habitats ne sont souvent pas stables dans l'espace et dans le temps, et les perturbations peuvent donc entraîner une succession secondaire des communautés microbiennes (Fierer, Nemergut, Knight, & Craine, 2010). Par exemple, les changements saisonniers de température de la glace de mer peuvent entraîner une dynamique microbienne différente en hiver et en été (Eric Collins, Rocap, & Deming, 2010 ; Luria, Amaral-Zettler, Ducklow, & Rich, 2016 ; Schmidt et al., 2007 ; Wu et al., 2016). Les environnements connectés qui font office de banques de semences microbiennes peuvent également être confrontés à des changements et peuvent potentiellement fournir de nouvelles espèces concurrentes. Il est difficile de démêler les facteurs qui déterminent les différentes étapes de la succession microbienne dans les environnements naturels en raison de l'imprévisibilité des paramètres environnementaux changeants, des diverses échelles spatio-temporelles auxquelles ces paramètres agissent et de la complexité des interactions microbiennes entre elles et avec leur environnement physicochimique (Zhou & Ning, 2017).

En raison des changements climatiques, nous assistons à une transition de la prédominance de la glace pluriannuelle (MYI) à la nouvelle glace de mer (FYI) dans les régions polaires (Kwok, 2018). Non seulement la FYI augmente en importance spatiale et temporelle, mais elle représente un modèle unique pour l'étude de l'écologie de la succession microbienne. Lorsque les températures baissent en automne, la formation de la FYI est couplée à une colonisation microbienne primaire (Vonnahme, Dietrich, & Hassett, 2020). En raison de l'exclusion du sel pendant la formation de la glace, il se forme un environnement semi-solide constitué d'une matrice de glace en phase solide perméable à un réseau de canaux de saumure liquide dense, hypersalin et riche en nutriments et matière organique (OM) (Hunke, Notz, Turner, & Vancoppenolle, 2011). Cette séparation spatiale offre des possibilités de cloisonnement de niches tant physiques que métaboliques. En outre, les conditions environnementales dominantes de la glace de mer nouvellement formée créent un environnement sélectif fort pour les pionniers microbiens permettant l'exploration des processus de filtrage environnemental (Eronen-Rasimus et al., 2015). La réduction de l'espace habitable de la saumure renforce l'interaction et facilite ainsi l'étude de l'exclusion compétitive par rapport au cloisonnement métabolique et de niche. Comme les régions polaires ont la saisonnalité la plus marquée du globe, les communautés microbiennes de la glace de mer sont soumises à de fortes perturbations (Vonnahme et al., 2020). Nous pouvons utiliser cette occurrence naturelle pour comprendre les schémas de succession à long terme. Les précipitations et les courants océaniques reconstituent les apports microbiens dans le système de glace de mer et permettent ainsi de savoir si la diversité et la structure des communautés sources contribuent à la dynamique des communautés microbiennes dans l'environnement du puits. Les précipitations, telles que les chutes de neige, ajoutent de nouvelles couches au profil de la glace de mer, ce qui fournit un autre habitat colonisable.

Bien que la microbiologie de la glace de mer ait été largement étudiée (Boetius, Anesio, Deming, Mikucki et Rapp, 2015 ; J. P. Bowman, 2013 ; Jeff S. Bowman, 2015 ; Deming et Eric Collins, 2016a ; Marcela Ewert et Deming, 2013 ; Vonnahme et al., 2020), nous nous concentrons dans ce chapitre sur la façon dont la littérature disponible nous aide à comprendre les processus de colonisation dans les écosystèmes naturels. Nous établissons comment les processus stochastiques et déterministes contribuent à la colonisation primaire de la glace de mer et comment les facteurs qui régissent la succession microbienne peuvent changer à des échelles temporelles et spatiales différentes. Nos objectifs sont de comprendre comment les relations inter et intra-domaines, les interactions microbiennes avec leur environnement physico-chimique et la diversité et la structure de l'environnement d'origine impactent l'établissement et le développement de la glace de mer et de ses habitats associés. Enfin, nous abordons l'importance de comprendre la dynamique des facteurs de succession microbienne dans les communautés microbiennes de la glace de mer pour faire des prévisions futures sur les effets de la fonte de la glace de mer et la rapidité avec laquelle les outils de séquençage et de bioinformatique en développement facilitent cette orientation de l'étude.

1.2 La colonisation primaire de la glace de mer

La glace de mer de première année se forme en fonction de la baisse automnale de la température. Au départ, de petits cristaux de glace frasil à flottabilité inefficace se forment dans l'eau dont la température est inférieure à 0 degré Celsius. Leur circulation dans la partie supérieure de la colonne d'eau, due aux courants océaniques, favorise leur adhérence aux particules de l'eau. Ces cristaux remontent ensuite à la surface au fur et à mesure de leur croissance, ce qui donne une couche appelée "glace grasse" (Omstedt, 1985). La glace grasse se consolide lorsque la température baisse, formant une jeune glace de mer qui est composée d'une matrice de glace entrelacée avec un réseau de saumure ; ce dernier peut contenir une grande abondance de micro-organismes (Deming & Eric Collins, 2016a ; S Grossmann & Dieckmann, 1994). La capacité d'un organisme à se disperser dans la jeune glace de mer pour la coloniser, est déterminée à la fois par des voies passives et actives, impliquant des facteurs tels que le filtrage environnemental, la motilité et les interactions biotiques. Il est bien connu que les bactéries et les algues sont enrichies dans la glace de mer par rapport à l'eau de mer sous-jacente, mais les mécanismes qui sous-tendent leur enrichissement ne sont pas clairs. Dans cette section, nous examinons les connaissances actuelles sur la façon dont les micro-organismes sont enrichis dans la glace de mer et sur les raisons de cet enrichissement.

1.2.1 Glace de frasil, associations d'algues et de bactéries

Garrison et al. (1983) ont proposé que les algues pouvaient coloniser les FYI pendant leur formation grâce à l'activité de balayage des cristaux de glace de frasil, en se basant sur des observations dans la mer de Weddell (David L. Garrison, Ackley, & Buck, 1983). Des expériences en laboratoire ont ensuite démontré la capacité de la glace de frasil à récolter et à concentrer les cellules d'algues en suspension (D. L. Garrison, Close, & Reimnitz, 1989). Dans les études de terrain, il a été démontré que la formation de glace retenait préférentiellement les grandes cellules (R. Gradinger & Ikävalko, 1998 ; Kauko et al., 2018 ; Riedel, Michel, Gosselin, & LeBlanc, 2007a). En outre, la formation expérimentale de la glace de mer suggère un enrichissement bactérien négligeable en l'absence de cellules d'algues (Weissenberger & Grossmann, 1998). Cela a conduit à l'hypothèse selon laquelle l'enrichissement des bactéries observé dans la glace de mer pourrait être le résultat de leur capacité à s'associer aux algues (S. Grossmann, 1994 ; S Grossmann & Dieckmann, 1994 ; Sönnke Grossmann & Gleitz, 1993 ; Weissenberger & Grossmann, 1998).

Bien que les bactéries et les algues soient connues pour leur association dans un cadre marin, les recherches traitant ce sujet dans la glace de mer sont limitées. En outre, l'hypothèse selon laquelle l'attachement des bactéries aux grandes cellules eucaryotes est incompatible avec l'observation d'une faible colonisation bactérienne des diatomées pélagiques (Crenn, Duffieux, & Jeanthon, 2018 ; Droop & Elson, 1966 ; Kaczmarska et al., 2005). Une étude portant à la fois sur les domaines bactérien et eucaryote dans la glace de mer a démontré que l'attachement des bactéries aux diatomées était faible et qu'aucune relation directe entre l'enrichissement bactérien et l'enrichissement autotrophe n'a été observée (Riedel et al., 2007a). De plus, les rares données provenant de la glace de mer hivernale de l'Antarctique montrent que la biomasse algale est un mauvais indicateur de la biomasse bactérienne (Stewart & Fritsen, 2004). D'autres preuves contradictoires démontrent un enrichissement bactérien dans la glace de mer artificielle, équivalent à celui observé précédemment, malgré une abondance extrêmement faible d'algues (Eronen-Rasimus et al., 2014). Collectivement, ces recherches suggèrent que si les cellules d'algues peuvent être entraînées dans la glace par un processus physique sélectif en fonction de leur taille, les bactéries sont enrichies par des moyens différents.

1.2.2 Fixation des particules, polysaccharides extracellulaires (EPS) et protéines de liaison à la glace (IBP)

L'entraînement des particules dans la glace de mer est conforme à l'hypothèse selon laquelle les particules en suspension dans l'eau de mer peuvent agir comme un système de livraison pour les cellules attachées (Lindemann, Hölemann, Korablev, & Zachek, 1999). L'observation d'un nombre élevé de cellules associées à des particules dans la glace de mer corrobore cette hypothèse, bien qu'il ne soit pas clair que des associations de particules se soient formées dans l'environnement pélagique, ce qui entraîne un biais dans le potentiel de colonisation observé (Karen Junge, Eicken, & Deming, 2004). En retour, la capacité des micro-organismes à s'attacher aux particules pourrait être un facteur clé de la colonisation primaire de la glace de mer.

Le polysaccharide extracellulaire (EPS) dans la glace de mer a fait l'objet de nombreuses études. Les EPS sont des polymères de structure diverse sécrétés par un large éventail d'organismes, dont de nombreuses bactéries et algues (Marcela Ewert & Deming, 2013). Outre son rôle dans la cryoconservation, les EPS peuvent favoriser la formation d'agrégats en augmentant la "viscosité" d'un organisme. Les estimations de l'adhésivité des exopolymères ont montré qu'elle était de 2 à 4 ordres de grandeur plus élevée que celle des autres particules (Passow, 2002). Cela entraînerait une augmentation des organismes producteurs d'EPS dans la glace de mer. Des concentrations élevées d'EPS dans la saumure par rapport à l'eau de mer sous-jacente ont été régulièrement signalées (Meiners, Gradinger, Fehling, Civitarese et Spindler, 2003 ; Riedel, Michel, Gosselin et LeBlanc, 2007b ; Graham J.C. Underwood et al., 2013). En outre, l'attachement des bactéries aux exopolymères a été directement observé dans des environnements pélagiques (Passow & Alldredge, 1994) et benthiques (G. J. C. Underwood, Paterson, & Parkes, 1995). Cependant, un lien direct entre le taux d'entraînement et les micro-organismes producteurs de EPS reste à démontrer.

Les protéines de nucléation de la glace (INP) et les protéines antigel (AFP) sont des protéines de liaison à la glace (IBP) qui peuvent faire partie du EPS sécrété par les micro-organismes marins adaptés au froid, notamment les bactéries, les champignons et les algues. Les INP sont capables d'induire la formation de glace à des températures inférieures à zéro (-5 à -12°C) plus élevées que les processus homogènes qui ont été estimés à - 42,2 ±0,3°C dans l'eau de mer artificielle (K. Junge & Swanson, 2008). La formation de cristaux de glace extracellulaires augmente la tolérance au gel en générant de la chaleur, qui peut ensuite être utilisée par l'organisme. Les AFP se lient aux cristaux de glace, ce qui peut aider à maintenir l'accès à l'oxygène et aux nutriments, ainsi qu'à protéger la cellule en servant de médiateur pour la croissance des plus petits cristaux, empêchant ainsi la recristallisation (Białkowska, Majewska, Olczak, & Twarda-Clapa, 2020). Par conséquent, la production de ces protéines dans les micro-organismes de la glace de mer peut favoriser la poursuite de l'activité et la survie à des températures autrement contraignantes. Bien que le rôle protecteur des protéines IBP dans les microorganismes de la glace de mer ait été examiné ailleurs, par exemple (Białkowska et al., 2020 ; Ewart, Lin, & Hew, 1999), le potentiel de ces protéines à aider la colonisation primaire de la glace de mer n'a pas fait l'objet de nombreuses études. Comme indiqué ci-dessus, l'EPS, qui contient potentiellement des fractions de protéines liant la glace, est de préférence entraîné dans la glace de mer. Collectivement, ces études fournissent des preuves circonstancielles de l'entraînement sélectif de certains micro-organismes en fonction de leur capacité à produire et à sécréter des protéines qui interagissent avec les cristaux de glace. Toutefois, il n'a pas encore été démontré si les organismes produisant du EPS et des protéines de glace sont entraînés dans la structure de la glace plus facilement que d'autres ou si les concentrations élevées de EPS observées dans la glace de mer sont dues à la réaction des organismes à la baisse de température et à l'augmentation de la salinité.

1.2.3 Vacuoles de gaz

Les vacuoles gazeuses sont constituées de structures protéiques intracellulaires creuses qui s'agglomèrent en grandes organelles et se trouvent presque exclusivement dans les bactéries aquatiques (Anthony E. Walsby, 1981). Le gonflement ou le dégonflement des vésicules de gaz à l'intérieur des cellules microbiennes assure la flottabilité (A E Walsby, 1972) et permet à son tour à l'organisme de flotter à une profondeur souhaitée dans la colonne d'eau. Ainsi, les vésicules gazeuses permettent la sélection des conditions favorables à la croissance, en particulier dans les habitats à stratification verticale photique, thermique ou chimique (John J. Gosink, Irgens et Staley, 1993). Les vacuoles gazeuses peuvent agir comme une alternative au flagelle pour la motilité, et ont été trouvées presque exclusivement chez des espèces sans flagelles. L'identification de vacuoles de gaz chez les bactéries occupant la partie supérieure de la colonne d'eau de l'Antarctique a conduit à l'hypothèse qu’elles jouent un rôle dans la colonisation de la glace de mer ou même à un mouvement au sein du système de glace de mer (J. J. Gosink & Staley, 1995). Cette hypothèse a été confirmée par la proportion plus élevée de bactéries vacuolées dans la glace (165 sur 177 isolats) par rapport à l'eau de mer libre sous-jacente (12 sur 177 isolats) identifiée dans une étude sur la glace de mer arctique (John J. Gosink et al., 1993). Depuis la découverte initiale, il a été démontré qu'un grand nombre de taxons archétypaux de la glace de mer produisent des vacuoles de gaz, notamment les protéobactéries alpha, bêta et gamma et le groupe des Flavobactéries-Cytophaga (J. J. Gosink & Staley, 1995). La question de savoir si l'incorporation de certaines espèces bactériennes dans la glace de mer est une conséquence directe de la possession de vacuoles de gaz n'a pas été étudiée.

1.2.4 Taxis

Pour réussir dans des environnements hétérogènes comme la glace de mer, les organismes doivent soit adapter leur métabolisme aux conditions locales, soit migrer et coloniser de nouvelles niches (Alexandre, 2010). La colonisation d'une niche donnée dépendra de la capacité de l'organisme à obtenir des nutriments et à résister aux mécanismes antagonistes des autres membres de la communauté (Kennedy, 1987). La capacité à percevoir les changements dans leur environnement immédiat par un certain nombre de réactions est appelée "taxes" et elle est définie comme les réactions de mouvement directionnel à la suite de stimuli externes. Cette capacité de perception a été proposée comme un avantage sélectif dans l'exploitation de nouvelles niches, en particulier dans les environnements non mixtes où des gradients chimiques spatiaux ou temporels persistent (Stocker & Seymour, 2012). De nombreuses bactéries sont capables de détecter des changements dans leur environnement chimique et de diriger leur mouvement de manière appropriée dans un processus connu sous le nom de chimiotaxie. Cette stratégie est considérée comme particulièrement importante pour les micro-organismes d'eau de mer afin de faciliter une réponse rapide aux nutriments éphémères dans un océan généralement dilué (Stocker & Seymour, 2012). Par conséquent, la formation de la glace de mer et de la saumure dense en nutriments attirera les micro-organismes chimiotactiques qui cherchent à exploiter une niche favorable. Il a été suggéré que les bactéries oligotrophes non mobiles dominent lorsque la productivité est faible (Smriga, Fernandez, Mitchell, & Stocker, 2016). Inversement, les copiotrophes mobiles dominent dans des conditions riches en nutriments. Par conséquent, on peut s'attendre à une forte proportion de copiotrophes mobiles dans la saumure pendant la formation de la glace de mer. En effet, les analyses métagénomiques de la glace de mer groenlandaise ont montré une prévalence plus élevée des gènes liés à la motilité cellulaire et à la copiotrophie dans les saumures par rapport à l'eau de mer sous-jacente (Maccario, Carpenter, Deming, Vogel, & Larose, 2019). En outre, des preuves expérimentales ont montré que le micro-organisme représentatif de la glace de mer, *Colwellia psychrerythraea*, est capable de se déplacer à des températures aussi basses que -10°C (Karen Junge, Eicken, & Deming, 2003) et qu'il présente une chimiohalotaxie (Showalter & Deming, 2018). Le mouvement des cellules bactériennes dans des conditions quasi *in situ* a été détecté à l'aide d'un microscope holographique numérique. La motilité en réponse à d'autres gradients environnementaux pourrait également mettre en évidence la colonisation primaire de la glace de mer. Par exemple, dans une stratégie thigmotatique (réponse au stimulus du contact), la bactérie antarctique *Marinomonas primoryensis* utilise la motilité et une protéine de liaison à la glace pour adhérer aux cristaux de glace à la surface de l'océan et rester dans un environnement relativement riche en oxygène et en nutriments (Guo et al., 2017). La détection et le mouvement vers la lumière (phototaxie), le pH (pH-taxie) et l'oxygène (aérotaxie) représentent tous des stratégies utilisées par les micro-organismes qui vivent dans l'océan (Stocker & Seymour, 2012) et qui pourraient affecter l'assemblage microbien dans la glace de mer, mais qui n'ont pas encore été explorées.

1.2.5 Stratégies multiples de colonisation de la glace de mer

La manière dont les micro-organismes sont transférés à la glace de mer nouvellement formée est probablement une combinaison de processus déterministes et stochastiques. Par exemple, des individus peuvent s'attacher à des agrégats de glace frasil par contact fortuit et à la merci de turbulences océaniques aléatoires. Cependant, les micro-organismes peuvent augmenter leurs chances d'incorporation par des processus déterministes tels que l'attachement des algues, la production de EPS, les taxis et la flottabilité. En retour, cela peut affecter le tri des communautés dans la glace de mer en ce qui concerne le cloisonnement des niches physiques et métaboliques. Par exemple, la possession de protéines de liaison à la glace peut entraîner un piégeage dans la matrice de glace, la motilité pourrait faciliter la rétention dans les canaux de saumure liquide ou les bactéries vacuolées pourraient contribuer à la stratification verticale des communautés. Toutefois, si le cloisonnement physique des niches dans la glace de mer est un phénomène connu, il n'a été pris en compte qu'en termes d'abondance des organismes (Marcela Ewert & Deming, 2013). On ne sait pas exactement comment les différents membres de la communauté se répartissent selon différentes résolutions fonctionnelles et taxonomiques, mais aussi comment les mécanismes de transfert peuvent être un moteur structurel clé dans ce processus.

1.3 Environnements source et puits

Dans le contexte de la formation de la glace de mer, l'eau de mer agit comme la principale source de micro-organismes, tandis que la glace de mer peut être qualifiée de puits. Des conditions ou événements météorologiques spécifiques peuvent créer des environnements sources et puits supplémentaires qui contribuent au profil de la glace de mer. L'augmentation de l'étendue de la neige saline et des fleurs de givre qui s'accumulent à la surface de la glace de mer (Barber et al., 2014a ; Maccario et al., 2019) correspond à l'augmentation prévue de l'étendue de l'FYI. Il s'agit de milieux nouveaux qui restent sous-explorés mais qui sont de plus en plus importants en termes de communautés microbiennes et de conséquences biogéochimiques (J. S. Bowman & Deming, 2010 ; J. S. Bowman, Larose, Vogel, & Deming, 2013a ; Maccario et al., 2019). En outre, ils présentent de nouveaux environnements de colonisation secondaire, ce qui ajoute un intérêt à l'utilisation du profil de la glace de mer comme modèle de succession microbienne.

1.3.1 Eau de mer

Une partition taxonomique distincte a été mise en évidence en laboratoire comme un facteur majeur au cours du processus de colonisation, malgré l'hétérogénéité taxonomique du matériel d'origine (Dann et al., 2019). Si cela était transposable aux systèmes naturels de glace de mer, les glaces de mer nouvellement colonisables présenteraient une structure de communauté taxonomique distincte malgré des sources variables. Le stade initial de colonisation pendant la formation de la glace de mer n'a pas été bien documenté ; cependant, une étude a démontré que les premières communautés de glace en crêpes (pancake ice) étaient similaires à celles des communautés d'eau libre, ce qui suggère que l'eau de mer de source est un facteur clé dans la structuration des communautés à ce stade (Eronen-Rasimus et al., 2015). Inversement, une fois que la glace de mer se consolide et arrive à maturité, les recherches montrent systématiquement que les communautés qu'elle abrite contrastent avec l'eau de mer sous-jacente et sont dominées par des taxons de glace de mer archétypaux (Boetius et al., 2015 ; Deming & Eric Collins, 2016a ; Eronen-Rasimus et al., 2015 ; S. Grossmann, 1994). La composition taxonomique et fonctionnelle des communautés de glace de mer présente plus de similitudes entre elles qu'avec l'eau de mer sous-jacente, indépendamment des grandes distances géographiques, comme le montre une étude biogéographique printanière couvrant l'Arctique canadien (Yergeau et al., 2017). L'ensemble de ces études suggère que la taxonomie et l'assemblage fonctionnel sont régis par des mécanismes de sélection conférés par l'environnement de la glace de mer ; cette sélection pourrait être attribuée à toute combinaison de facteurs physico-chimiques (par exemple, la lyse par choc osmotique) et d'interactions biologiques (par exemple, l'exclusion compétitive). Ces résultats se sont traduits par des expériences en mésocosme où la formation de glace de mer artificielle, utilisant de l'eau de mer qui n'est généralement pas soumise à la formation annuelle de glace de mer, a entraîné l'apparition de bactéries archétypales de la glace de mer qui n'ont pas été détectées dans le matériau d'origine (Eronen-Rasimus et al., 2014). Ainsi, il est suggéré que si l'eau de mer de source peut être une variable structurante majeure au cours des toutes premières étapes de la formation de la glace de mer, les proportions relatives des populations bactériennes dans la source ne confondent pas l'émergence d'une communauté spécifique à la glace de mer.

1.3.2 Neige et atmosphère

Dans l'Arctique, la majeure partie de la surface de la glace de mer, estimée à 7× 10^6 km^2 en hiver (Nghiem et al., 2007), est couverte de neige qui s'accumule entre l'automne et le printemps (Sturm, Holmgren, & Perovich, 2002 ; Warren et al., 1999). Non seulement la neige a été récemment reconnue comme un écosystème microbien important, mais le réchauffement des régions polaires pourrait entraîner une plus grande contribution de la masse de glace de mer sous forme de neige métamorphique (Granskog, Vihma, Pirazzini, & Cheng, 2006). Sur une échelle de temps plus courte, les chutes de neige modifient la physico-chimie de la glace de mer, par exemple par l'extinction des UV, et fournissent une nouvelle source de populations microbiennes (Maccario et al., 2019). Bien que le transfert des micro-organismes de la neige à la glace de mer n'ait pas été étudié, en partie en raison de la biomasse minimale qu'elle contient par rapport aux saumures de la glace de mer, il existe un potentiel de colonisation et de croissance des micro-organismes de la neige dans la glace de mer. En effet, ces microorganismes peuvent être adaptés à la vie dans un environnement glacé car de nombreuses contraintes environnementales sont partagées. Par exemple, les micro-organismes de la glace de mer, de la neige et de l'atmosphère sont soumis à de forts gradients environnementaux en termes d'UV, de température et de disponibilité des nutriments

(Maccario, Sanguino, Vogel et Larose, 2015). Le peu de données disponibles sur la structure des communautés de l'atmosphère et de la neige recouvrant la glace de mer limite notre capacité actuelle à déchiffrer leurs contributions relatives à la succession des communautés de glace de mer.

Par rapport au MYI, le FYI contient une plus grande fraction de saumure qui peut être rejetée à la surface, apportant des nutriments, des sels et des micro-organismes à la neige qui la couvre (Barber, Reddan, & LeDrew, 1995). Ce processus se traduit par une couche de neige saline qui devient un environnement de plus en plus important en fonction de la perte d'épaisseur de la glace de mer et du passage du MYI au FYI (diminution du MYI de 50% depuis 2002, le MYI couvre maintenant moins d'un tiers de l'océan Arctique) (Kwok, 2018). La glace de première année domine maintenant la couverture de glace, représentant près de 80 % de la banquise de mars 2018 (D. Perovich et al., 2018). Les propriétés physiques et chimiques de la neige saline ont été étudiées de manière relativement approfondie (par exemple Domine et al., 2004 ; Huang & Jaeglé, 2017 ; Peterson et al., 2019 ; Yang et al., 2008) et pourtant seuls deux articles ont examiné son contenu microbien (M. Ewert, Carpenter, Colangelo-Lillis, & Deming, 2013 ; Maccario et al., 2019). Néanmoins, la neige saline peut agir comme un environnement de puits dans les profils de glace de mer tandis que la saumure représente un afflux de micro-organismes compétitifs. Les premières études sur la neige saline montrent des abondances élevées de bactéries et de EPS par rapport à la neige peu saline. Bien que les premiers travaux aient suggéré que des abondances bactériennes élevées étaient attribuées aux organismes provenant de la saumure, Maccario et al. (2019) ont émis l'hypothèse que la supplémentation en nutriments stimulait la croissance des micro-organismes de la neige. Cette hypothèse est cohérente avec l'observation selon laquelle les micro-organismes retrouvés dans la saumure n'étaient pas répandus dans la neige salée. La même étude a mis en évidence le manque de diversité présent dans la neige saline par rapport aux systèmes sous et sur-jacents couplés à un enrichissement spécifique en certains taxons. La question de savoir si l'assemblage des communautés dans la neige saline est cohérent dans le temps et dans l'espace n'a pas été étudiée, mais fournirait des informations utiles concernant un écosystème de plus en plus important à mesure que le climat change.

1.3.3 Fleurs de givre

Les fleurs de givre sont des structures délicates qui se forment presque exclusivement à la surface de la jeune glace de mer. Si l'atmosphère qui recouvre la glace de mer est saturée en eau et qu'il existe un gradient de température suffisant, des cristaux de glace se forment sur les imperfections de la surface de la glace (Florent Domine, 2005). Ceux-ci agissent en tant que points de nucléation à partir desquels les fleurs de gel se développent en structures dendritiques pouvant atteindre plusieurs centimètres de hauteur. Leurs tiges creuses s'écoulent par mèche à la surface de la saumure, ce qui crée un environnement extrêmement hypersalin (>120 ppt) (Barber et al., 2014b). De même, les bactéries sont transportées de la saumure aux fleurs de givre à des densités étonnamment élevées (3 à 6 fois plus que dans la glace sous-jacente) (Bowman & Deming, 2010). Cette découverte a alimenté quelques recherches qui ont donné un premier aperçu de la dynamique structurelle des communautés bactériennes des fleurs de givre (Barber et al., 2014b ; J. S. Bowman et al., 2013a ; Jeff S. Bowman, Berthiaume, Armbrust, & Deming, 2014 ; Mortazavi, Attiya, & Ariya, 2015). Il est intéressant de noter que la composition taxonomique des fleurs de givre est distincte de celle de la glace et de l'océan sous-jacents. Parallèlement à la neige saline, les communautés de fleurs de givre semblent être moins diversifiées et présentent un enrichissement spécifique de certains taxons. En outre, le séquençage métagénomique d'un échantillon singulier a montré que ces organismes ont le potentiel de contribuer à divers cycles biogéochimiques, notamment le cycle du mercure et la production de sulfure de diméthyle (DMS) (Jeff S. Bowman et al., 2014). Bien que la source des organismes des fleurs de givre semble provenir de la saumure, il n'est pas clair si la colonisation par l'atmosphère y contribue. De même, l'influence des fleurs de givre sur les communautés microbiennes des environnements connectés n'est pas connue, bien que l'impact de leur composition chimique sur la chimie atmosphérique ait été documenté (F. Domine et al., 2004 ; Huang & Jaeglé, 2017). Les fleurs de givre peuvent être recouvertes par des chutes de neige, fondre et s'infiltrer dans la glace de mer ou être emportées par des vents forts. À leur tour, elles peuvent servir de sources d'ensemencement pour les communautés locales de glace de mer et de neige et agir comme des vecteurs de dispersion plus ou moins longs pour les micro-organismes dans l'atmosphère.

1.4 L'environnement physio-chimique, les interactions biologiques et la succession microbienne à différentes échelles temporelles

Les micro-organismes sont confrontés à une sélection supplémentaire par de forts gradients environnementaux à travers diverses échelles spatiales et temporelles une fois transportés vers la glace de mer pendant sa formation (Bowman, 2015 ; Marcela Ewert & Deming, 2013 ; Vonnahme et al., 2020). L'un des changements les plus drastiques de l'environnement physico-chimique de la glace de mer est dû à la forte saisonnalité des hémisphères nord et sud. À plus long terme, si la glace de mer persiste pendant plus d'une année, les contraintes environnementales imposées aux micro-organismes sont à nouveau susceptibles de changer. À leur tour, les changements de l'environnement physico-chimique auront un impact sur la manière dont les populations microbiennes interagissent, ce qui régira la dynamique des communautés microbiennes et, par conséquent, les cycles biogéochimiques associés.

1.4.1 Conditions sélectives régissant la succession microbienne (automne et hiver)

En automne, des cristaux de glace frasil se forment et se consolident en jeune glace de mer (Petrich et Eicken 2010), ce qui entraîne la formation de canaux de saumure riches en nutriments et en matières organiques. Une phase de latence, caractérisée par une diminution globale de l'activité algale et bactérienne, est associée à cette transition ; peut-être en conséquence d'une baisse de température, d'une augmentation de la salinité et/ou d'une carence en systèmes efficaces d'absorption des nutriments (Grossmann et Dieckmann 1994). En hiver, l'absence de lumière, les salinités élevées et les températures froides exacerbent encore les défis auxquels sont confrontés les micro-organismes de la glace de mer. La production bactérienne globale est faible par rapport aux autres saisons, mais supérieure à celle des microalgues, ce qui entraîne une hétérotrophie nette (Grossmann et Dieckmann 1994 ; Wells et Deming 2006a, b). Après le gel initial, la transition de l'automne à l'hiver se caractérise par une diminution de la lumière et de la température. La basse température réduit le volume de la saumure, diminuant l'espace habitable et forçant les interactions (Deming & Eric Collins, 2016a). On sait peu de choses sur les interactions entre les microorganismes de la glace de mer en automne et en hiver. Néanmoins, il existe un potentiel de compétition et de collaboration. Par exemple, l'espace habitable restreint peut accroître la concurrence pour certaines ressources si deux organismes occupent la même niche métabolique, ce qui conduirait à terme à une exclusion compétitive (Zhou & Ning, 2017). Cependant, l'une des principales sources de matière organique dans les saumures de glace de mer est constituée par les polysaccharides des algues qui sont généralement trop complexes pour qu'une seule espèce bactérienne puisse les décomposer. Par conséquent, les micro-organismes présentes dans les saumures de glace de mer peuvent former des collaborations métaboliques pour utiliser des polysaccharides d'algues complexes, dans un processus appelé partage des ressources (Kabisch et al., 2014). Ce phénomène a été étudié dans les efflorescences algales (Williams et al., 2013), mais n'a pas encore été étudié dans la glace de mer. Si la teneur en nutriments est relativement élevée dans les saumures de la glace de mer, ce qui peut entraîner une compétition dans les environnements froids, comme récemment observé dans les manteaux de neige terrestres (Bergk Pinto, Maccario, Dommergue, Vogel, & Larose, 2019a), la complexité de la disponibilité des substrats pourrait favoriser une communauté plus collaborative. Les micro-organismes peuvent également modifier le métabolisme en fonction de leurs besoins, en formant des biofilms pendant l'hiver (Thomas & Dieckmann, 2002). Cet arrangement mutualiste pourrait constituer une stratégie de survie dominante dans les environnements extrêmes (Ramanan, Kim, Cho, Oh et Kim, 2016).

Le manque d'espace habitable peut également augmenter les interactions bactériennes avec les eucaryotes et les virus. Alors que la prédation a un impact minime sur les communautés de glace de mer (Gowing et al., 2004), les virus sont jusqu'à 100 fois enrichis par rapport à l'eau de mer sous-jacente - tant au printemps (Maranger, Bird, & Kim Juniper, 1994) qu'à l'automne (Collins & Deming, 2011b). Des ratios entre les virus et les bactéries (VBR) de 10-72 ont été observés au printemps dans l'Arctique (Maranger et al., 1994) et de 0,7-119 dans la banquise d'été de l'Antarctique (Gowing et al., 2004). La seule étude sur la glace d'automne a montré des VBR extrêmement élevés (moyenne 846, maximum), ce qui suggèrent un entraînement préférentiel ou une production active de virus au sein de la glace (Collins & Deming, 2011c). En revanche, des VBR de 3 à 18 ont été observés dans la glace d'hiver de l'Arctique (Wells & Deming, 2006). Les cycles lysogéniques en fonction des faibles taux de production virale pourraient dominer pendant les mois froids ou les organismes des glaces de mer hivernales pourraient posséder des mécanismes de résistance aux phages qui favorisent leur stabilité dans la glace de mer. Yu et al. ont postulé que les phages augmentaient la capacité de l'hôte à survivre pendant les hivers polaires lorsque les teneurs en nutriments sont limitées par le ralentissant du métabolisme de l'hôte et en augmentant sa capacité à trouver de nouvelles sources de nutriments au sein de la structure hétérogène de la glace de mer (Yu et al., 2015). Bien que les efforts de culture aient permis de caractériser les virus de la glace de mer, d'identifier leur gamme d'hôtes, ainsi que leur dépendance à l'égard des températures froides pour l'infection (Borriss, Helmke, Hanschke et Schweder, 2003 ; Luhtanen et al., 2014 ; Senčilo, Luhtanen, Saarijärvi, Bamford et Roine, 2015), nous connaissons peu de choses sur la diversité des systèmes de défense de l'hôte. La faible fréquence des espaceurs CRISPR dans les isolats de glace de mer adaptés au froid a conduit à l'hypothèse selon laquelle les EPS pourraient être un mécanisme de défense de l'hôte primaire (Feng, Powell, Wilson, & Bowman, 2013). Cependant, il reste à savoir si cela se traduit au sein des communautés naturelles de glace de mer.

Les adaptations innées qui permettent aux organismes de persister et même de rester actifs pendant les mois contraignants de l'automne et de l'hiver, comme les protéines actives au froid, la stabilité des membranes et les cryoprotecteurs, ont été détaillées ailleurs (Rodrigues & Tiedje, 2008). Toutefois, il est clair que les pionniers de la glace de mer nouvellement formée sont sélectionnés en fonction de leur capacité à survivre et à se développer. Les interactions bactériennes, à la fois inter et intra-domaine, peuvent être cruciales pour façonner la dynamique des communautés microbiennes.

1.4.2 Conditions sélectives régissant la succession microbienne (printemps et été)

Les conditions physico-chimiques dominantes changent suite à l'augmentation saisonnière du rayonnement solaire et de la température. Ces conditions favorisent la prolifération printanière des algues de la glace de mer et, de ce fait, une augmentation marquée de la production primaire par rapport à la dominance hétérotrophe de l'automne et de l'hiver (Vonnahme et al., 2020). Les efflorescences algales offrent un grand espace pour les interactions entre domaines qui pourraient être un moteur important des communautés bactériennes de la glace de mer, en particulier dans la partie inférieure de la glace où les efflorescences dominent. Alors que la production sous la glace augmente via la production primaire d'algues, les couches supérieures de la glace connaissent une activité hétérotrophe encore plus importante attribuée à la partie bactérienne de la communauté (Rolf Gradinger & Zhang, 1997). L'augmentation des températures facilite l'utilisation de la vieille matière organique ou EPS, désormais considérée comme inessentielle dans son rôle protecteur (Arrigo et al., 2014 ; Deming & Eric Collins, 2016a ; Vonnahme et al., 2020)

Si les relations entre les algues et les bactéries sont généralement appréciées dans l'ensemble de l'écologie microbienne, leur couplage dans la glace de mer est incertain. Pendant les mois d'hiver et d'automne, lorsque la productivité primaire est faible, il peut y avoir un découplage entre les bactéries et les algues (Stewart & Fritsen, 2004). Toutefois, la force de la relation peut être plus importante au printemps et en été, lorsque la production primaire est élevée et qu'elle est rapidement acheminée dans la boucle microbienne (Vonnahme et al., 2020). Si la structure et la succession des communautés microbiennes ont été intimement liées à des stades spécifiques d’efflorescences algales dans les habitats pélagiques froids, reflétant des interactions complexes entre domaines, cela reste à explorer dans un environnement de glace de mer (Fadeev et al., 2018 ; Rapp, Fernández-Méndez, Bienhold, & Boetius, 2018). Toutefois, les interactions spécifiques induites par la production de cobalamine dans les bactéries sont déterminantes pour la croissance du phytoplancton dans les communautés côtières de la bordure de la glace de mer de l'Antarctique, ce qui souligne l'importance de la coopération entre les domaines dans la productivité primaire (Bertrand et al., 2015).

À la surface de la glace, où l'hétérotrophie règne, la différence marquée dans l'environnement physico-chimique va modifier la dynamique des interactions bactériennes. Par exemple, la matière organique qui s'épuise rapidement et qui était soumise à une lente dégradation pendant les mois les plus froids pourrait favoriser une compétition accrue entre des individus ayant des besoins nutritionnels similaires. Cela se reflète dans les données intersaisonnières limitées disponibles, où les taxons réputés pour leur capacité à utiliser rapidement des substrats augmentent en abondance (Kaartokallio et al., 2008). De nouvelles niches écologiques apparaissent à la suite de l'augmentation de l'activité, ce qui offre de nouvelles possibilités de relations inédites. Par exemple, l'augmentation des taux de matière organique en ammonium et en nitrate épuisera les niveaux d'oxygène, ce qui entraînera des poches d'anoxie où la dénitrification et l'annamox devraient jouer un rôle (Rysgaard & Glud, 2004 ; Rysgaard, Glud, Sejr, Blicher, & Stahl, 2008).

1.4.3 Glace pluriannuelle

Des échelles temporelles plus longues entraînent des différences plus marquées dans les conditions physico-chimiques de la glace de mer. Cela se manifeste par deux types de glace de mer distincts : la glace pluriannuelle (MYI), définie comme la glace de mer qui persiste plus d'une saison de fonte, et la glace de première année (FYI) qui subit une fonte complète au cours du printemps et de l'été. Ces deux types de glace présentent des caractéristiques d'habitat différentes (par exemple, l'épaisseur, la salinité de la glace, l'âge et l'albédo) qui peuvent avoir un impact sur les processus biologiques (Weeks et Ackley 1986). Les deux types de glace sont soumis à des fluctuations saisonnières, mais les conditions météorologiques peuvent avoir des effets contrastés sur les types de glace. Par exemple, lorsque le FYI subit une fonte complète pendant la saison chaude, des bassins d'eau de fonte (de glace et de neige) se forment à la surface du MYI et peuvent percoler à travers la matrice de glace (Deming & Eric Collins, 2016b). Le MYI offre un environnement plus stable que le FYI, qui peut être caractérisé par des gradients plus prononcés de température, de lumière, de volume de saumure et de salinité (Hatam, Lange, Beckers, Haas et Lanoil, 2016). La stabilité des conditions dans le MYI se reflète dans la stabilité de sa communauté bactérienne (Hatam et al., 2016). La stratification de la glace pluriannuelle composée d'au moins deux couches, une ancienne couche qui survit à la saison de fonte et une nouvelle couche inférieure formée pendant la période de gel saisonnier, ajoute à la complexité de la question. De nettes différences dans les assemblages bactériens qui colonisent ces couches ont été mises en évidence et sont principalement attribuées à l'origine de la glace, les adaptations aux conditions locales jouant un rôle relativement mineur (Hatam et al., 2014). Cependant, le potentiel fonctionnel des communautés MYI ou la manière dont le changement de la composition de la communauté et la stabilité relative des conditions affectent les interactions microbiennes n'a pas été étudié.

1.4.4 Interaction par transfert horizontal de gènes

Les organismes vivant dans les canaux salins de la glace de mer connaîtront des taux de contact cellule-cellule plus élevés en raison de l'espace limité et des densités cellulaires élevées par rapport à la vie en milieu pélagique. La densité cellulaire a été mise en évidence comme un déterminant majeur du transfert horizontal de gènes par conjugaison (HGT) dans d'autres environnements (Elsas & Bailey, 2002) et, à ce titre, pourrait offrir une forme importante d'interaction et d'adaptation pour les micro-organismes vivant dans des saumures à espace limité. Bien que les taux de conjugaison n'aient pas été déterminés dans la glace de mer, des preuves circonstancielles par l'identification de plasmides dans des micro-organismes archétypaux de la glace de mer suggèrent que les HGT pourraient avoir lieu. Par exemple, il a été démontré que 30 % des bactéries isolées de la glace de mer de l'Antarctique contiennent des plasmides (Kobori, Sullivan et Shizuya, 1984) et certains isolats bactériens de la saumure de l'Arctique portent des plasmides conférant une résistance au mercure (Møller et al., 2014). En outre, un grand pool d'ADN extracellulaire qui pourrait servir de médiateur à la transformation a été mesuré dans la glace de mer, mais les auteurs ont suggéré que ce processus pourrait être d'une importance limitée car la diffusivité modélisée de l'ADN libre est au moins d'un ordre de grandeur inférieur à celle de l'ADN encapsulé dans des virus (Collins & Deming, 2011a). Néanmoins, une étude en microcosme a montré que l'efficacité de la transformation était 50 fois plus élevée à l'interface solide-liquide que dans le liquide seul (Lorenz, Aardema, & Wackernagel, 1988), ce qui renforce l'hypothèse selon laquelle la glace de mer serait un « hotspot » pour les HGT. De plus, étant donné les concentrations élevées de virus à bactéries dans la glace de mer, le contact avec les virus est susceptible de se produire plus fréquemment. La transduction pourrait donc être une voie de HGT plus dominante dans les communautés actives de la glace de mer. Enfin, des preuves d'événements HGT dans les micro-organismes de la glace de mer ont été enregistrées. Par exemple, il est probable que la HGT soit responsable de l'acquisition inter-domaines de protéines de liaison à la glace (Janech, Krell, Mock, Kang, & Raymond, 2006 ; Raymond, Fritsen, & Shen, 2007) et de protéines antigel (Bayer-Giraldi, Uhlig, John, Mock, & Valentin, 2010 ; Kiko, 2010) dans tous les domaines de la vie.

1.4.5 L'environnement physico-chimique et la dynamique de la succession

Les pressions sélectives imposées par l'environnement physico-chimique peuvent contrôler la dynamique des communautés microbiennes. Les contraintes environnementales imposées aux microorganismes de la glace de mer sont extrêmement variables à la fois à l'échelle spatiale et temporelle et pourraient à leur tour régir la façon dont les microorganismes se répartissent. Comprendre comment les populations et les communautés réagissent aux conditions changeantes nous aidera à prédire les changements de communauté en conséquence de l'évolution des types de glace dominants dans l'Arctique. Toutefois, les études menées pendant la période de gel automnal et tout au long de l'hiver sont rares. De plus, bien qu'une poignée d'études aient identifié des variations taxonomiques dans la structure des communautés au cours des différentes saisons, la diversité fonctionnelle des bactéries et des algues reste largement inexplorée.

1.5 Le séquençage de nouvelle génération comme outil dans l'écologie de la succession des glaces de mer

Pour dépasser les limites présentées par les méthodes traditionnelles telles que la culture et la microscopie, des technologies de séquençage ont été mises en place, offrant une plate-forme permettant de déduire les relations phylogénétiques et la capacité fonctionnelle. Les technologies NGS diffèrent des anciennes méthodes de séquençage Sanger en ce sens qu'elles permettent une analyse massivement parallèle, une production à haut débit à des coûts relativement faibles. Les technologies NGS peuvent être appliquées pour générer des génomes par séquençage du génome entier (WGS) de micro-organismes en culture ou même pour étudier le contenu génomique total des communautés microbiennes qui font le pont entre les trois domaines de la vie sans qu'il soit nécessaire de procéder à des cultures.

1.5.1 Séquençage du génome entier (WGS)

L'avènement des NGS a permis la construction plus facile, plus rapide et moins coûteuse de génomes complets. Les WGS provenant de bactéries isolées fournissent des informations importantes sur le fonctionnement des écosystèmes. Par exemple, le WGS permet d'évaluer l'impact de l'adaptation et de l'évolution sur les changements génomiques ou de comprendre pourquoi des organismes particuliers prospèrent dans certaines conditions environnementales. L'analyse de génomes entiers fournit une résolution qui permet de discriminer des populations étroitement liées (Delmont & Eren, 2018). À son tour, la génomique comparative des populations peut mettre en évidence le degré de représentativité des individus par rapport à leurs populations associées et l'importance de processus tels que le transfert horizontal de gènes (HGT) et le réarrangement génomique. Par exemple, le séquençage du génome du micro-organisme archétypal de la glace de mer *Colwellia psychrerythraea 34H* a permis de mieux comprendre les stratégies d'adaptation, comme la synthèse de composés susceptibles de conférer une cryotolérance (par exemple les PHA, les composés de type cyanophycine, la glycine bétaïne) et la capacité de produire de grandes quantités d'enzymes extracellulaires (Methé et al., 2005). Dans une étude séparée, la génomique comparative de trois souches de *C. psychrerythraea 34H* a démontré une grande quantité de synthèse du gène de l'ARNr 16S, mais des profils d'utilisation des sources de carbone distincts reflétant leur site d'isolement (Techtmann et al., 2016). Une approche génomique comparative a également été mise en œuvre pour résoudre les différences entre deux espèces étroitement apparentées d'*Octadecabacter*, isolées de la glace de mer à des pôles opposés. Bien que leur gène ARNr 16S soit similaire à plus de 99 %, moins de 80 % de leur contenu génétique était comparable (Vollmers et al., 2013). En outre, la diversité et l'abondance des éléments transposables dans les deux génomes indiquent l'importance de la plasticité du génome et des éléments génétiques mobiles sur leur évolution.

Toutes les études mentionnées précédemment n'ont pas été réalisées à l'aide des technologies NGS ; cependant, leur application au WGS permettra de résoudre un plus grand nombre de génomes sous-représentés dans les bases de données actuelles. En outre, la capacité de générer rapidement et à moindre coût des génomes à partir de souches étroitement apparentées fournira une plus grande plate-forme pour les études de comparaison génomique, ce qui permettra d'étudier les pressions sélectives qui régissent des processus tels que la biogéographie microbienne au niveau génomique.

1.5.2 Séquençage des gènes marqueurs (MGS)

Le MGS désigne l'amplification de fragments de séquences d'ADN isolés directement de l'environnement sans qu'il soit nécessaire de cultiver des microorganismes. En général, les gènes ARNr 16S et ARNr 18S/espaceur transcrit interne ribosomique (ITS) sont ciblés pour déduire la phylogénie des communautés bactériennes et eucaryotes, respectivement. L'avènement des NGS en 2005 et la diminution correspondante des coûts de séquençage ont entraîné une augmentation spectaculaire des études de NGS en écologie microbienne. Toutefois, son application à la glace de mer a été limitée et la plupart des études de la glace de mer sur les EMN ont mis en œuvre des technologies pré-ENG telles que ARISA et RFLP. Ces études ont permis une première caractérisation du biote microbien de la glace de mer et ont mis en évidence la variabilité des communautés en fonction de l'espace (Brinkmeyer et al., 2003 ; Yergeau et al., 2017), du temps (Eric Collins, Rocap, & Deming, 2010) et du type de glace

(Hatam, Lange, Beckers, Haas, & Lanoil, 2016). De plus, la sélection répétitive de phylotypes d'ARNr 16S marins rares dans la glace de mer, indépendamment de la géographie, suggère fortement un mécanisme de sélection basé sur une niche. Les raisons qui sous-tendent l'enrichissement spécifique des taxons ne sont pas encore pleinement réalisées.

Le développement des technologies de séquençage s'accompagne d'un progrès rapide des outils bioinformatiques. Les interactions entre les microbes et l'environnement peuvent être déduites de moyens relativement simples tels que des analyses de corrélation entre les variables environnementales mesurées et l'abondance des taxons. D'autre part, les interactions microbe-microbes sont importantes sur le plan écologique, mais généralement moins étudiées que les relations microbe-environnement. Les analyses de réseaux basées sur les corrélations ont gagné en popularité en raison de l'augmentation des ensembles de données de séquençage et de la disponibilité croissante d'outils bioinformatiques spécifiques aux données d'amplicon. L'identification de relations présumées par le biais d'analyses de réseau a permis de démêler des interactions microbiennes complexes et de reconnaître des associations métaboliques potentielles, des espèces clés ou des groupes de taxons cooccurrents à des échelles temporelles et spatiales (Bergk Pinto, Maccario, Dommergue, Vogel et Larose, 2019b ; Chafee et al. 2018 ; Milici et al. 2016 ; Peura, Bertilsson, Jones et Eiler, 2015). Si ces approches ont été appliquées aux communautés microbiennes de divers environnements, ces méthodes doivent encore être appliquées aux données sur la glace de mer. En outre, les interactions entre domaines ont une importance fondamentale dans les écosystèmes marins (Fadeev et al., 2018) et leurs interactions affectent le cycle biogéochimique et, en fin de compte, le climat (Ramanan et al., 2016). L'importance de ces interactions peut être soulignée dans la glace de mer où la réduction de l'espace habitable entraîne des interactions plus importantes. Cependant, les bactéries et les eucaryotes sont rarement étudiés simultanément dans la glace de mer.

1.5.3 Métagénomique

Le séquençage métagénomique fait référence au séquençage non ciblé de tous les génomes microbiens présents dans un échantillon et, à ce titre, peut être utilisé pour étudier simultanément la composition taxonomique et le potentiel fonctionnel des communautés microbiennes. Cet outil est couramment utilisé dans l'étude de divers habitats, mais son application aux environnements de glace de mer reste limitée. Néanmoins, la première étude métagénomique était basée sur un seul métagénome de FYI et un autre provenant d'un échantillon de fleurs de givre (Bowman et al., 2014). Bien qu'il soit difficile d'établir le rôle des communautés microbiennes dans la glace de mer et les fleurs de givre à partir d'un seul échantillon, l'étude a démontré certaines fonctions d'intérêt potentiel. Par exemple, la communauté microbienne était capable de catabolisme de DMSP et de réduction du mercure, ce qui pourrait avoir des conséquences importantes sur le cycle biogéochimique polaire. Une étude biogéographique a mis en évidence une nette ségrégation de la glace de mer et de l'eau de mer à la fois en termes de taxonomie et de fonction (Yergeau et al., 2017). Toutefois, l'étude a été limitée à la partie inférieure de la glace de mer, mais la stratification verticale, du moins sur le plan taxonomique, est un phénomène bien décrit (Hatam et al., 2014). Enfin, la dernière étude métagénomique de la glace de mer a été la première à évaluer un profil complet de la glace de mer, y compris la neige et l'atmosphère qui la recouvrent (Maccario et al., 2019). Il a été démontré que les communautés microbiennes sont stratifiées verticalement en termes de caractéristiques taxonomiques et fonctionnelles, la neige étant plus proche des communautés atmosphériques que des communautés de glace de mer, de saumure ou d'eau de mer. Les communautés de neige ont montré une grande polyvalence dans l'utilisation des sources de carbone, avec des gènes liés à la dégradation de l'amidon, du saccharose et de certains composés xénobiotiques. Cette étude a également identifié la couche de neige saline comme une niche unique, potentiellement alimentée par des nutriments résultant de l'expulsion de la saumure, qui soutient probablement la croissance de membres copiotrophes, psycho- et halotolérants de la neige tels que Glaciecola (Maccario et al., 2019). Toutefois, cette étude a été réalisée sur un seul site au Groenland et des recherches supplémentaires sont nécessaires afin de valider cette hypothèse.

Les génomes de référence fournissent des informations sur les propriétés métaboliques, la phylogénie, la fonctionnalité, le cycle biogéochimique et les interactions, mais il n'en existe que très peu pour les micro-organismes de la glace de mer (Deming & Eric Collins, 2016a). Bien que les bactéries de la glace de mer soient généralement considérées comme pouvant être cultivées, l'absence de séquences génomiques complètes et le biais en faveur de certains taxons limitent la portée de l'étude. En outre, la culture de micro-organismes peut forcer des changements dans le contenu génomique et, par conséquent, limiter son applicabilité aux écosystèmes naturels. Les progrès des outils bioinformatiques ont permis d'assembler des ébauches de génomes à partir d'ensembles de données métagénomiques. De multiples génomes assemblés par métagénomique (MAG) peuvent être récupérés à partir d'ensembles de données et ils sont souvent très différents de toute espèce isolée, ce qui offre un net avantage par rapport aux approches traditionnelles (Barnum et al., 2018 ; Delmont, Murat Eren, Vineis, & Post, 2015 ; Wilkins, Ettinger, Jospin, & Eisen, 2019). Par conséquent, cette ligne d'étude constitue une opportunité de résoudre les génomes des représentants de la glace de mer qui manquent actuellement dans les bases de données telles que *Flavobacterium*, *Polaribacter*, *Pelagibacter*, *Loktanella* et *Sulfitobacter* (Deming & Eric Collins, 2016a) et pourrait fournir un nouvel aperçu des processus biogéochimiques qui ont lieu dans la glace de mer. Par exemple, l'oxydation complète de l'ammoniac en nitrate via le nitrite (c'est-à-dire le comammox) a été déterminée dans un seul MAG (van Kessel et al. 2015). Ces outils pourraient fournir un aperçu des interactions entre domaines, comme la répartition des ressources et des niches, grâce à l'étude d'îlots génomiques spécialisés (Kabisch et al., 2014), ou découvrir des réponses adaptatives/évolutives aux changements environnementaux par le biais du HGT (Feng et al., 2013).

1.6 Conclusion

La dynamique de la succession des microbiomes de la glace de mer et de la neige n'est pas bien comprise, mais elle est cruciale pour prédire le développement des écosystèmes. Au cours de la succession primaire, les taxons pionniers sont des copiotrophes potentiellement mobiles qui recherchent activement l'environnement salin riche en nutriments, des épibiontes qui se déplacent avec de plus grandes cellules d'algues et/ou des organismes producteurs de EPS et de protéines de glace qui améliorent leur taux d'entraînement en se collant ou en se liant aux cristaux de glace. Les taxons pionniers qui produisent de l'EPS façonnent physiquement l'espace de niche disponible tout en contribuant à la réserve de matière organique disponible pour l'activité hétérotrophe. Il doit également s'agir d'organismes capables de résister aux fortes variations des paramètres physico-chimiques tant lors de l'induction de la glace de mer que lors du changement de saison. D'autres horizons d'importance émergente (fleurs de givre et de neige saline) peuvent agir pour modifier la dynamique biologique, chimique et physique du profil de la glace de mer et d'autres écosystèmes par le biais de la dispersion. Par conséquent, l'écologie de la succession des glaces de mer pourrait bénéficier de trois grandes orientations de recherche.

Premièrement, nous devons démêler les facteurs qui favorisent la colonisation primaire et commencer à dissiper les hypothèses avec lesquelles les preuves du contraire s'accumulent. Par exemple, l'incorporation bactérienne induite par l'attachement des algues n'a que peu de preuves directes, les recherches ne permettant pas d'établir une corrélation. Nous devrions plutôt nous concentrer sur des hypothèses prometteuses telles que le lien entre la motilité ou la capacité à produire du EPS et l'incorporation dans la glace de mer.

Ensuite, la dynamique de succession des microbiomes de la glace de mer est liée aux changements de l'environnement physiochimique, comme l'intensité des UV, la salinité et la disponibilité du substrat. Les contributions relatives de ces déterminants représentent des lacunes fondamentales dans la compréhension des trajectoires de développement des écosystèmes et seront cruciales pour les prévisions du changement climatique.

Enfin, les environnements émergents tels que la neige saline et les fleurs de givre sont riches en biomasse et leur étendue spatiale et temporelle augmente. Les quelques études disponibles suggèrent un enrichissement spécifique de certains taxons dans ces deux environnements d'interface. Cependant, l'impact de cette chute spectaculaire de la diversité est inconnu. Les études futures devraient chercher à comprendre les facteurs écologiques sous-jacents de cet enrichissement afin de mieux prévoir les changements au niveau des écosystèmes qui sont accélérés par le réchauffement climatique.

Chapter 1: Bibliography – Microbial succession ecology of sea ice

1.1 Introduction

Microorganisms contain members from all the domains of life (e.g. bacteria, fungi, viruses, protozoans, algae). These microbes form communities with complex interactions that drive many globally important processes such as biogeochemistry (Bertrand et al., 2015), pathogenesis (Chow, Tang, & Mazmanian, 2011; Rogers, Hoffman, Carroll, & Bruce, 2013; Vayssier-Taussat et al., 2014), symbiosis (Beinart, 2019; Kamagata & Narihiro, 2016; Raina, Fernandez, Lambert, Stocker, & Seymour, 2019; Wein et al., 2019) and bioremediation (Paliwal, Puranik, & Purohit, 2012; Pandey, Chauhan, & Jain, 2009). Their contribution to such a wide range of processes is partly owed to their ability to colonise any habitat, even the most inhospitable (Bull & Goodfellow, 2019; Horikoshi & Bull, 2011). Although colonisation by microorganisms is a continuous process, primary colonisation will shape how microbial communities establish within new environments and govern the distribution, composition and latter development of the community. Clearly, how microorganisms sort themselves in the initial stages of community formation is fundamental to microbial ecology. Whilst colonisation processes have been studied extensively in culture (Ben-Jacob, Cohen, & Levine, 2000; Burrows, Elbert, Lawrence, & Pöschl, 2009; Dann, Clanahan, Paterson, & Mitchell, 2019; Grossart, Kiørboe, Tang, & Ploug, 2003; Hallatschek & Nelson, 2010; Kerr, Neuhauser, Bohannan, & Dean, 2006; Kerr, Riley, Feldman, & Bohannan, 2002; Lloyd & Allen, 2015; Venegas-Ortiz, Allen, & Evans, 2014), our understanding of how they operate in natural communities is poorly understood. Sea-ice is an excellent model for elucidating colonization processes and microbial community adaptation, given its rapid formation, seasonality and multiple habitable niches with steep physico-chemical gradients.

Stages of colonisation can be broken down into a number of broad categories. Initially, microorganisms must be passively or actively transferred to the new niche environment. Therefore, factors such as the diversity and structure of the source community is thought to influence the diversity and structure of the sink community (Lindström & Langenheder, 2012; Zhou & Ning, 2017). Innate adaptations in specific populations may also contribute to their ability to pioneer new habitats (e.g. chemotaxis) (Josenhans & Schweinitzer, 2010; Raina et al., 2019). Once present, microbial pioneers will be subject to a new sphere of physiochemical parameters. Successful pioneers must overcome potentially novel selection pressures such as different temperature, salinity, light and nutrient gradients (Lindström & Langenheder, 2012). This leads to environmental exclusion of certain species which are unable to survive once present (e.g. lysed due to differences in osmotic pressures) (Zhou & Ning, 2017). Additionally, novel intra- species interactions (competition, co-operation and mutualism) can drive early establishment of community structure and diversity. Long term-competition is thought to result in either (1) *competitive exclusion*, where one population dominates and the other faces extinction, (2) *metabolic niche partitioning*, where competitors are able to coexist because they specialize on different resource types or (3) *physical niche partitioning*, where competitors disperse into different territorial niches (e.g. in semi-solid environments) (Bauer, Kainz, Carmona-Gutierrez, & Madeo, 2018).

Habitats are often not spatially and temporally stable and thus disturbance can lead to secondary succession in microbial communities (Fierer, Nemergut, Knight, & Craine, 2010). For instance, seasonal changes in temperature may force different microbial dynamics in winter and summer sea ice (Eric Collins, Rocap, & Deming, 2010; Luria, Amaral-Zettler, Ducklow, & Rich, 2016; Schmidt et al., 2007; Wu et al., 2016). Connected environments which act as microbial seed banks may also face change and potentially deliver novel, competing species. Disentangling the factors that drive the different stages of microbial succession in natural environments is difficult due to the unpredictability of changing environmental parameters, the various spatio-temporal scales at which these parameters act and the complexity of microbial interactions with each other and their physiochemical environment (Zhou & Ning, 2017).

In line with climate change, there is a transition from a dominance of multi-year ice (MYI) to new and young sea ice (FYI) in the Polar Regions (Kwok, 2018). Not only is FYI increasing in its spatial and temporal importance, but it represents a unique model to study microbial succession ecology. As temperatures drop in autumn, formation of FYI is coupled with primary microbial colonisation (Vonnahme, Dietrich, & Hassett, 2020). Due to the exclusion of salt during ice formation, a semi-solid environment is formed consisting of a solid phase ice-matrix permeated by a network of hypersaline, nutrient and organic matter (OM) dense liquid brine channels (Hunke, Notz, Turner, & Vancoppenolle, 2011). This spatial separation of environmental structures and the diversity of metabolic substrate availability presents opportunities for both physical and metabolic niche partitioning. Moreover, the prevailing environmental conditions of newly formed sea ice creates a strong selective environment for microbial pioneers allowing exploration of environmental filtering processes (Eronen-Rasimus et al., 2015). Additionally, the reduced habitable space of the brine forces increased interaction and thus facilitates the study of competitive exclusion versus metabolic and niche partitioning. Because Polar Regions have the most distinct seasonality across the globe, sea ice microbial communities are subject to strong seasonal disturbances (Vonnahme et al., 2020). We can use this natural occurrence to understand longer-term succession patterns. Precipitation events and ocean currents replenish microbial inputs into the sea ice system and as such offer insight into whether the diversity and structure of source communities contribute to the microbial community dynamics of the sink environment. Precipitation events such as snowfall add new layers to the sea ice profile which provide another colonisable and contrasting habitat for study.

Whilst sea ice microbiology has been extensively reviewed (Boetius, Anesio, Deming, Mikucki, & Rapp, 2015; J. P. Bowman, 2013; Jeff S. Bowman, 2015; Deming & Eric Collins, 2016a; Marcela Ewert & Deming, 2013; Vonnahme et al., 2020), we focus this chapter on how the available literature aids our understanding of the colonisation processes in natural ecosystems. We establish how stochastic and deterministic processes contribute to primary colonisation of sea ice and how the factors that govern microbial succession might change over temporal and spatial scales. Our objectives are to understand how inter and intra domain relationships, microbial interactions with their physical-chemical environment and the diversity and structure of the source environment affects the establishment and development of sea ice and its associated habitats. Finally, we touch upon how understanding the dynamics of microbial succession drivers in sea ice microbial communities is important for making future predictions on the effects of sea ice melt and how rapidly developing sequencing and bioinformatics tools are facilitating this direction of study.

1.2 Primary colonisation of sea ice

First year sea ice forms in line with the autumnal decrease in temperature. Initially, small frazil ice crystals with ineffective buoyancy form in super-cooled (i.e., water temperatures below 0 degrees Celsius) water. Their circulation through the upper water column due to ocean currents promotes their adherence to particles and cells in the water. These crystals then rise to the surface as they grow, resulting in a slushy layer known as grease ice (Omstedt, 1985). The grease ice consolidates as temperature drops, forming young sea ice which is composed of an ice matrix interlaced with a brine network; the latter can contain a high abundance of microorganisms (Deming & Eric Collins, 2016a; S Grossmann & Dieckmann, 1994). An organism's ability to disperse and colonise, referred to here as its successful movement and establishment in young sea ice, is driven both by passive and active pathways, involving factors such as environmental filtering, motility and biotic interactions. It is well known that both bacteria and algae are enriched in sea ice relative to the underlying seawater however, the mechanisms underpinning their enrichment is unclear. In this section we discuss the current understanding of how and why microorganisms are enriched in sea ice.

1.2.1 Frazil Ice, algae and bacterial associations

Garrison *et* al. (1983) first proposed that algae were able to colonise FYI during formation through the scavenging activity of frazil ice crystals based on observations in the Weddell Sea (David L. Garrison, Ackley, & Buck, 1983). This was followed by laboratory experiments demonstrating the ability of frazil ice to harvest and concentrate suspended algal cells (D. L. Garrison, Close, & Reimnitz, 1989). In field studies, ice formation has been shown to preferentially retain larger cells (R. Gradinger & Ikävalko, 1998; Kauko et al., 2018; Riedel, Michel, Gosselin, & LeBlanc, 2007a). Additionally, experimental formation of sea ice suggested negligible bacterial enrichment in the absence of algal cells (Weissenberger & Grossmann, 1998). This has led to the proposition that the enrichment of bacteria observed in sea ice could be the result of their ability to associate with algae (S. Grossmann, 1994; S Grossmann & Dieckmann, 1994; Sönnke Grossmann & Gleitz, 1993; Weissenberger & Grossmann, 1998).

Although bacteria and algae are known to have an intimate relationship in a marine setting, the research addressing this topic in sea ice is limited. Furthermore, the hypothesis that bacterial attachment to larger eukaryotic cells is inconsistent with the observation of low bacterial colonisation of pelagic diatoms (Crenn, Duffieux, & Jeanthon, 2018; Droop & Elson, 1966; Kaczmarska et al., 2005). One study addressing both the bacterial and eukaryotic domains in sea ice, demonstrated that bacterial attachment to diatoms was low and no direct relationship between bacterial enrichment and large autotroph enrichment was observed (Riedel et al., 2007a). Moreover, rare data from winter sea ice of the Antarctic show that algal biomass is a poor indicator of bacterial biomass (Stewart & Fritsen, 2004). Other conflicting evidence demonstrates bacterial enrichment in artificial sea ice, equivalent to that previously observed, despite extremely low abundance of algae (Eronen-Rasimus et al., 2014). Collectively, this research suggests that whilst algal cells may be entrained in ice by a physical size-selective process, bacteria are enriched by different means.

1.2.2 Particle attachment, extracellular polysaccharide (EPS) and ice binding proteins (IBPs)

Particulate matter entrainment in sea ice is consistent with the hypothesis that suspended particles in seawater may act as a delivery system for attached cells (Lindemann, Hölemann, Korablev, & Zachek, 1999). Observations of high numbers of particle-associated cells in sea ice corroborate this hypothesis, although it is unclear whether particle-associations were formed in the pelagic environment, leading to a bias in observed colonisation potential (Karen Junge, Eicken, & Deming, 2004). In turn, the ability of microorganisms to attach to particulate matter may be a key driver of primary sea ice colonisation.

Extracellular polysaccharide (EPS) in sea ice has received a large amount of attention as it has been implicated in a number of important roles. EPS are structurally diverse polymers secreted by a wide range of organisms including many bacteria and algae (Marcela Ewert & Deming, 2013). Along with its role in cryo-preservation, EPS may drive aggregate formation or cell attachment to ice crystals or particles by increasing an organism’s ‘stickiness’. Estimates of exopolymer stickiness were shown as 2 to 4 orders of magnitude higher than the stickiness of other particles (Passow, 2002). This would lead to an increase of EPS-producing organisms in sea ice. High concentrations of EPS in brine relative to the underlying seawater has been consistently reported (Meiners, Gradinger, Fehling, Civitarese, & Spindler, 2003; Riedel, Michel, Gosselin, & LeBlanc, 2007b; Graham J.C. Underwood et al., 2013). Moreover, the attachment of bacteria to exopolymers has been directly observed in pelagic (Passow & Alldredge, 1994) and benthic (G. J. C. Underwood, Paterson, & Parkes, 1995) environments. However, a direct link between entrainment rate and EPS-producing microorganisms has yet to be evidenced.

Ice nucleating proteins (INPs) and antifreeze proteins (AFPs) are ice-binding proteins (IBPs) that can form part of the EPS secreted by cold-adapted marine microorganisms including bacteria, fungi and algae. INPs are able to induce ice formation at higher subzero temperatures (-5 to -12°C) than homogeneous processes that have been estimated to occur at – 42.2 ±0.3°C in artificial seawater (K. Junge & Swanson, 2008). The formation of extracellular ice crystals increases tolerance to freezing by generating heat, which can then be used by the organism. AFPs bind to ice crystals, which can help maintain access to oxygen and nutrients, as well as protect the cell by mediating the growth of smaller crystals, thereby preventing recrystallization (Białkowska, Majewska, Olczak, & Twarda-Clapa, 2020). Therefore, the production of these proteins in sea ice microorganisms can aid continued activity and survival at otherwise constraining temperatures. While the protective role of IBPs in sea ice microorganisms has been reviewed elsewhere *e.g.* (Białkowska et al., 2020; Ewart, Lin, & Hew, 1999), the potential for these proteins to aid primary colonisation of sea ice has not been the focus of many studies. As addressed above, EPS, potentially containing ice-binding protein fractions, is preferentially entrained in sea ice. Collectively, these studies provide circumstantial evidence for the selective entrainment of some micrororganisms based on their ability to produce and secrete proteins that interact with ice crystals. However, it has not yet been shown whether EPS and IBP-producing organisms are entrained into the ice structure more readily than others or if the high concentrations of EPS observed in sea ice are driven by organism response to decreasing temperature and increasing salinity.

1.2.3 Gas vacuoles

Gas vacuoles are comprised of hollow, intracellular proteinaceous structures that agglomerate into large organelles and are found almost exclusively in aquatic bacteria (Anthony E. Walsby, 1981). Inflation or deflation of gas vesicles inside microbial cells provide buoyancy (A E Walsby, 1972) and in turn allows the organism to float at a desired depth in the water column. Thus gas vesicles provide a strategy to select conditions that are favourable for growth, especially in habitats with vertical photic, thermal, or chemical stratification (John J. Gosink, Irgens, & Staley, 1993). Gas vacuoles can act as an alternative to flagellum for movement, and have almost exclusively been found in non-flagellated species. The identification of gas vacuoles in bacteria occupying the upper water column of

Antarctic waters led to the hypothesis that gas vacuoles are coupled to sea ice colonisation or even with movement within the sea ice system (J. J. Gosink & Staley, 1995). This hypothesis was further supported by the higher proportion of gas vacuolated bacteria in ice (165 out of 177 isolates) relative to the underlying open sea water (12 of 177 isolates) identified in an Arctic sea ice study (John J. Gosink et al., 1993). Since the initial discovery, a large number of archetypal sea ice taxa have been shown to produce gas vacuoles including alpha, beta, and gamma *Proteobacteria* and the *Flavobacteria-Cytophaga* group (J. J. Gosink & Staley, 1995). Whether the incorporation of certain bacterial species in sea ice is a direct consequence of gas vacuole possession has not been studied.

1.2.4 Taxis

In order to succeed in heterogeneous environments such as sea ice, organisms must either tune their metabolism to the local conditions or migrate and colonize new niches (Alexandre, 2010). Colonisation of a given niche will depend on the organism's capacity to successfully compete for limiting nutrients and to resist antagonistic mechanisms of other members of the community (Kennedy, 1987). The capacity to sense changes in their immediate environment through a number of responses is referred to as "taxes" and defined as the directional movement responses as a result of external stimuli. It has been proposed as a selective advantage in exploiting new niches, especially in unmixed environments where spatial or temporal chemical gradients persist (Stocker & Seymour, 2012). Many bacteria are capable of detecting changes in their chemical environment and appropriately directing their movement in a process known as chemotaxis. This strategy is thought of as particularly important in seawater microorganisms to facilitate a rapid response to ephemeral nutrient patches, plumes and gradients in a generally dilute ocean (Stocker & Seymour, 2012). Therefore, the formation of sea ice and corresponding nutrient-dense brine will attract chemotactic microorganisms seeking to exploit a favourable niche. It has been suggested that non-motile oligotrophic bacteria dominate when productivity is low (Smriga, Fernandez, Mitchell, & Stocker, 2016). Conversely, motile copiotrophs dominate in nutrient rich conditions. Therefore, we would expect a high proportion of motile copiotrophs in the brine during sea ice formation. Indeed, metagenomic analyses of Greenlandic sea ice showed a higher prevalence of genes related to cell motility and copiotrophy in the brines relative to the underlying seawater (Maccario, Carpenter, Deming, Vogel, & Larose, 2019).

Furthermore, experimental evidence has shown that representative sea ice microorganism *Colwellia psychrerythraea* is capable of motility at temperatures as low as -10°C (Karen Junge, Eicken, & Deming, 2003)and displays chemohalotaxis (Showalter & Deming, 2018). Movement of bacterial cells in near *in situ* conditions were detected using a digital holographic microscope. Motility in response to other environmental gradients could also underscore primary colonisation of sea ice. For instance, in a unique thigmotatic strategy (response to the stimulus of contact), the Antarctic bacterium *Marinomonas primoryensis* uses motility and an ice binding protein to adhere to ice crystals at the ocean surface and remain in a relatively oxygen and nutrient rich environment (Guo et al., 2017). The detection and movement towards light (phototaxis), pH (pH-taxis) and oxygen (aerotaxis) all represent strategies used by ocean-dwelling microorganisms (Stocker & Seymour, 2012) that could affect microbial assembly in sea ice but have not yet been explored.

1.2.5 Multiple strategies of sea ice colonisation

How microorganisms are transferred to newly forming sea ice is likely a combination of deterministic and stochastic processes. For instance, individuals may become attached to rising frazil ice aggregates by chance contact and at the mercy of random ocean turbulence. However, microorganisms may increase their chances of incorporation through deterministic processes such as algal attachment, EPS production, taxis and buoyancy. In turn, this may affect community sorting in sea ice with regards to physical and metabolic niche partitioning. For instance, the possession of ice binding proteins may result in entrapment within the ice matrix, motility could facilitate retention in the liquid brine channels or gas vacuolated bacteria could contribute to the vertical stratification of communities. However, while physical niche partitioning in sea ice is a known phenomenon, it has only been considered in terms of organism abundance (Marcela Ewert & Deming, 2013). It is not only uncertain how different members of the community partition at different functional and taxonomic resolutions but also how transfer mechanisms may be a key structural driver in this process.

1.3 Source and sink environments

Within the context of sea ice formation, seawater acts as the major source of microorganisms whilst sea ice can be termed the sink environment. Specific meteorological conditions or events can result in additional source and sink environments that contribute to the sea ice profile. Consistent with the predicted increased range of FYI is the increased extent of saline snow and frost flowers which accumulate on the surface of sea ice (Barber et al., 2014a; Maccario et al., 2019). These are novel environments that remain underexplored yet increasingly important in terms of their microbial communities and biogeochemical consequences (J. S. Bowman & Deming, 2010; J. S. Bowman, Larose, Vogel, & Deming, 2013a; Maccario et al., 2019). Additionally, they present novel secondary colonisation environments, adding interest to using the sea ice profile as a model for microbial succession.

1.3.1 Seawater

Distinct taxonomic partitioning has been evidenced in laboratory systems as a major factor during the colonisation process, despite taxonomic heterogeneity in the source material (Dann et al., 2019). If this was translatable to natural sea ice systems, newly colonisable sea ice would evidence distinct taxonomic community structure despite variable sources. The initial colonisation stage during sea ice formation has not been well documented however, one study demonstrated that early pancake ice communities were similar to that of open water communities, suggesting that source seawater is a key factor in community structuring at this stage (Eronen-Rasimus et al., 2015). Conversely, once sea ice consolidates and matures, research consistently shows that the communities it harbours contrast to the underlying seawater and is dominated by archetypal sea ice taxa (Boetius et al., 2015; Deming & Eric Collins, 2016a; Eronen-Rasimus et al., 2015; S. Grossmann, 1994). Taxonomic and functional composition of sea ice communities show more similarity to each other than their underlying seawater, independent of large geographic distances as evidenced by a springtime biogeographic study spanning the Canadian Arctic (Yergeau et al., 2017). Collectively, these studies suggest that taxonomy and functional assemblage is driven by selective mechanisms expressly conferred by the sea ice environment; this selection could be attributed to any combination of physiochemical factors (e.g. lysis by osmotic shock) and biological interactions (e.g. competitive exclusion). These results were reflected in mesocosm experiments where the formation of artificial sea ice, using seawater not usually subject to annual sea ice formation, resulted in the appearance of archetypal sea ice bacteria that were undetected in the source material (Eronen-Rasimus et al., 2014). As such, it is suggested that while source seawater may be a major structuring variable during the very initial stages of sea ice formation, the relative proportions of bacterial populations in the source do not confound the emergence of a sea ice specific community.

1.3.2 Snow and Atmosphere

In the Arctic, the majority of the sea ice surface, estimated at 7× 10^6 km^2 in winter (Nghiem et al., 2007), is covered with snow that accumulates between fall and spring (Sturm, Holmgren, & Perovich, 2002; Warren et al., 1999). Not only has snow been recently recognised as an important microbial ecosystem, but milder climates in the Polar Regions could lead to a greater contribution of sea ice mass in the form of metamorphic snow (Granskog, Vihma, Pirazzini, & Cheng, 2006). On a shorter timescale, snowfall changes the physical chemistry of the sea ice, such as through UV extinction, and delivers a novel source of microbial populations to the profile (Maccario et al., 2019). While transfer of microorganisms from the snow to the sea ice has not been studied, in part owing to the minimal biomass it contains relative to sea ice brines, there is potential for snow microorganisms to colonise and grow in the sea ice. Indeed, these microorganisms may be innately adapted to life in an icy environment as many environmental constraints are shared. For instance, sea ice, snow and atmosphere microorganisms are subject to strong environmental gradients in UV, temperature and nutrient availability (Maccario, Sanguino, Vogel, & Larose, 2015). The limited data available on the community structure of the atmosphere and snow overlying sea ice curtails our current ability to decipher their relative contributions to sea ice community succession.

Relative to MYI, FYI contains a larger fraction of brine which can be rejected on to the surface, delivering nutrients, salts and microorganisms to any overlying snow (Barber, Reddan, & LeDrew, 1995). This process results in a saline snow layer which is becoming an increasingly important environment in line with the loss in sea ice thickness and the shift from MYI to FYI (decrease in MYI of 50% since 2002, MYI now covers less than one-third of the Arctic Ocean) (Kwok, 2018). . First year ice now dominates the ice cover, representing nearly 80% of the March 2018 ice pack (D. Perovich *et al.*, 2018). Physical and chemical properties of saline snow have been studied in relatively extensive detail (e.g. Domine et al., 2004; Huang & Jaeglé, 2017; Peterson et al., 2019; Yang et al., 2008) and yet only two papers have considered its microbial content (M. Ewert, Carpenter, Colangelo-Lillis, &

Deming, 2013; Maccario et al., 2019). Nevertheless, saline snow may act as a sink environment in sea ice profiles while brine represents an influx of competitive microorganisms. Initial studies of saline snow show elevated abundances of bacteria and EPS relative to the overlying-low salinity snow. Although early work suggested elevated bacterial counts were attributed to brine organisms, Maccario *et al*. (2019) hypothesised that nutrient supplementation stimulated the growth of snow microorganisms. This is consistent with the observation that suitably adapted microorganisms in the brine were not prevalent in the saline snow. The same study highlighted the lack of diversity present in brine-wetted snow relative to both under- and over- lying systems coupled with taxa-specific enrichment. Whether community assemblage in saline snow is consistent across time and space has not been investigated but would provide useful information regarding an increasingly important ecosystem as the climate changes.

1.3.3 Frost flowers

Frost flowers are delicate structures that form almost exclusively on the surface of young sea ice. If the atmosphere overlying the sea ice is water-saturated and a sufficient temperature gradient exists, ice crystals form on imperfections on the ice surface (Florent Domine, 2005). These serve as nucleation points from which frost flowers grow into dendritic structures that can reach many centimetres in height. Their hollow stems wick up brine skim from the surface resulting in an extremely hypersaline environment (>120 ppt) (Barber et al., 2014b). Similarly, bacteria are transported from the brine skim to frost flowers in surprisingly densities 3– 6-fold higher than in underlying ice (Bowman & Deming, 2010). This discovery fuelled a handful of papers which offered an initial insight into the structural dynamics of frost flower bacterial communities (Barber et al., 2014b; J. S. Bowman et al., 2013a; Jeff S. Bowman, Berthiaume, Armbrust, & Deming, 2014; Mortazavi, Attiya, & Ariya, 2015). Interestingly, the taxonomic composition of frost flowers is distinct from the underlying ice and ocean. In parallel to saline snow, frost flower communities appear to be less diverse and show taxa-specific enrichment. Moreover, metagenomic sequencing of a singular sample showed that these organisms have the potential to contribute to various biogeochemical cycles including mercury cycling and dimethyl sulfide (DMS) production (Jeff S. Bowman et al., 2014). Although the source of frost flower organisms appears to originate from the brine, it is unclear whether colonization via the atmosphere contributes. Equally, the influence frost flowers have on the microbial communities of connected environments is not known, although the impact of their chemical composition on atmospheric chemistry has been documented (F. Domine et al., 2004; Huang & Jaeglé, 2017). Frost flowers may be covered by snowfall events, melt and infiltrate the sea ice or be blown away by strong winds. In turn, they may act as seeding sources for local sea ice and snow communities and act as long or short dispersal vectors for microorganisms in the atmosphere.

1.4 The physio-chemical environment, biological interactions and microbial succession at different time scales

Whether microorganisms are passively or actively transported to the sea ice during formation, they face further selection by strong environmental gradients across various spatial and temporal scales reviewed by (Bowman, 2015; Marcela Ewert & Deming, 2013; Vonnahme et al., 2020). One of the most drastic changes in the physio-chemical environment of sea ice is due to the strong seasonality at the Northern and Southern hemispheres. At longer timescales, if sea ice persists throughout more than one year, the environmental constraints imposed on microorganisms are again subject to change. In turn, the changes in the physical-chemical environment will impact how microbial populations interact, which will govern microbial community dynamics and consequently biogeochemical cycling.

1.4.1 Selective conditions governing microbial succession (autumn and winter)

In autumn, frazil ice crystals form and consolidate to young sea ice (Petrich and Eicken 2010) leading to the formation of nutrient and organic-rich brine channels. A lag phase, characterised by an overall decrease in algal and bacterial activity, is associated with this transition; possibly as a consequence of decreased temperature, increased salinities and/or a deficiency in effective nutrient uptake systems (Grossmann and Dieckmann 1994). In winter, the absence of light, high salinities, and cold temperatures further exacerbate the challenges faced by sea ice microorganisms. Overall bacterial production is low compared to other seasons, but higher than that of microalgae, leading to net heterotrophy (Grossmann and Dieckmann 1994 Wells and Deming 2006a, b). Following the initial freeze-up, the autumn to winter transition is characterised by decreasing light availability and temperature.

The low temperature constricts brine volume, decreasing the habitable space and forcing interactions (Deming & Eric Collins, 2016a). Little is known about the interactions between microorganisms in autumn and winter sea ice. Nevertheless, there is potential for both competition and collaboration. For instance, the restricted habitable space may heighten competition for certain resources if two organisms occupy the same metabolic niche, which would eventually lead to competitive exclusion (Zhou & Ning, 2017). However, one of the major sources of organic matter in sea ice brines is algal polysaccharides that are generally too manifold for a single bacterial species to decompose. Therefore, co-occurring species in sea ice brines may form metabolic collaborations to utilise complex algal polysaccharides, in a process termed resource-partitioning (Kabisch et al., 2014). This phenomenon has been investigated in algal blooms (Williams et al., 2013), but is yet to be explored in sea ice. While nutrient content is relatively high in sea ice brines, which may drive competition in cold environments, as recently observed in terrestrial snowpacks (Bergk Pinto, Maccario, Dommergue, Vogel, & Larose, 2019a), complexity of substrate availability and activity constraining temperatures and salinities could be conducive to a more collaborative community. Microorganisms may also alter each other's metabolism to suit their needs, by forming biofilms during the winter (Thomas & Dieckmann, 2002). This mutualistic arrangement has been hypothesized to be a dominant survival strategy in extreme environments (Ramanan, Kim, Cho, Oh, & Kim, 2016).

The lack of habitable space may also increase bacterial interactions with eukaryotes and viruses. Whilst grazing is thought to have minimal impact on sea ice communities (Gowing et al., 2004), viruses are enriched as much as 100 times relative to the underlying seawaterboth in spring (Maranger, Bird, & Kim Juniper, 1994) and autumn ice (Collins & Deming, 2011b). VBRs of 10–72 were observed in spring in the Arctic (Maranger et al., 1994), 0.7–119 in Antarctic summer pack ice (Gowing et al., 2004), which translates to activity maxima and in turn facilitates viral production rates. The only study on autumn ice showed extremely high VBRs (mean 846, maximum) that suggests preferential entrainment or active production of viruses within the ice (Collins & Deming, 2011c). Contrastingly, VBRs of 3 to 18 were observed in Arctic winter ice (Wells & Deming, 2006). Lysogenic cycles could dominate over the cold months in line with low production rates or winter sea-ice organisms may possess phage resistance mechanisms that aid their stability in sea ice. Yu et al postulated that phage increased host fitness during nutrient-limited polar winters by slowing down host metabolism and increasing its capacity to find new nutrient sources within the heterogeneous structure of the sea ice (Yu et al., 2015). Whilst culturing efforts have characterised sea ice viruses, identified their narrow-host ranges and dependency on cold temperatures for infection (Borriss, Helmke, Hanschke, & Schweder, 2003; Luhtanen et al., 2014; Senčilo, Luhtanen, Saarijärvi, Bamford, & Roine, 2015), little is known about the extent of host defence systems. The low frequency of CRISPR-spacers in cold-adapted sea ice isolates led to the hypothesis that EPS may be a primary host defence mechanism (Feng, Powell, Wilson, & Bowman, 2013). However, whether this translates to natural sea ice communities is yet to be explored.

Innate adaptions that allow organisms to persist and even remain active over the constraining autumn and winter months, such as cold-active proteins, membrane stability and cryoprotectants have been detailed elsewhere (Rodrigues & Tiedje, 2008). However, it is clear that pioneers of new sea ice are further selected by their capacity to survive and grow in the following months. Bacterial interactions, both inter and intra domain, may be crucial to shaping microbial community dynamics and in turn, set the platform for the more productive spring and summer months.

1.4.2 Selective conditions governing microbial succession (spring and summer)

The prevailing physiochemical conditions change following seasonal increase in solar radiation and temperature. These conditions lend to spring sea ice algal blooms and as such see a marked increase in primary production relative to the heterotrophic dominance of autumn and winter (Vonnahme et al., 2020). Algal blooms offer a large space for cross-domain interactions which could be a major driver of sea ice bacterial communities, especially in the lower portion of the ice where the blooms dominate. While under ice production increases via algal primary production, the upper ice layers experiences even greater heterotrophic activity attributed to the bacterial portion of the community (Rolf Gradinger & Zhang, 1997). Increasing temperatures facilitate utilisation of old organic matter or EPS, now deemed inessential in its protective role (Arrigo et al., 2014; Deming & Eric Collins, 2016a; Vonnahme et al., 2020)

While the relationships between algae and bacteria are generally appreciated across microbial ecology, their coupling in sea ice is uncertain. In the winter and autumn months when primary productivity is low, there may be an uncoupling of bacteria and algae (Stewart & Fritsen, 2004). However, the strength of the relationship may be more prominent in the spring and summer months when primary production is high and in turn, is quickly funnelled through the microbial loop (Vonnahme et al., 2020). While microbial community structure and succession has been intimately linked to specific bloom stages in cold pelagic habitats, reflecting complex inter-domain interactions, this has yet to be explored in a sea ice environment (Fadeev et al., 2018; Rapp, Fernández-Méndez, Bienhold, & Boetius, 2018). However, specific interactions driven by cobalamin production in bacteria is deterministic for phytoplankton growth in coastal Antarctic sea ice edge communities underpinning the importance of cross-domain co-operation in primary productivity (Bertrand et al., 2015).

At the surface of the ice where heterotrophy still reigns, the marked difference in the physiochemical-environment will change the dynamics of bacterial-bacterial interactions. For instance, the now rapidly depleting organic matter which was subject to slow degradation over the colder months may promote increased competition amongst individuals with similar nutritional requirements. This is reflected in the limited cross-season data available where taxa renowned for their ability to rapidly utilise substrates increase in abundance (Kaartokallio et al., 2008). New ecological niches arise as a result of increased activity, providing further opportunity for novel relationships. For instance, the increased rates of organic matter into ammonium and nitrate will deplete oxygen levels resulting in pockets of anoxia where denitrification and annamox are predicted to have a role (Rysgaard & Glud, 2004; Rysgaard, Glud, Sejr, Blicher, & Stahl, 2008) .

1.4.3 Multiyear ice

Longer temporal scales result in larger marked differences in physiochemical conditions in sea ice. This manifests as two distinct sea ice types; multiyear ice (MYI) defined as sea ice that persists more than one melt season and first year ice (FYI) which undergoes complete melt during the Spring and Summer. These two ice types provide different habitat characteristics (e.g., thickness, ice bulk salinities, age, and albedo) that can impact biological processes (Weeks and Ackley 1986). Both MYI and FYI are subject to seasonal fluctuations however, meterological conditions can have contrasting effects on the ice types. For instance, where FYI undergoes complete melt in the warmer season, meltwater ponds (from both ice and snow) form on the surface of MYI and can percolate the ice matrix (Deming & Eric Collins, 2016b). MYI provides a more stable environment relative to FYI which can be characterised by its steeper gradients in temperature, light, brine volume and salinity (Hatam, Lange, Beckers, Haas, & Lanoil, 2016). The consistency of conditions in MYI is reflected in the stability of its bacterial community membership and composition (Hatam et al., 2016). Adding complexity to the matter is the stratification of multiyear ice consisting of at least two layers; an old layer that survives the melt season and a new bottom layer formed during seasonal freeze-up. Clear differences in the bacterial assemblages that colonise these layers have been evidenced and mostly attributed to ice origin with adaptations to local conditions playing a relatively minor role (Hatam et al., 2014). However, the functional potential of MYI communities or how the change in community composition and the relative stability of conditions affects microbial interactions has not been studied.

1.4.4 Interaction via horizontal gene transfer

Organisms inhabiting the brine channels of sea ice will experience greater cell-cell contact rates as a consequence of limited space and high cell density estimates relative to the sparsity of pelagic living. Cell density has been highlighted as a major determinant of conjugation mediated horizontal gene transfer (HGT) in other environments (Elsas & Bailey, 2002) and as such could offer a significant form of interaction and adaptation for microorganisms inhabiting space-limited brines. Although conjugation rates have not been determined in sea ice, circumstantial evidence through the identification of plasmids in archetypal sea ice microorganisms provides a basis for its role. For example, 30% of bacteria isolated from Antarctic sea ice were shown to contain plasmids (Kobori, Sullivan, & Shizuya, 1984) and some bacterial isolates of Arctic brine carry plasmids conferring resistance to mercury (Møller et al., 2014). Furthermore, a large extracellular DNA pool that could mediate transformation was measured in sea ice, but the authors suggested that this process might be of limited importance as modelled diffusivity of free DNA is at least an order of magnitude lower than that of viral-encapsulated DNA (Collins & Deming, 2011a). Nevertheless, a microcosm study showed that transformation efficiency was enhanced by 50X at the solid-liquid interface than in the liquid alone (Lorenz, Aardema, & Wackernagel, 1988), which further supports the hypothesis of the sea ice as a hotspot for HGT. However, given the high virus to bacterial concentrations in sea ice, contact with viruses is likely to occur more frequently. therefore transduction might be a more dominant HGT pathway in active sea ice communities. Finally, evidence for HGT events in sea ice microorganisms have been recorded. For instance, it is likely that HGT was responsible for the inter-domainacquisition of ice-binding proteins (Janech, Krell, Mock, Kang, & Raymond, 2006; Raymond, Fritsen, & Shen, 2007) and antifreeze proteins (Bayer-Giraldi, Uhlig, John, Mock, & Valentin, 2010; Kiko, 2010) across the domains of life.

1.4.5 The physical-chemical environment and succession dynamics

Selective pressures imposed by the physiochemical environment can control microbial community dynamics. The environmental constraints imposed on sea ice microorganisms are extremely variable at both spatial and temporal scales and in turn could govern how microorganisms partition themselves. Understanding how populations and communities react to changing conditions will help us predict community shifts as a consequence of the changing dominant ice types in the Arctic. However, studies during the autumn freeze-up and throughout the winter are rare. Moreover, although a handful of studies have identified taxonomic variances in community structure in different seasons, bacterial and algal functional diversity remains largely unexplored.

1.5 Next-generation sequencing as a tool in sea ice succession ecology

To surpass the limitations presented by traditional methods such as culturing and microscopy, sequencing technologies have been established providing a platform to infer phylogenetic relationships and functional capacity. The NGS technologies differ from the earlier Sanger sequencing methodologies in that they provide massively parallel analysis, high throughput production at relatively low costs. NGS technologies can be applied to generate genomes by whole genome sequencing (WGS) of cultured microorganisms or even to study the total genomic content of microbial communities that bridges the three domains of life without the need to culture.

1.5.1 Whole genome sequencing (WGS)

The advent of NGS has enabled the easier, faster and cheaper construction of complete genomes. WGS from isolated bacteria provide important insights into ecosystem functioning. For instance, WGS provides a means to assess how adaptation and evolution impact genomic change or why particular organisms prosper under certain environmental conditions. Analysing entire genomes provides a resolution which enables discrimination between closely related populations (Delmont & Eren, 2018). In turn, comparative genomics of populations can highlight the degree to which individuals are representative of their associated populations and the importance of processes such as horizontal gene transfer (HGT) and genomic re-arrangement. For instance, genome sequencing of the archetypal sea ice microorganism Colwellia psychrerythraea 34H provided insight into adaptive lifestyle strategies such as the synthesis of compounds that may confer cryotolerance (e.g. PHA's, cyanophycin-like compounds, glycine betaine) and the capacity to produce large amounts of extracellular enzymes (Methé et al., 2005). In a separate study, comparative genomics of three strains of C. psychrerythraea 34H demonstrated a large amount of 16S rRNA gene synteny yet distinct carbon source utilisation profiles reflective of their isolation site (Techtmann et al., 2016). A comparative genomic approach has also been implemented to resolve the differences in two closely related species of *Octadecabacter*, isolated from sea ice at opposite poles. Despite sharing >99% similarity in their 16S rRNA gene, less than 80% of their gene content was comparable (Vollmers et al., 2013). Moreover, the diversity and abundance of transposable elements in both genomes indicate the importance of genome plasticity and mobile genetic elements on their evolution.

Not all the aforementioned studies were performed using NGS technologies however, their application to WGS will help resolve a larger number of genomes underrepresented in the current databases. Moreover, the ability to rapidly and cheaply generate genomes from closely related strains will provide a larger platform for genome-genome comparison studies aiding the ability to investigate the selective pressures that drive processes such as microbial biogeography on a genomic level.

1.5.2 Marker Gene Sequencing (MGS)

MGS refers to the amplification of sequence fragments from DNA isolated directly from the environment without the prerequisite of culturing microorganisms. Typically, the 16S rRNA genes and 18S rRNA/ribosomal internal transcribed spacer (ITS) genes are targeted to infer phylogeny of bacterial and eukaryotic communities, respectively. The advent of NGS in 2005 and the corresponding decrease in sequencing costs has led to a dramatic increase in MGS studies in microbial ecology. However, its application to sea ice has been limited and most MGS studies of sea ice have implemented pre-NGS technologies such as ARISA and RFLP (Table 1). These studies have enabled an initial characterisation of microbial biota of sea ice and highlighted community variability in relation to space (Brinkmeyer et al., 2003; Yergeau et al., 2017), time (Eric Collins, Rocap, & Deming, 2010) and ice type (Hatam, Lange, Beckers, Haas, & Lanoil, 2016). Moreover, the repetitive selection of rare marine 16S rRNA phylotypes in sea ice independent of geography strongly suggest a niche-based selection mechanism. The reasons underpinning taxa-specific enrichment are yet to be fully realised.

Sequencing target and reference	Ice and sequencing type	Location	Purpose of study	Main findings
16S rRNA (Jeff S. Bowman et al., 2012)	MYI 454	Arctic	Taxonomic profiling	**high degree of diversity, identification of novel clades not previously found in sea ice**
16S rRNA (Hatam et al., 2014)	MYI FLX pyro-sequencing	Arctic	Taxonomic shifts in bacteria in relation to sea ice depth	**Bacterial community structure is defined by layers which correspond to ice type**
16S and 18S rRNA (Torstensson et al., 2015)	FYI (pack) 454	Antarctica	Link microbial diversity and environment	**Bacterial richness is strongly correlated to temperature and salinity**
16S rRNA (Eronen-Rasimus et al., 2015)	First year drift ice T-RFLP, Sanger	Sub-Arctic	Bacterial community structure in developmental stages of drift ice from open water to consolidated sea ice.	**temporal development of sea ice bacterial communities during formation and consolidation**

16S rRNA (Brown & Bowman, 2006)	FYI, MYI RFLP Sanger	Antarctica, Arctic	Determine if algal assemblages result in shifts in bacterial community structure	**Common bacterial genera at both poles, dominance of bacterial heterotrophic taxa, variety of autotrophic heterotrophic nanoplankton**, **increasing algal density** linked to **increasing proportion of CFB clones**
16S rRNA (Maas et al., 2012)	FYI RFLP Sanger	Antarctica	Identify dominant species in Antarctic sea ice	**Dominance of chemoheterotrophs and potential phototrophs**
16S rRNA, metagenomes (Yergeau et al., 2017)	FYI Ion Sequencing	Arctic	Characterise bacterial community structure of the bottom layer of sea ice and seawater and relate taxonomic and functional profiles to environmental parameters	**Seawater and sea ice exhibited large differences in terms of chemistry, community composition and functional gene content, dominance of algae in sea ice samples**, **salinity and nutrients** correlate with **abundance of various microbial taxa**
16S rRNA (R. Cowie, Williams, Maas, Voyles, & Ryan, 2014)	Land-fast sea ice T-RFLP, Sanger	Antarctica	Identify environmental drivers of bacterial community structure and function. Identify the role bacteria play in the sea-ice microbial loop	**Nutrient concentrations, bacterial community structure and metabolic activity stratify in sea ice. Uncoupling of bacteria and algae evidenced by a lack of correlation between bacterial numbers and chl a**
16S rRNA (J. S. Bowman, Larose, Vogel, & Deming, 2013b)	Frost flowers, FYI Microarray	Arctic	Determine whether microbial community structures varied between sea ice and frost flowers	**Distinct community structures of young sea ice and frost flowers observed by a taxonomic enrichment of Rhizobiales in frost flowers**
18S rRNA (Bachy, López-García, Vereshchaka, & Moreira, 2011)	MYI Sanger	Arctic	Compare the diversity and vertical distribution of microbial eukaryotes of sea ice, overlying snow and underlying seawater	**Eukaryotic communities were distinctive along a vertical gradient**. **Sea ice and seawater were dominated by heterotrophic dinoflagellates** with **little representation of phototrophic organisms**.

				Identification of the possible importance of mixotophy
Metagenomes (Maccario et al., 2019)	FYI Illumina Miseq	Arctic	Relate functional potential to selection processes in seawater, sea ice, snow and the atmosphere	**Heterotrophy dominated all layers of the profile but presence Cyanobacteria in the surface snow suggests a role for primary production. Highlighted the importance of saline snow as a unique microbial habitat. Photochemical stress** is **a major driver in the upper** snow **layers**
16S rRNA (R. O. M. Cowie, Maas, & Ryan, 2011)	Land-fast sea ice Sanger	Antarctica	Detect and quantify archaeal abundance and diversity	**Archaea are present but represent low abundant organisms in sea ice**. The **vast majority of Archaea are from the phylum Thaumarchaeota**
Proteorhodopsin (PR) genes (Koh et al., 2010)	N/A Sanger	Antarctica	Detect present and active PR genes in sea ice bacteria	**sea ice-derived bacterial sequences are distinct and may reflect physiological adaptation. Sea ice bacteria are actively expressing PR mRNA**
Metagenomes (Jeff S. Bowman et al., 2014)	Frost flowers, FYI Illumina Miseq	Arctic	Understand the functional potential of bacterial communities of frost flowers, specifically the enriched genera Rhizobiales	**Partitioning of metabolic function between sea ice and frost flowers. Potential for DMS and betaine glycine catabolism, and halocarbon production in Rhizobiales. Overall community capable of DMSP catabolism and mercury reduction**
16S rRNA (Hatam et al., 2016)	MYI, FYI FLX-Titanium amplicon pyrosequencing	Arctic	Compare bacterial community structure of FYI and MYI	**Larger variability in community membership and structure of FYI relative to MYI**
18S rRNA (Stecher et al., 2016)	MYI, FYI 454	Arctic	Survey the present and active community of sea ice and compare MYI and FYI	**Sea ice protist communities varied based on ice type**. **Active members of the community were disproportionately under-represented when only analysing total community structure**

18S rRNA (Majaneva, Rintala, Piisilä, Fewer, & Blomster, 2012)	Pancake ice, land-fast ice, pack ice Sanger	Sub-Arctic	Survey the nano-sized eukaryotic in winter sea ice	**Sea ice protist diversity is high and dominated by heterotrophic flagellates (underestimated by microscopy)**. **Highlights the importance of grazers and parasitic dinoflagellate (Syndiniales) and fungal richness**
18S rRNA **(Comeau et al., 2013)**	Drift ice, land-fast ice	Arctic	Compare protist communities of land-fast and drift ice	**Diverse community of heterotrophic and mixotrophic eukaryotes**. **Key taxa identified as marker taxa for differing ice types**
18S rRNA **(Belevich et al., 2018)**	Land-fast sea ice Illumina MiSeq	Sub-Arctic	composition and abundance of photosynthetic picoeukaryotes	Dominance of **Chlorophyta**. Identification of a new clade of Micromonas with a wider distribution than previously identified.
18S rRNA **(Gast et al., 2006)**	FYI, water column, meltwater layer (slush)	Antarctica	Assess biodiversity of sea ice, slush and seawater and isolate novel organisms	**Identification and isolation of novel and abundant dinoflagellate**

In line with the development of sequencing technologies is the rapid advancement of bioinformatic tools. Microbe-environment interactions can be inferred from relatively straightforward means such as through correlation analyses of measured environmental variables and taxa abundance. On the other hand, microbe-microbe interactions are ecologically important but generally less studied relative to microbe-environment relationships. Correlation-based network analyses have gained popularity as a result of the larger sequencing datasets and increasing availability of bioinformatic tools specific to amplicon data. Identifying putative relationships via network analyses has helped disentangle complex microbial interactions and in turn recognise potential metabolic associations, keystone species or clusters of co-occurring taxa over temporal and spatial scales (Bergk Pinto, Maccario, Dommergue, Vogel, & Larose, 2019b; Chafee et al., 2018; Milici et al., 2016; Peura, Bertilsson, Jones, & Eiler, 2015). While these approaches have been applied to microbial communities of various environments, these methods are yet to be applied to sea ice data. Furthermore, cross-domain interactions have a fundamental importance in marine ecosystems (Fadeev et al., 2018) and their interactions affect biogeochemical cycling and ultimately the climate (Ramanan et al., 2016). The importance of these interactions may be emphasised in sea ice where reduced habitable space forces greater interactions. However, bacteria and eukaryotes are seldom studied simultaneously in sea ice.

1.5.3 Metagenomics

Shotgun metagenomic sequencing refers to the untargeted sequencing of all microbial genomes present in a sample and as such can be used to simultaneously investigate the taxonomic composition and functional potential of microbial communities. Where metagenomic surveys have been routinely employed in studies of diverse habitats, its application to sea ice environments remains limited. Nevertheless, the first metagenomics study was based on a single metagenome from FYI and one from a frost flower sample (Bowman et al., 2014). Although it is difficult to establish the role of microbial communities in sea ice and frost flowers from a single sample, the study did demonstrate some functions of potential interest. For instance, the overall community was capable of DMSP catabolism and mercury reduction which could have important consequences on polar biogeochemical cycling. A larger biogeographic study highlighted possible drivers of community function and a clear segregation of sea ice and seawater both in terms of taxonomy and function (Yergeau et al., 2017). However, the study was limited to the bottom portion of sea ice, yet vertical stratification, at least taxonomically, is a well described phenomenon (Hatam et al., 2014). Finally, the latest metagenomic study of sea ice was the first to assess a complete profile of sea ice including the overlying snow and atmosphere (Maccario et al., 2019). Microbial communities were shown to be vertically stratified both in terms of their taxonomic and functional characteristics, with snow being more similar to atmospheric communities than to sea ice, brine or seawater communities. Snow communities showed a high versatility in carbon source utilization, with genes related to the degradation of starch, sucrose and some xenobiotic compounds. This study also identified the saline snow layer as a unique niche, potentially fed with nutrients resulting from brine expulsion, that likely supports the growth of copiotrophic, psychro- and halotolerant snow members such as Glaciecola (Maccario et al., 2019). However, this study was based on a single site in Greenland and further research is required to validate this hypothesis.

Reference genomes provide information regarding metabolic properties, phylogeny, functionality, biogeochemical cycling and interactions and yet only a handful exist for sea ice microorganisms (Deming & Eric Collins, 2016a). Although sea ice bacteria are generally considered amenable to culture, the lack of comprehensive genome sequences and the bias toward certain taxa limits the scope of study. Additionally, culturing microorganisms can force changes in genomic content and in turn circumscribe its applicability to natural ecosystems. Advances in bioinformatics tools have provided the ability to assemble draft genomes from metagenomic datasets. Multiple metagenomic assembled genomes (MAGs) can be recovered from datasets and are often distantly related to any isolated species, offering a clear advantage over traditional approaches (Barnum et al., 2018; Delmont, Murat Eren, Vineis, & Post, 2015; Wilkins, Ettinger, Jospin, & Eisen, 2019). Therefore, this line of study provides an opportunity to resolve genomes from sea ice representatives currently lacking in the databases such as Flavobacterium, Polaribacter, Pelagibacter, Loktanella and Sulfitobacter (Deming & Eric Collins, 2016a)and could provide a novel insight into the biogeochemical processes taking place in sea ice. For instance, the complete oxidation of ammonia to nitrate via nitrite (i.e., comammox) was determined by the detection of necessary genes in a single MAG (van Kessel et al. 2015). These tools might provide insight into cross-domain interactions, such as resource and niche partitioning, via the study of specialized genomic islands (Kabisch et al., 2014), or uncover adaptive/evolutionary responses to environmental change via HGT (Feng et al., 2013).

1.6 Conclusion

Successional dynamics of sea ice and snow microbiomes are not well understood but are crucial to predicting ecosystem development. During primary succession, pioneering taxa are potentially motile copiotrophs that actively seek out the nutrient rich brine environment, epibionts that hitchhike with larger algal cells and/or EPS-producing and IBP/INP-producing organisms that enhance their ice-entrainment rates through sticking or binding to ice-crystals. Early-colonising taxa that produce EPS physically shape available niche space whilst also contributing to the pool of organic matter available for heterotrophic activity. They must also be organisms that are able to withstand the strong changes in physio-chemical parameters both on induction to the sea ice and as the seasons change. Additional horizons of emerging importance (snow and frost flowers) may act to change the biological, chemical and physical dynamics of the sea ice profile and other ecosystems through dispersal. Therefore, succession ecology in sea ice could benefit from three major research directions.

Firstly, we must disentangle the factors that drive primary colonisation and begin to dispel hypotheses with which evidence to the contrary is gathering. For instance, bacterial incorporation driven through algal attachment has little direct evidence, research that does not support a correlation and is inconsistent with low rates of bacteria-algal attachment in pelagic systems. Instead, we should focus on promising hypotheses such as linking motility or the ability to produce EPS and sea ice incorporation.

Next, successional dynamics of sea ice microbiomes are related to changes in the physiochemical environment; such as UV intensity, salinity and substrate availability. The relative contributions of these determinants represent fundamental knowledge gaps in understanding the trajectories in ecosystem development and will be crucial for climate change predictions.

Finally, emerging environments such as saline snow and frost flowers are rich with biomass and increasing in spatial and temporal extent. The handful of studies available suggest specific taxa-enrichment in both these interface environments. However, the impact of this dramatic drop in diversity is unknown. Future studies should seek to understand the underlying ecological drivers for this enrichment in order to better predict ecosystem-level changes that are accelerated by climate change.

1.7 Research objectives and hypotheses

The overarching aim of this thesis was to assess microbial colonisation, community adaptation and dynamics in relation to sea ice formation and the developing habitats at different spatial and temporal scales. Given that factors such as the prevailing conditions, local geography and ocean currents can have a large impact on sea ice formation, they can obscure our understanding of how microbial life incorporates and assembles in sea ice and its associated environments. In order to understand how both eukaryotic and bacterial populations established and partitioned in young sea ice and frost flowers, we used a controlled sea-ice chamber experiment (Chapter 1). A secondary objective was to determine whether bacteria of the rare marine biosphere are selected for within sea ice. We suggest that colonisation and cell enrichment in sea ice brines is not a product of algal association, but instead a function of post-formation selection mechanisms related to the prevailing environmental conditions. We hypothesize that the structuring variables that drive sea ice communities are universal and that seawater not naturally subject to annual sea ice formation would mirror taxonomic and functional selection of natural assemblages.

To further study the underlying selection methods within sea ice communities, we carried out two field studies in Greenland. In Chapter 2, we focus on first year sea ice and hypothesize that the functional signature of bacterial and eukaryotic communities coupled to physico-chemical measurements could help infer dominant environmental and biological drivers of community structure. We predict that, in addition to changes in salt concentrations and organic content (including complex carbohydrates and sugars) and other physico-chemical drivers, such as UV, communities are subjected to top down and bottom up control. This research was extended to include multi-year and land fast ice communities in Chapter 3. We hypothesize that the presence of brine is dominant structuring variable in sea ice communities that selects for microorganisms with copiotrophic lifestyles.

Frost flowers and brine fed snow represent ecosystems of emerging importance in line with the transition from a MYI dominated landscape to one in which FYI prevails. A taxa-specific enrichment has been evidenced in both these environments. We hypothesize that this phenomenon is recurrent across time and space. Chapter 1 addresses taxa-specific enrichment in artificial frost flowers whilst chapters 2 and 3 validate this observation in naturally occurring saline snow horizons. The genomic determinants that underscore the enrichment and coupled decline in diversity are addressed through functional profiling of the seawater-sea ice-snow-atmosphere continuum. We extend our understanding of post-depositional processes in snow overlying sea ice in chapter 3 by comparing sea ice and snow profiles from geographically distinct MYI and FYI sites.

Finally, utilising the knowledge garnered in chapters 2 and 3, we aimed to recover genomes of dominant saline snow associated taxa. We utilised modern bioinformatic tools to assemble draft genomes from available metagenomic data and placed these genomes in the context of their environment by comparisons to closely related organisms for which genome sequences were available. We hypothesize that in order to inhabit the saline snow, these organisms must either acquire, or already contain, functions that would provide them with a selective advantage, such as mechanisms to encounter UV stress.

1.8 Bibliography

Alexandre, G. (2010, August 1). Coupling metabolism and chemotaxis-dependent behaviours by energy taxis receptors. *Microbiology*, Vol. 156, pp. 2283–2293. https://doi.org/10.1099/mic.0.039214-0

Arrigo, K. R., Perovich, D. K., Pickart, R. S., Brown, Z. W., van Dijken, G. L., Lowry, K. E., … Swift, J. H. (2014). Phytoplankton blooms beneath the sea ice in the Chukchi sea. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *105*, 1–16. https://doi.org/10.1016/j.dsr2.2014.03.018

Bachy, C., López-García, P., Vereshchaka, A., & Moreira, D. (2011). Diversity and Vertical Distribution of Microbial Eukaryotes in the Snow, Sea Ice and Seawater Near the North Pole at the End of the Polar Night. *Frontiers in Microbiology*, *2*(MAY), 106. https://doi.org/10.3389/fmicb.2011.00106

Barber, D. G., Ehn, J. K., Pućko, M., Rysgaard, S., Deming, J. W., Bowman, J. S., … Søgaard, D. H. (2014a). Frost flowers on young Arctic sea ice: The climatic, chemical, and microbial significance of an emerging ice type. *Journal of Geophysical Research: Atmospheres*, *119*(20), 11,593-11,612. https://doi.org/10.1002/2014JD021736

Barber, D. G., Ehn, J. K., Pućko, M., Rysgaard, S., Deming, J. W., Bowman, J. S., … Søgaard, D. H. (2014b). Frost flowers on young Arctic sea ice: The climatic, chemical, and microbial significance of an emerging ice type. *Journal of Geophysical Research: Atmospheres*, *119*(20), 11,593-11,612. https://doi.org/10.1002/2014JD021736

Barber, D. G., Reddan, S. P., & LeDrew, E. F. (1995). Statistical characterization of the geophysical and electrical properties of snow on Landfast first-year sea ice. *Journal of Geophysical Research*, *100*(C2), 2673. https://doi.org/10.1029/94JC02200

Barnum, T. P., Figueroa, I. A., Carlström, C. I., Lucas, L. N., Engelbrektson, A. L., & Coates, J. D. (2018). Genome-resolved metagenomics identifies genetic mobility, metabolic interactions, and unexpected diversity in perchlorate-reducing communities. *ISME Journal*, *12*(6), 1568–1581. https://doi.org/10.1038/s41396-018-0081-5

Bauer, M. A., Kainz, K., Carmona-Gutierrez, D., & Madeo, F. (2018, May 1). Microbial wars: Competition in ecological niches and within the microbiome. *Microbial Cell*, Vol. 5, pp. 215–219. https://doi.org/10.15698/mic2018.05.628

Bayer-Giraldi, M., Uhlig, C., John, U., Mock, T., & Valentin, K. (2010). Antifreeze proteins in polar sea ice diatoms: Diversity and gene expression in the genus Fragilariopsis. *Environmental Microbiology*, *12*(4), 1041–1052. https://doi.org/10.1111/j.1462-2920.2009.02149.x

Beinart, R. A. (2019). The Significance of Microbial Symbionts in Ecosystem Processes. *MSystems*, *4*(3). https://doi.org/10.1128/msystems.00127-19

Belevich, T. A., Ilyash, L. V, Milyutina, I. A., Logacheva, & M. D., Goryunov, & D. V, & Troitsky, & A. V. (2018). Photosynthetic Picoeukaryotes in the Land-Fast Ice of the White Sea, Russia. *Microbial Ecology*, *75*, 582–597. https://doi.org/10.1007/s00248-017-1076-x

Ben-Jacob, E., Cohen, I., & Levine, H. (2000). Cooperative self-organization of microorganisms. *Advances in Physics*, *49*(4), 395–554. https://doi.org/10.1080/000187300405228

Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M., & Larose, C. (2019a). Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.02492

Bergk Pinto, B., Maccario, L., Dommergue, A., Vogel, T. M., & Larose, C. (2019b). Do Organic Substrates Drive Microbial Community Interactions in Arctic Snow? *Frontiers in Microbiology*, *10*, 2492. https://doi.org/10.3389/fmicb.2019.02492

Bertrand, J. C., Bonin, P., Caumette, P., Gattuso, J. P., Grégori, G., Guyoneaud, R., … Poly, F. (2015). Biogeochemical cycles. In *Environmental Microbiology: Fundamentals and Applications* (pp. 511–617). https://doi.org/10.1007/978-94-017-9118-2_14

Białkowska, A., Majewska, E., Olczak, A., & Twarda-Clapa, A. (2020). Ice Binding Proteins: Diverse Biological Roles and Applications in Different Types of Industry. *Biomolecules*, *10*(2), 274. https://doi.org/10.3390/biom10020274

Boetius, A., Anesio, A. M., Deming, J. W., Mikucki, J. A., & Rapp, J. Z. (2015, November 1). Microbial ecology of the cryosphere: Sea ice and glacial habitats. *Nature Reviews Microbiology*, Vol. 13, pp. 677–690. https://doi.org/10.1038/nrmicro3522

Borriss, M., Helmke, E., Hanschke, R., & Schweder, T. (2003). Isolation and characterization of marine psychrophilic phage-host systems from Arctic sea ice. *Extremophiles*, *7*(5), 377–384. https://doi.org/10.1007/s00792-003-0334-7

Bowman, J. P. (2013). Sea-ice microbial communities. In *The Prokaryotes: Prokaryotic Communities and Ecophysiology* (Vol. 9783642301230, pp. 139–161). https://doi.org/10.1007/978-3-642-30123-0_46

Bowman, J. S., & Deming, J. W. (2010). Elevated bacterial abundance and exopolymers in saline frost flowers and implications for atmospheric chemistry and microbial dispersal. *Geophysical Research Letters*, *37*(13), n/a-n/a. https://doi.org/10.1029/2010GL043020

Bowman, J. S., Larose, C., Vogel, T. M., & Deming, J. W. (2013a). Selective occurrence of Rhizobiales in frost flowers on the surface of young sea ice near Barrow, Alaska and distribution in the polar marine rare biosphere. *Environmental Microbiology Reports*, *5*(4), 575–582. https://doi.org/10.1111/1758-2229.12047

Bowman, J. S., Larose, C., Vogel, T. M., & Deming, J. W. (2013b). Selective occurrence of *Rhizobiales* in frost flowers on the surface of young sea ice near Barrow, Alaska and distribution in the polar marine rare biosphere. *Environmental Microbiology Reports*, *5*(4), 575–582. https://doi.org/10.1111/1758-2229.12047

Bowman, Jeff S. (2015, October 13). The relationship between sea ice bacterial community structure and biogeochemistry: A synthesis of current knowledge and known unknowns. *Elementa*, Vol. 3. https://doi.org/10.12952/journal.elementa.000072

Bowman, Jeff S., Berthiaume, C. T., Armbrust, E. V., & Deming, J. W. (2014). The genetic potential for key biogeochemical processes in Arctic frost flowers and young sea ice revealed by metagenomic analysis. *FEMS Microbiology Ecology*, *89*(2), 376–387.

https://doi.org/10.1111/1574-6941.12331

Bowman, Jeff S., Rasmussen, S., Blom, N., Deming, J. W., Rysgaard, S., & Sicheritz-Ponten, T. (2012). Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene. *ISME Journal*, *6*(1), 11–20. https://doi.org/10.1038/ismej.2011.76

Brown, M. V., & Bowman, J. P. (2006). A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiology Ecology*, *35*(3), 267–275. https://doi.org/10.1111/j.1574-6941.2001.tb00812.x

Bull, A. T., & Goodfellow, M. (2019). Dark, rare and inspirational microbial matter in the extremobiosphere: 16 000 m of bioprospecting campaigns. *Microbiology (United Kingdom)*, Vol. 165, pp. 1252–1264. https://doi.org/10.1099/mic.0.000822

Burrows, S. M., Elbert, W., Lawrence, M. G., & Pöschl, U. (2009). and Physics Bacteria in the global atmosphere – Part 1 : Review and synthesis of literature data for different ecosystems. *Atmospheric Chemistry and Physics*, *9*(1987), 9263–9280. https://doi.org/10.5194/acp-9-9263-2009

Chafee, M., Fernàndez-Guerra, A., Buttigieg, P. L., Gerdts, G., Eren, A. M., Teeling, H., & Amann, R. I. (2018). Recurrent patterns of microdiversity in a temperate coastal marine environment. *ISME Journal*, *12*(1), 237–252. https://doi.org/10.1038/ismej.2017.165

Chow, J., Tang, H., & Mazmanian, S. K. (2011, August). Pathobionts of the gastrointestinal microbiota and inflammatory disease. *Current Opinion in Immunology*, Vol. 23, pp. 473–480. https://doi.org/10.1016/j.coi.2011.07.010

Collins, R. E., & Deming, J. W. (2011a). Abundant dissolved genetic material in Arctic sea ice Part I: Extracellular DNA. *Polar Biology*, *34*(12), 1819–1830. https://doi.org/10.1007/s00300-011-1041-y

Collins, R. E., & Deming, J. W. (2011b). Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up. *Polar Biology*, *34*(12), 1831–1841. https://doi.org/10.1007/s00300-011-1008-z

Collins, R. E., & Deming, J. W. (2011c). Abundant dissolved genetic material in Arctic sea ice

Part II: Viral dynamics during autumn freeze-up. *Polar Biology*, *34*(12), 1831–1841. https://doi.org/10.1007/s00300-011-1008-z

Comeau, A. M., Philippe, B., Thaler, M., Gosselin, M., Poulin, M., & Lovejoy, C. (2013). Protists in Arctic drift and land-fast sea ice. *Journal of Phycology*, *49*(2), 229–240. https://doi.org/10.1111/jpy.12026

Cowie, R. O. M., Maas, E. W., & Ryan, K. G. (2011). Archaeal diversity revealed in Antarctic sea ice. *Antarctic Science*, *23*(6), 531–536. https://doi.org/10.1017/S0954102011000368

Cowie, R., Williams, G., Maas, E., Voyles, K., & Ryan, K. (2014). Antarctic sea-ice microbial communities show distinct patterns of zonation in response to algal-derived substrates. *Aquatic Microbial Ecology*, *73*(2), 123–134. https://doi.org/10.3354/ame01710

Crenn, K., Duffieux, D., & Jeanthon, C. (2018). Bacterial Epibiotic Communities of Ubiquitous and Abundant Marine Diatoms Are Distinct in Short- and Long-Term Associations. *Frontiers in Microbiology*, *9*. https://doi.org/10.3389/fmicb.2018.02879

D. Perovich1, W. Meier, M. Tschudi, S. Farrell, S. Hendricks, S. Gerland, … M. Webster. (2018). Sea Ice [in Arctic Report Card 2018]. Retrieved April 15, 2020, from https://www.arctic.noaa.gov/Report-Card

Dann, L., Clanahan, M., Paterson, J., & Mitchell, J. (2019). Distinct niche partitioning of marine and freshwater microbes during colonisation. *FEMS Microbiology Ecology*, *95*. https://doi.org/10.1093/femsec/fiz098

Delmont, T. O., & Eren, E. M. (2018). Linking pangenomes and metagenomes: The Prochlorococcus metapangenome. *PeerJ*, *2018*(1). https://doi.org/10.7717/peerj.4320

Delmont, T. O., Murat Eren, A., Vineis, J. H., & Post, A. F. (2015). Genome reconstructions indicate the partitioning of ecological functions inside a phytoplankton bloom in the Amundsen Sea, Antarctica. *Frontiers in Microbiology*, *6*(OCT). https://doi.org/10.3389/fmicb.2015.01090

Deming, J. W., & Eric Collins, R. (2016a). Sea ice as a habitat for Bacteria, Archaea and viruses. In *Sea Ice: Third Edition* (pp. 326–351).

https://doi.org/10.1002/9781118778371.ch13

Deming, J. W., & Eric Collins, R. (2016b). Sea ice as a habitat for Bacteria, Archaea and viruses. In *Sea Ice* (pp. 326–351). https://doi.org/10.1002/9781118778371.ch13

Domine, F., Sparapani, R., Ianniello, A., & Beine, H. J. (2004). The origin of sea salt in snow on Arctic sea ice and in coastal regions. *Atmospheric Chemistry and Physics*, *4*(9/10), 2259–2271. https://doi.org/10.5194/acp-4-2259-2004

Domine, Florent. (2005). Specific surface area, density and microstructure of frost flowers. *Geophysical Research Letters*, *32*(13), L13502. https://doi.org/10.1029/2005GL023245

Droop, M. R., & Elson, K. G. R. (1966). Are pelagic diatoms free from bacteria? [31]. *Nature*, Vol. 211, pp. 1096–1097. https://doi.org/10.1038/2111096a0

Elsas, J. D., & Bailey, M. J. (2002). The ecology of transfer of mobile genetic elements. *FEMS Microbiology Ecology*, *42*(2), 187–197. https://doi.org/10.1111/j.1574-6941.2002.tb01008.x

Eric Collins, R., Rocap, G., & Deming, J. W. (2010). Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Environmental Microbiology*, *12*(7), 1828–1841. https://doi.org/10.1111/j.1462-2920.2010.02179.x

Eronen-Rasimus, E., Kaartokallio, H., Lyra, C., Autio, R., Kuosa, H., Dieckmann, G. S., & Thomas, D. N. (2014). Bacterial community dynamics and activity in relation to dissolved organic matter availability during sea-ice formation in a mesocosm experiment. *MicrobiologyOpen*, *3*(1), 139–156. https://doi.org/10.1002/mbo3.157

Eronen-Rasimus, E., Lyra, C., Rintala, J. M., Jürgens, K., Ikonen, V., & Kaartokallio, H. (2015). Ice formation and growth shape bacterial community structure in Baltic Sea drift ice. *FEMS Microbiology Ecology*, *91*(2). https://doi.org/10.1093/femsec/fiu022

Ewart, K. V., Lin, Q., & Hew, C. L. (1999, February). Structure, function and evolution of antifreeze proteins. *Cellular and Molecular Life Sciences*, Vol. 55, pp. 271–283. https://doi.org/10.1007/s000180050289

Ewert, M., Carpenter, S. D., Colangelo-Lillis, J., & Deming, J. W. (2013). Bacterial and extracellular polysaccharide content of brine-wetted snow over Arctic winter first-year sea ice. *Journal of Geophysical Research: Oceans*, *118*(2), 726–735. https://doi.org/10.1002/jgrc.20055

Ewert, Marcela, & Deming, J. W. (2013). Sea ice microorganisms: Environmental constraints and extracellular responses. *Biology*, *2*(2), 603–628. https://doi.org/10.3390/biology2020603

Fadeev, E., Salter, I., Schourup-Kristensen, V., Nöthig, E. M., Metfies, K., Engel, A., … Bienhold, C. (2018). Microbial communities in the east and west fram strait during sea ice melting season. *Frontiers in Marine Science*, *5*(NOV), 429. https://doi.org/10.3389/fmars.2018.00429

Feng, S., Powell, S. M., Wilson, R., & Bowman, J. P. (2013). Extensive Gene Acquisition in the Extremely Psychrophilic Bacterial Species Psychroflexus torquis and the Link to Sea-Ice Ecosystem Specialism. *Genome Biol. Evol. 6(1):133–148.*, *6*(1), 133–148. https://doi.org/10.1093/gbe/evt209

Fierer, N., Nemergut, D., Knight, R., & Craine, J. M. (2010). *Changes through time: integrating microorganisms into the study of succession*. https://doi.org/10.1016/j.resmic.2010.06.002

Garrison, D. L., Close, A. R., & Reimnitz, E. (1989). Algae concentrated by frazil ice: Evidence from laboratory experiments and field measurements. *Antarctic Science*, *1*(4), 313–316. https://doi.org/10.1017/S0954102089000477

Garrison, David L., Ackley, S. F., & Buck, K. R. (1983). A physical mechanism for establishing algal populations in frazil ice. *Nature*, *306*(5941), 363–365. https://doi.org/10.1038/306363a0

Gast, R. J., Moran, D. M., Beaudoin, D. J., Blythe, J. N., Dennett, M. R., & Caron, D. A. (2006). ABUNDANCE OF A NOVEL DINOFLAGELLATE PHYLOTYPE IN THE ROSS SEA, ANTARCTICA1. *Journal of Phycology*, *42*(1), 233–242. https://doi.org/10.1111/j.1529-8817.2006.00183.x

Gosink, J. J., & Staley, J. T. (1995). Biodiversity of gas vacuolate bacteria from antarctic sea ice and water. *Applied and Environmental Microbiology*, *61*(9), 3486–3489. https://doi.org/10.1128/aem.61.9.3486-3489.1995

Gosink, John J., Irgens, R. L., & Staley, J. T. (1993). Vertical distribution of bacteria in arctic sea ice. *FEMS Microbiology Letters*, *102*(2), 85–90. https://doi.org/10.1016/0378-1097(93)90003-K

Gowing, M., Garrison, D., Gibson, A., Krupp, J., Jeffries, M., & Fritsen, C. (2004). Bacterial and viral abundance in Ross Sea summer pack ice communities. *Marine Ecology Progress Series*, *279*, 3–12. https://doi.org/10.3354/meps279003

Gradinger, R., & Ikävalko, J. (1998). Organism incorporation into newly forming Arctic sea ice in the Greenland Sea. *Journal of Plankton Research*, *20*(5), 871–886. https://doi.org/10.1093/plankt/20.5.871

Gradinger, Rolf, & Zhang, Q. (1997). Vertical distribution of bacteria in Arctic sea ice from the Barents and Laptev Seas. *Polar Biology*, *17*(5), 448–454. https://doi.org/10.1007/s003000050139

Granskog, M. A., Vihma, T., Pirazzini, R., & Cheng, B. (2006). Superimposed ice formation and surface energy fluxes on sea ice during the spring melt-freeze period in the Baltic Sea. *Journal of Glaciology*, *52*(176), 119–127. https://doi.org/10.3189/172756506781828971

Grossart, H. P., Kiørboe, T., Tang, K., & Ploug, H. (2003). Bacterial colonization of particles: Growth and interactions. *Applied and Environmental Microbiology*, *69*(6), 3500–3509. https://doi.org/10.1128/AEM.69.6.3500-3509.2003

Grossmann, S. (1994). Bacterial activity in sea ice and open water of the Weddell Sea, Antarctica: A microautoradiographic study. *Microbial Ecology*, *28*(1), 1–18. https://doi.org/10.1007/BF00170244

Grossmann, S, & Dieckmann, G. S. (1994). Bacterial Standing Stock, Activity, and Carbon Production during Formation and Growth of Sea Ice in the Weddell Sea, Antarctica. *Applied and Environmental Microbiology*, *60*(8), 2746–2753. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16349347

Grossmann, Sönnke, & Gleitz, M. (1993). Microbial responses to experimental sea-ice formation: Implications for the establishment of Antarctic sea-ice communities. *Journal of Experimental Marine Biology and Ecology*, *173*(2), 273–289. https://doi.org/10.1016/0022-0981(93)90058-V

Guo, S., Stevens, C. A., Vance, T. D. R., Olijve, L. L. C., Graham, L. A., Campbell, R. L., … Davies, P. L. (2017). Structure of a 1.5-MDa adhesin that binds its Antarctic bacterium to diatoms and ice. *Science Advances*, *3*(8), e1701440. https://doi.org/10.1126/sciadv.1701440

Hallatschek, O., & Nelson, D. R. (2010). Life at the front of an expanding population. *Evolution*, *64*(1), 193–206. https://doi.org/10.1111/j.1558-5646.2009.00809.x

Hatam, I., Charchuk, R., Lange, B., Beckers, J., Haas, C., & Lanoil, B. (2014). Distinct bacterial assemblages reside at different depths in Arctic multiyear sea ice. *FEMS Microbiology Ecology*, *90*(1), 115–125. https://doi.org/10.1111/1574-6941.12377

Hatam, I., Lange, B., Beckers, J., Haas, C., & Lanoil, B. (2016). Bacterial communities from Arctic seasonal sea ice are more compositionally variable than those from multi-year sea ice. *ISME Journal*, *10*(10), 2543–2552. https://doi.org/10.1038/ismej.2016.4

Horikoshi, K., & Bull, A. T. (2011). Prologue: Definition, Categories, Distribution, Origin and Evolution, Pioneering Studies, and Emerging Fields of Extremophiles. In *Extremophiles Handbook* (pp. 3–15). https://doi.org/10.1007/978-4-431-53898-1_1

Huang, J., & Jaeglé, L. (2017). Wintertime enhancements of sea salt aerosol in polar regions consistent with a sea ice source from blowing snow. *Atmospheric Chemistry and Physics*, *17*(5), 3699–3712. https://doi.org/10.5194/acp-17-3699-2017

Hunke, E. C., Notz, D., Turner, A. K., & Vancoppenolle, M. (2011). The multiphase physics of sea ice: A review for model developers. *Cryosphere*, *5*(4), 989–1009. https://doi.org/10.5194/tc-5-989-2011

Janech, M. G., Krell, A., Mock, T., Kang, J. S., & Raymond, J. A. (2006). Ice-binding proteins from sea ice diatoms (Bacillariophyceae). *Journal of Phycology*, *42*(2), 410–416. https://doi.org/10.1111/j.1529-8817.2006.00208.x

Josenhans, C., & Schweinitzer, T. (2010, July). Bacterial energy taxis: A global strategy? *Archives of Microbiology*, Vol. 192, pp. 507–520. https://doi.org/10.1007/s00203-010-0575-7

Junge, K., & Swanson, B. D. (2008). High-resolution ice nucleation spectra of sea-ice bacteria:

implications for cloud formation and life in frozen environments. *Biogeosciences*, *5*(3), 865–873. https://doi.org/10.5194/bg-5-865-2008

Junge, Karen, Eicken, H., & Deming, J. W. (2003). Motility of Colwellia psychrerythraea strain 34H at subzero temperatures. *Applied and Environmental Microbiology*, *69*(7), 4282–4284. https://doi.org/10.1128/AEM.69.7.4282-4284.2003

Junge, Karen, Eicken, H., & Deming, J. W. (2004). Bacterial Activity at -2 to -20°C in Arctic Wintertime Sea Ice. *Applied and Environmental Microbiology*, *70*(1), 550–557. https://doi.org/10.1128/AEM.70.1.550-557.2004

Kaartokallio, H., Tuomainen, J., Kuosa, H., Kuparinen, J., Martikainen, P. J., & Servomaa, K. (2008). Succession of sea-ice bacterial communities in the Baltic Sea fast ice. *Polar Biology*, *31*(7), 783–793. https://doi.org/10.1007/s00300-008-0416-1

Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., … Schweder, T. (2014). Functional characterization of polysaccharide utilization loci in the marine Bacteroidetes “Gramella forsetii” KT0803. *ISME Journal*, *8*(7), 1492–1502. https://doi.org/10.1038/ismej.2014.4

Kaczmarska, I., Ehrman, J. M., Bates, S. S., Green, D. H., Léger, C., & Harris, J. (2005). Diversity and distribution of epibiotic bacteria on Pseudo-nitzschia multiseries (Bacillariophyceae) in culture, and comparison with those on diatoms in native seawater. *Harmful Algae*, *4*(4), 725–741. https://doi.org/10.1016/j.hal.2004.10.001

Kamagata, Y., & Narihiro, T. (2016). Symbiosis studies in microbial ecology. *Microbes and Environments*, Vol. 31, pp. 201–203. https://doi.org/10.1264/jsme2.ME3103rh

Kauko, H. M., Olsen, L. M., Duarte, P., Peeken, I., Granskog, M. A., Johnsen, G., … Assmy, P. (2018). Algal Colonization of Young Arctic Sea Ice in Spring. *Frontiers in Marine Science*, *5*(JUN), 199. https://doi.org/10.3389/fmars.2018.00199

Kennedy, M. J. (1987). Role of Motility, Chemotaxis, and Adhesion in Microbial Ecology. *Annals of the New York Academy of Sciences*, *506*(1), 260–273. https://doi.org/10.1111/j.1749-6632.1987.tb23825.x

Kerr, B., Neuhauser, C., Bohannan, B. J. M., & Dean, A. M. (2006). Local migration promotes competitive restraint in a host-pathogen "tragedy of the commons." *Nature*, *442*(7098), 75–78. https://doi.org/10.1038/nature04864

Kerr, B., Riley, M. A., Feldman, M. W., & Bohannan, B. J. M. (2002). Local dispersal promotes biodiversity in a real-life game of rock-paper-scissors. *Nature*, *418*(6894), 171–174. https://doi.org/10.1038/nature00823

Kiko, R. (2010). Acquisition of freeze protection in a sea-ice crustacean through horizontal gene transfer? *Polar Biology*, *33*(4), 543–556. https://doi.org/10.1007/s00300-009-0732-0

Kobori, H., Sullivan, C. W., & Shizuya, H. (1984). Bacterial plasmids in antarctic natural microbial assemblages. *Applied and Environmental Microbiology*, *48*(3), 515–518. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16346621

Koh, E. Y., Atamna-Ismaeel, N., Martin, A., Cowie, R. O. M., Beja, O., Davy, S. K., … Ryan, K. G. (2010). Proteorhodopsin-bearing bacteria in Antarctic sea ice. *Applied and Environmental Microbiology*, *76*(17), 5918–5925. https://doi.org/10.1128/AEM.00562-10

Kwok, R. (2018). Arctic sea ice thickness, volume, and multiyear ice coverage: Losses and coupled variability (1958-2018). *Environmental Research Letters*, *13*, 105005. https://doi.org/10.1088/1748-9326/aae3ec

Lindemann, F., Hölemann, J. A., Korablev, A., & Zachek, A. (1999). Particle Entrainment into Newly Forming Sea Ice—Freeze-Up Studies in October 1995. In *Land-Ocean Systems in the Siberian Arctic* (pp. 113–123). https://doi.org/10.1007/978-3-642-60134-7_12

Lindström, E. S., & Langenheder, S. (2012, February). Local and regional factors influencing bacterial community assembly. *Environmental Microbiology Reports*, Vol. 4, pp. 1–9. https://doi.org/10.1111/j.1758-2229.2011.00257.x

Lloyd, D. P., & Allen, R. J. (2015). Competition for space during bacterial colonization of a surface. *Journal of the Royal Society Interface*, *12*(110). https://doi.org/10.1098/rsif.2015.0608

Lorenz, M. G., Aardema, B. W., & Wackernagel, W. (1988). Highly efficient genetic transformation of Bacillus subtilis attached to sand grains. *Journal of General Microbiology*, *134*(1), 107–112. https://doi.org/10.1099/00221287-134-1-107

Luhtanen, A. M., Eronen-Rasimus, E., Kaartokallio, H., Rintala, J. M., Autio, R., & Roine, E. (2014). Isolation and characterization of phage-host systems from the Baltic Sea ice. *Extremophiles*, *18*(1), 121–130. https://doi.org/10.1007/s00792-013-0604-y

Luria, C. M., Amaral-Zettler, L. A., Ducklow, H. W., & Rich, J. J. (2016). Seasonal succession of free-living bacterial communities in coastal waters of the western antarctic peninsula. *Frontiers in Microbiology*, *7*(NOV), 1731. https://doi.org/10.3389/fmicb.2016.01731

Maas, E., Simpson, A., Martin, A., Thompson, S., Koh, E., Davy, S., … O'Toole, R. (2012). Phylogenetic analyses of bacteria in sea ice at Cape Hallett, Antarctica. *New Zealand Journal of Marine and Freshwater Research*, *46*(1), 3–12. https://doi.org/10.1080/00288330.2011.579981

Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M., & Larose, C. (2019). Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-38744-y

Maccario, L., Sanguino, L., Vogel, T. M., & Larose, C. (2015). Snow and ice ecosystems: Not so extreme. *Research in Microbiology*, *166*(10), 782–795. https://doi.org/10.1016/j.resmic.2015.09.002

Majaneva, M., Rintala, J. M., Piisilä, M., Fewer, D. P., & Blomster, J. (2012). Comparison of wintertime eukaryotic community from sea ice and open water in the Baltic Sea, based on sequencing of the 18S rRNA gene. *Polar Biology*, *35*(6), 875–889. https://doi.org/10.1007/s00300-011-1132-9

Maranger, R., Bird, D. F., & Kim Juniper, S. (1994). *Viral and bacterial dynamics in Arctic sea ice during the spring algal bloom near Resolute, N* (Vol. 111).

Meiners, K., Gradinger, R., Fehling, J., Civitarese, G., & Spindler, M. (2003). Vertical distribution of exopolymer particles in sea ice of the Fram Strait (Arctic) during autumn. *Marine Ecology Progress Series*, *248*, 1–13. https://doi.org/10.3354/meps248001

Methé, B. A., Nelson, K. E., Deming, J. W., Momen, B., Melamud, E., Zhang, X., … Fraser, C.

M. (2005). The psychrophilic lifestyle as revealed by the genome sequence of Colwellia psychrerythraea 34H through genomic and proteomic analyses. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(31), 10913–10918. https://doi.org/10.1073/pnas.0504766102

Milici, M., Deng, Z. L., Tomasch, J., Decelle, J., Wos-Oxley, M. L., Wang, H., … Wagner-Döbler, I. (2016). Co-occurrence analysis of microbial taxa in the Atlantic ocean reveals high connectivity in the free-living bacterioplankton. *Frontiers in Microbiology*, *7*(MAY), 649. https://doi.org/10.3389/fmicb.2016.00649

Møller, A. K., Barkay, T., Hansen, M. A., Norman, A., Hansen, L. H., Sørensen, S. J., … Kroer, N. (2014). Mercuric reductase genes (merA) and mercury resistance plasmids in High Arctic snow, freshwater and sea-ice brine. *FEMS Microbiology Ecology*, *87*(1), 52–63. https://doi.org/10.1111/1574-6941.12189

Mortazavi, R., Attiya, S., & Ariya, P. A. (2015). Arctic microbial and next-generation sequencing approach for bacteria in snow and frost flowers: selected identification, abundance and freezing nucleation. *Atmospheric Chemistry and Physics*, *15*(11), 6183–6204. https://doi.org/10.5194/acp-15-6183-2015

Nghiem, S. V., Rigor, I. G., Perovich, D. K., Clemente-Colón, P., Weatherly, J. W., & Neumann, G. (2007). Rapid reduction of Arctic perennial sea ice. *Geophysical Research Letters*, *34*(19). https://doi.org/10.1029/2007GL031138

Omstedt, A. (1985). Modelling frazil ice and grease ice formation in the upper layers of the ocean. *Cold Regions Science and Technology*, *11*(1), 87–98. https://doi.org/10.1016/0165-232X(85)90009-6

Paliwal, V., Puranik, S., & Purohit, H. J. (2012, February). Integrated perspective for effective bioremediation. *Applied Biochemistry and Biotechnology*, Vol. 166, pp. 903–924. https://doi.org/10.1007/s12010-011-9479-5

Pandey, J., Chauhan, A., & Jain, R. K. (2009, March 1). Integrative approaches for assessing the ecological sustainability of in situ bioremediation. *FEMS Microbiology Reviews*, Vol. 33, pp. 324–375. https://doi.org/10.1111/j.1574-6976.2008.00133.x

Passow, U. (2002). Transparent exopolymer particles (TEP) in aquatic environments. *Progress* *in Oceanography*, Vol. 55, pp. 287–333. https://doi.org/10.1016/S0079-6611(02)00138-6

Passow, U., & Alldredge, A. L. (1994). Distribution, size and bacterial colonization of transparent exopolymer particles (TEP) in the ocean. *Marine Ecology Progress Series*, *113*(1–2), 185–198. https://doi.org/10.3354/meps113185

Peterson, P. K., Hartwig, M., May, N. W., Schwartz, E., Rigor, I., Ermold, W., … Pratt, K. A. (2019). Snowpack measurements suggest role for multi-year sea ice regions in Arctic atmospheric bromine and chlorine chemistry. *Elem Sci Anth*, *7*(1), 14. https://doi.org/10.1525/elementa.352

Peura, S., Bertilsson, S., Jones, R. I., & Eiler, A. (2015). Resistant microbial cooccurrence patterns inferred by network topology. *Applied and Environmental Microbiology*, *81*(6), 2090–2097. https://doi.org/10.1128/AEM.03660-14

Raina, J. B., Fernandez, V., Lambert, B., Stocker, R., & Seymour, J. R. (2019, May 1). The role of microbial motility and chemotaxis in symbiosis. *Nature Reviews Microbiology*, Vol. 17, pp. 284–294. https://doi.org/10.1038/s41579-019-0182-9

Ramanan, R., Kim, B. H., Cho, D. H., Oh, H. M., & Kim, H. S. (2016, January 1). Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnology Advances*, Vol. 34, pp. 14–29. https://doi.org/10.1016/j.biotechadv.2015.12.003

Rapp, J. Z., Fernández-Méndez, M., Bienhold, C., & Boetius, A. (2018). Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Frontiers in Microbiology*, *9*(MAY), 1035. https://doi.org/10.3389/fmicb.2018.01035

Raymond, J. A., Fritsen, C., & Shen, K. (2007). An ice-binding protein from an Antarctic sea ice bacterium. *FEMS Microbiology Ecology*, *61*(2), 214–221. https://doi.org/10.1111/j.1574-6941.2007.00345.x

Riedel, A., Michel, C., Gosselin, M., & LeBlanc, B. (2007a). Enrichment of nutrients, exopolymeric substances and microorganisms in newly formed sea ice on the Mackenzie shelf. *Marine Ecology Progress Series*, *342*, 55–67. https://doi.org/10.3354/meps342055

Riedel, A., Michel, C., Gosselin, M., & LeBlanc, B. (2007b). Enrichment of nutrients, exopolymeric substances and microorganisms in newly formed sea ice on the Mackenzie shelf. *Marine Ecology Progress Series*, *342*, 55–67. https://doi.org/10.3354/meps342055

Rodrigues, D. F., & Tiedje, J. M. (2008, March 15). Coping with our cold planet. *Applied and Environmental Microbiology*, Vol. 74, pp. 1677–1686. https://doi.org/10.1128/AEM.02000-07

Rogers, G. B., Hoffman, L. R., Carroll, M. P., & Bruce, K. D. (2013, June). Interpreting infective microbiota: The importance of an ecological perspective. *Trends in Microbiology*, Vol. 21, pp. 271–276. https://doi.org/10.1016/j.tim.2013.03.004

Rysgaard, S., & Glud, R. N. (2004). Anaerobic N2 production in Arctic sea ice. *Limnology and Oceanography*, *49*(1), 86–94. https://doi.org/10.4319/lo.2004.49.1.0086

Rysgaard, S., Glud, R. N., Sejr, M. K., Blicher, M. E., & Stahl, H. J. (2008). Denitrification activity and oxygen dynamics in Arctic sea ice. *Polar Biology*, *31*(5), 527–537. https://doi.org/10.1007/s00300-007-0384-x

Schmidt, S. K., Costello, E. K., Nemergut, D. R., Cleveland, C. C., Reed, S. C., Weintraub, M. N., … Martin, A. M. (2007, June). Biogeochemical consequences of rapid microbial turnover and seasonal succession in soil. *Ecology*, Vol. 88, pp. 1379–1385. https://doi.org/10.1890/06-0164

Senčilo, A., Luhtanen, A. M., Saarijärvi, M., Bamford, D. H., & Roine, E. (2015). Cold-active bacteriophages from the Baltic Sea ice have diverse genomes and virus-host interactions. *Environmental Microbiology*, *17*(10), 3628–3641. https://doi.org/10.1111/1462-2920.12611

Showalter, G. M., & Deming, J. W. (2018). Low-temperature chemotaxis, halotaxis and chemohalotaxis by the psychrophilic marine bacterium Colwellia psychrerythraea 34H. *Environmental Microbiology Reports*, *10*(1), 92–101. https://doi.org/10.1111/1758-2229.12610

Smriga, S., Fernandez, V. I., Mitchell, J. G., & Stocker, R. (2016). Chemotaxis toward phytoplankton drives organic matter partitioning among marine bacteria. *Proceedings* *of the National Academy of Sciences of the United States of America*, *113*(6), 1576–1581. https://doi.org/10.1073/pnas.1512307113

Stecher, A., Neuhaus, S., Lange, B., Frickenhaus, S., Beszteri, B., Kroth, P. G., & Valentin, K. (2016). rRNA and rDNA based assessment of sea ice protist biodiversity from the central Arctic Ocean. *European Journal of Phycology*, *51*(1), 31–46. https://doi.org/10.1080/09670262.2015.1077395

Stewart, F. J., & Fritsen, C. H. (2004). Bacteria-algae relationships in Antarctic sea ice. *Antarctic Science*, *16*(2), 143–156. https://doi.org/10.1017/S0954102004001889

Stocker, R., & Seymour, J. R. (2012). *Ecology and Physics of Bacterial Chemotaxis in the Ocean*. https://doi.org/10.1128/MMBR.00029-12

Sturm, M., Holmgren, J., & Perovich, D. K. (2002). Winter snow cover on the sea ice of the Arctic Ocean at the Surface Heat Budget of the Arctic Ocean (SHEBA): Temporal evolution and spatial variability. *Journal of Geophysical Research C: Oceans*, *107*(10), SHE 23-1. https://doi.org/10.1029/2000jc000400

Techtmann, S. M., Fitzgerald, K. S., Stelling, S. C., Joyner, D. C., Uttukar, S. M., Harris, A. P., … Hazen, T. C. (2016). Colwellia psychrerythraea Strains from Distant Deep Sea Basins Show Adaptation to Local Conditions. *Frontiers in Environmental Science*, *4*(MAY), 33. https://doi.org/10.3389/fenvs.2016.00033

Thomas, D. N., & Dieckmann, G. S. (2002, January 25). Antarctic sea ice - A habitat for extremophiles. *Science*, Vol. 295, pp. 641–644. https://doi.org/10.1126/science.1063391

Torstensson, A., Dinasquet, J., Chierici, M., Fransson, A., Riemann, L., & Wulff, A. (2015). Physicochemical control of bacterial and protist community composition and diversity in Antarctic sea ice. *Environmental Microbiology*, *17*(10), 3869–3881. https://doi.org/10.1111/1462-2920.12865

Underwood, G. J. C., Paterson, D. M., & Parkes, R. J. (1995). The measurement of microbial carbohydrate exopolymers from intertidal sediments. *Limnology and Oceanography*, *40*(7), 1243–1253. https://doi.org/10.4319/lo.1995.40.7.1243

Underwood, Graham J.C., Aslam, S. N., Michel, C., Niemi, A., Norman, L., Meiners, K. M., … Thoma, D. N. (2013). Broad-scale predictability of carbohydrates and exopolymers in Antarctic and Arctic sea ice. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39), 15734–15739. https://doi.org/10.1073/pnas.1302870110

Vayssier-Taussat, M., Albina, E., Citti, C., Cosson, J.-F., Jacques, M.-A., Lebrun, M.-H., … Candresse, T. (2014). Shifting the paradigm from pathogens to pathobiome: new concepts in the light of meta-omics. *Frontiers in Cellular and Infection Microbiology*, *4*(MAR), 29. https://doi.org/10.3389/fcimb.2014.00029

Venegas-Ortiz, J., Allen, R. J., & Evans, M. R. (2014). Speed of invasion of an expanding population by a horizontally transmitted trait. *Genetics*, *196*(2), 497–507. https://doi.org/10.1534/genetics.113.158642

Vollmers, J., Voget, S., Dietrich, S., Gollnow, K., Smits, M., Meyer, K., … Daniel, R. (2013). Poles apart: Arctic and Antarctic Octadecabacter strains share high genome plasticity and a new type of xanthorhodopsin. *PloS One*, *8*(5), e63422. https://doi.org/10.1371/journal.pone.0063422

Vonnahme, T. R., Dietrich, U., & Hassett, B. T. (2020). Progress in Microbial Ecology in Ice-Covered Seas. In *YOUMARES 9 - The Oceans: Our Research, Our Future* (pp. 261–277). https://doi.org/10.1007/978-3-030-20389-4_14

Walsby, A E. (1972). *Structure and Function of Gas Vacuoles*. Retrieved from http://mmbr.asm.org/

Walsby, Anthony E. (1981). Gas-Vacuolate Bacteria (Apart from Cyanobacteria). In *The Prokaryotes* (pp. 441–447). https://doi.org/10.1007/978-3-662-13187-9_28

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Bryazgin, N. N., Aleksandrov, Y. I., … Colony, R. (1999). Snow Depth on Arctic Sea Ice. *Http://Dx.Doi.Org/10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2*. https://doi.org/10.1175/1520-0442(1999)012<1814:SDOASI>2.0.CO;2

Wein, T., Romero Picazo, D., Blow, F., Woehle, C., Jami, E., Reusch, T. B. H., … Dagan, T. (2019, October 1). Currency, Exchange, and Inheritance in the Evolution of Symbiosis.

Trends in Microbiology, Vol. 27, pp. 836–849. https://doi.org/10.1016/j.tim.2019.05.010

Weissenberger, J., & Grossmann, S. (1998). Experimental formation of sea ice: Importance of water circulation and wave action for incorporation of phytoplankton and bacteria. *Polar Biology*, *20*(3), 178–188. https://doi.org/10.1007/s003000050294

Wells, L. E., & Deming, J. W. (2006). Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environmental Microbiology*, *8*(6), 1115–1121. https://doi.org/10.1111/j.1462-2920.2006.00984.x

Wilkins, L. G. E., Ettinger, C. L., Jospin, G., & Eisen, J. A. (2019). Metagenome-assembled genomes provide new insight into the microbial diversity of two thermal pools in Kamchatka, Russia. *Scientific Reports*, *9*(1), 1–15. https://doi.org/10.1038/s41598-019-39576-6

Williams, T. J., Wilkins, D., Long, E., Evans, F., Demaere, M. Z., Raftery, M. J., & Cavicchioli, R. (2013). The role of planktonic Flavobacteria in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics. *Environmental Microbiology*, *15*(5), 1302–1317. https://doi.org/10.1111/1462-2920.12017

Wu, Z. yan, Lin, W. xiong, Li, J. juan, Liu, J. fu, Li, B. lian, Wu, L. kun, … Zhang, Z. xing. (2016). Effects of seasonal variations on soil microbial community composition of two typical zonal vegetation types in the Wuyi Mountains. *Journal of Mountain Science*, *13*(6), 1056–1065. https://doi.org/10.1007/s11629-015-3599-2

Yang, X., Pyle, J. A., & Cox, R. A. (2008). Sea salt aerosol production and bromine release: Role of snow on sea ice. *Geophysical Research Letters*, *35*(16). https://doi.org/10.1029/2008GL034536

Yergeau, E., Michel, C., Tremblay, J., Niemi, A., King, T. L., Wyglinski, J., … Greer, C. W. (2017). Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Scientific Reports*, *7*(1), 42242. https://doi.org/10.1038/srep42242

Yu, Z. C., Chen, X. L., Shen, Q. T., Zhao, D. L., Tang, B. L., Su, H. N., … Zhang, Y. Z. (2015). Filamentous phages prevalent in Pseudoalteromonas spp. Confer properties advantageous to host survival in Arctic sea ice. *ISME Journal*, *9*(4), 871–881. https://doi.org/10.1038/ismej.2014.185

Zhou, J., & Ning, D. (2017). Stochastic Community Assembly: Does It Matter in Microbial Ecology? *Microbiology and Molecular Biology Reviews*, *81*(4). https://doi.org/10.1128/mmbr.00002-17

Chapter 2: Uncovering the driving forces of ice and frost flower microbial communities in a controlled sea-ice chamber

Layton R, France J, Kaiser J, Vogel T.M and Larose C

2.0 Introduction

Sea ice is expansive, covers over 16 million km^2, which is approximately 5% of the world's ocean surface (Stroeve and Notz, 2018). Yet, with changes in climatic conditions, a general decline in sea ice cover and a transition from multiyear sea ice (MYI) to new and young sea ice is occurring more frequently. First year ice (FYI), which forms and melts annually, now dominates the ice cover and represented nearly 80% of the March 2018 ice pack (D. Perovich et al., 2018). The predicted increase in the spatial extent of FYI gives rise to the potential for frost flower fields to form. These are delicate dendritic structures that form almost exclusively on FYI if conditions permit (Style and Worster, 2009). Frost flowers have recently been identified as unique microbial habitats capable of supporting large bacterial densities despite their transiency and high salinity values (>120 ppt) (Style and Worster, 2009; Bowman and Deming, 2010; Bowman et al., 2013). Not only does abundance of bacteria in natural frost flowers increase over their lifetime, even compared to the underlying sea ice (3-6 fold higher) but a taxa-specific enrichment is observed (Bowman et al., 2013, 2014). While frost flower communities have the potential to participate in biogeochemical cycling, only one functional study on a singular sample is available to date (Bowman *et al.*, 2014). The enrichment of these microorganisms might be an important mechanism in maintaining their population numbers in oceans. Therefore, understanding why (i.e. what other than salinity?) and what organisms preferentially colonise sea ice during the transition from open water to the early stages of sea ice will govern the distribution, composition and development of the sea ice community.

During formation, a plethora of particles, compounds, microorganisms, and viruses can be entrained and entrapped in the sea ice. This is a size selective process where larger particles and cells are preferentially incorporated into the ice (Gradinger and Ikävalko, 1998; Rózańska, Poulin and Gosselin, 2008). A number of studies have suggested that frazil ice that forms in the underlying waters during the initial stages of sea ice formation acts to scavenge planktonic algae (Garrison, Close and Reimnitz, 1989; Grossmann and Gleitz, 1993). In turn, bacterial enrichment in sea ice is explained by their association with algae or other larger particles (Grossmann and Gleitz, 1993). However, sea ice recruitment studies have concentrated on eukaryotes, while only hinting at the potential mechanisms by which bacteria are enriched in sea ice. Moreover, open water and pancake ice are not discernible from the underlying seawater in terms of their bacterial community composition prior to consolidation of sea ice (Eronen-Raismus *et al.*, 2015). Thus, while frazil ice may act to scavenge algae, the selection of a sea ice specific bacterial community seems to occur at a later stage of sea ice formation.

A number of alternative hypotheses exist to explain the occurrence of a specific community in young, newly formed sea ice. Some microorganisms may actively move toward newly formed ice or increase their chances of incorporation during formation through strategies such as ice nucleation. For example, the archetypal sea ice microorganism *Colwellia psychrerythraea* strain 34H was shown to be motile at temperatures as low as −10°C (Junge, Eicken and Deming, 2003). Once incorporated in the sea ice, brine expulsion back into the seawater or on to the sea ice surface may affect the bacterial community in the ice and frost flowers. For example, microorganisms with retention strategies such as EPS production, leading to 'EPS plugs', or the ability to attach to ice crystals can predominate (Ewert and Deming, 2013). Supporting microscopic work of intact sea ice brine channels estimates that >50% of bacteria in sea ice are attached to surfaces (Junge, Eicken and Deming, 2004). Following recruitment and retention, selective prevailing conditions such as increased salinity, exposure to reactive oxygen species (ROS), increased viral stress and increased competition may select for unique innate adaptations such as resistance to osmotic stress, viral infection and antibiotic production. The reduced habitable space presented by the brine environment might increases cell-cell contact and in turn the potential for horizontal gene transfer.

As algae are scavenged by frazil ice crystals during sea ice formation, we would expect to find a larger number of algae associated with the solid phase of sea ice (*i.e*. the sea ice matrix). Taxa and functions related to bacterial and algal symbiosis would be likely to be enriched in solid phase sea ice communities. However, given the lack of this selection in bacterial community pre-consolidation, we suggest that colonisation and cell enrichment in sea ice brines is not a product of algal association, but is a function of post-formation selection mechanisms. Bacteria that make up the distinctive sea ice communities are uniquely adapted to actively seek out nutrient-rich environments, mitigate brine expulsion and/or resist physio-chemical conditions such as osmotic stress. To test these hypotheses, we performed a sea-ice-chamber-based experiment that circumvents some of the complexity inherent in natural microbial assemblages that can obscure ecological patterns. We expected that natural sea ice growth and development, and associated microbial ecology, can be reliably replicated in a laboratory setting. Moreover, we anticipated that genomic signatures of transport, retention and adaptation in the frost flower and brine environments would be reflected in their respective metagenomes. These metagenomics datasets were created by exploiting both amplicon and shotgun metagenomic sequencing techniques.

2.1 Methods

2.1.1 Sea Ice Chamber

The sea-ice facility consists of a 3,500 litre capacity glass tank, housed in an environmental chamber that is capable of temperatures between -55°C and +30°C. For this experiment, the tank was filled with 3,000 litres of un-filtered sea water taken from approximately 24 km (15 miles) offshore in the English Channel. The tank was instrumented with a range of instrumentation (Table S2.1; Supplementary Materials) designed to monitor the ice growth, in-ice conditions, under-ice ocean conditions and above ice meteorological conditions for the duration of the experiment. The data from the instrumentation is fed back to logging devices in real-time to monitor the experiment (and controlled changes if needed) to ensure the desired ice-forming conditions occur. A full schematic set up of the experiment is shown in **Figure 2.1**.

2.1.2 Growth of Sea Ice and Frost Flowers

As described in Style & Worster (2009) the conditions for forming frost flowers are expected to occur with a difference in air temperature and ice surface temperature of <-10°C for a

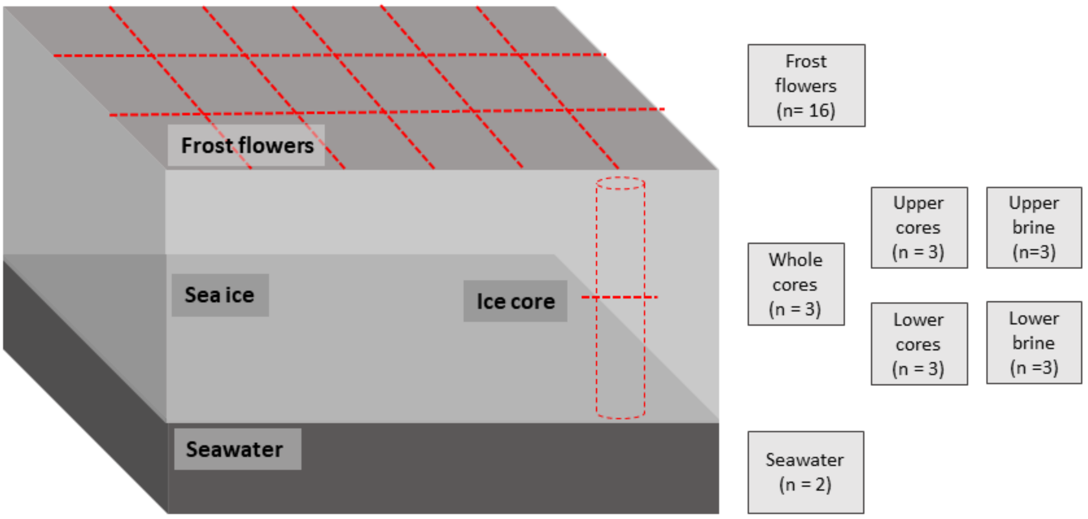


Figure 2.1: Sampling schematic of sea ice chamber

relative humidity of 60% (the average humidity during the experiment) with wind speeds below 5 m/s. Therefore, the ice growth conditions for the majority of the experiment were set at -26°C (lower temperatures were attempted, but resulted in the slight super-cooling of the water column) with an average wind speed of 0.3 m/s above the tank. By the end of the experiment, the ice surface temperature had reached -12°C, so the frost-flower-forming conditions were maintained at all times.

Air could only exchange between the laboratory and chamber through small ports designed to prevent pressure differences building up between the laboratory and chamber. The base of the tank was heated and insulated to maintain the ocean at temperatures above freezing. Sea-ice only grows due to cooling at the surface of the ocean, and thus replicates the ice-forming conditions seen in the natural world. Ice growth rates throughout the experiment are shown in Figure S2.1 (Supplementary Materials). Given the air temperatures, expected thermodynamic growth rates in a natural environment above the sea-ice would be ~80 mm/day (Lebedev, 1938). The slower growth rate in the tank was probably due to the extra energy being put into the system through the artificial heating to maintain the ocean temperature and the gradual increase in salinity throughout the experiment.

During the experiment, the sea-ice was periodically illuminated by visible light for approximately 30 minutes every 12 hours from LED lighting covering the spectral range from 400nm to 750nm to allow measurements of PAR (Photosynthetically Active Radiation) transmission through the ice. The periodic lighting of the chamber was also used for monitoring the frost flower growth, which is shown in Figure S2.2 (Supplementary Materials). The end of the experiment was defined when the ice thickness reached ~20 cm with frost flowers covering approximately 90% of the surface.

2.1.3 Sampling

Twelve hours before the experiment concluded, atmospheric samples were collected for microbial community analysis using a KNF LABOPORT N 86 KT.18 electric mini vacuum pump with a flow rate of 5.5 L/min. A total volume of 3.63 m^3 was collected on a 0.2µM sterile Millipore filter. No significant DNA was detected.

Frost flowers were subsequently harvested by dividing the chamber into 18 replicates (3x6 grid). Frost flowers from each section were collected using a sterile scraper, transferred to a sterile sampling bag and packaged for transport. The sea-ice was then cored using an ethanol-rinsed Kovacs corer a total of 18 times (once per replicate), only sampling undisturbed sea-ice away from the heavily instrumented parts of the tank. Four cores were kept whole and stored immediately and the remaining twelve were split length ways into two equal sections, representing the top and bottom layers of sea ice. The top and bottom segments were independently collated into six replicates (**Figure 2.1**). In order to separate the ice from the cores into discrete layers, a slow melt process at 4°C was employed to section the ice cores into bins based on salinity. This concept has been employed previously by Fripiat *et al.* (2007) and separates the ice by melting point with most of the saline brine drawn off first and the bulk pure ice last. This procedure accesses the habitat niches of high salinity brine channels without having to directly sample them in-situ.

Finally, 1 L of sea water was harvested before and after the experiment. The former was stored at 4°C for the duration of the experiment while the post-experiment sample was immediately packaged for transport. All samples were transported in well insulated material. However, the samples were allowed to melt during transportation over a period of 24-48 hours. Once completely melted, samples were immediately processed.

2.1.4 Basic Chemical Analyses and Filtration

All samples were subsampled independently for pH and salinity measurements, with the exception of frost flowers. Due to the limited volume achieved from frost flower melt, 10%

of each frost flower sample was sub-sampled into 3 independent replicates. The remaining sample was filtered onto a 0.2 µM filter using a vacuum filtration system. The filters were folded and placed in sterile micro centrifuge tubes, which were stored frozen (-20°C) and in the dark until DNA extraction.

2.1.5 DNA Extraction, Sequencing and qPCR

DNA was extracted from filters using the DNeasy PowerWater Kit (Qiagen) following the manufacturer’s instructions. Subsequent to DNA quantification using the QubitTM dsDNA HS Assay Kit (Thermo Fisher Scientific), the V3–V4 regions of the 16S rRNA gene were amplified using the universal primers Bac_341F and Bac_805R modified with the Illumina overhang adapter sequences. All primers are listed in Table S2.2 (Supplementary Materials). Similarly, the Euk-528F and Euk-706R primers tagged with the same Illumina adapters were used to amplify the V4-V5 region of the 18S rRNA gene. Both PCR’s were performed using the Platinum PCR SuperMix (Invitrogen) and consisted of the following cycles and conditions: 35 cycles at 92°C 30 s, 55°C 30 s and 72°C. The 16S and 18S rRNA gene sequencing libraries were prepared using the 16S rRNA gene Library Preparation Workflow recommended by Illumina. Samples with sufficient remaining DNA concentrations were subject to metagenomic sequencing. Metagenomic libraries were prepared using the Nextera XT Library Prep Kit and Workflow (Illumina) according to manufacturer’s instructions. All sequencing libraries were subject to paired end sequencing on a MiSeq sequencer (Illumina) at our laboratory in Lyon.

16S and 18S rRNA gene copy numbers were estimated by qPCR using the primer sets listed in Table S2.2 (Supplementary Materials). Standards for 16S and 18S rRNA gene quantification were obtained from PCR products amplified from *E.coli* or *C. potsdamis* respectively. Following the cloning of the 16S and 18S rRNA sequence fragments in a plasmid, DNA quantification was performed with the Broad-Range Qubit Fluorometric Quantification (Thermo Fisher Scientific) and standards diluted to the applicable qPCR concentrations. qPCR was carried out using the SensiFast SYBR No-Rox kit (Bioline) on a Rotorgene 3000 machine (Qiagen). The reaction mixture contained 10 µL of SYBR master mix, 2 µL of DNA and RNAse-free water to a final volume of 20 µL. The 2-step qPCR program consisted of; enzyme activation (95°C; 2 min), then 35 cycles of 5 s at 95°C and 20 s at 60°C for hybridization and elongation respectively. Subsequently, melt curve analyses were performed at of 1°C increments between 55 to 95°C. Gene copy numbers were calculated from the standard curves of 16S and 18S rRNA genes by using the 1Ct (cycle threshold) method.

2.1.6 Dataset Generation – Amplicon Data

Following primer removal from raw sequences (cutadapt v2.8; Martin, 2011), sequencing data was subject to DADA2 analysis processed following the DADA2 (version 1.12) pipeline (Callahan *et al.*, 2016). Briefly, sequence read quality profiles were inspected, and sequences filtered with the filterAndTrim function. Based on their quality scores, sequences were truncated (16S: 231 and 227 bp; 18S: 220 and 200 bp) and the maximum error rate set to 3 and 5 for forward and reverse sequence reads, respectively. The resulting reads were denoised subject to an error model learned from the data and paired ends were merged. Chimeras were removed using the “consensus” method. The 16S Ribosomal Database Project’s (trainset 16) and 18S Silva (v132) databases were used for taxonomy assignment by RDP’s naive Bayesian classifier within DADA2.

2.1.7 Amplicon Analyses

Both 16S and 18S rRNA gene annotated datasets were imported into R (R Development Core Team, 2011) and manipulated using the R package Phyloseq (McMurdie and Holmes, 2013). Samples with less than 1000 reads and any singletons within the remaining samples were removed before proceeding with further analyses. Using Phyloseq and ggplot2 (Wickham, 2016) for figure generation, we computed the relative abundances of taxa and plotted stacked histograms. We calculated the Bray–Curtis dissimilarity distance and visualized the

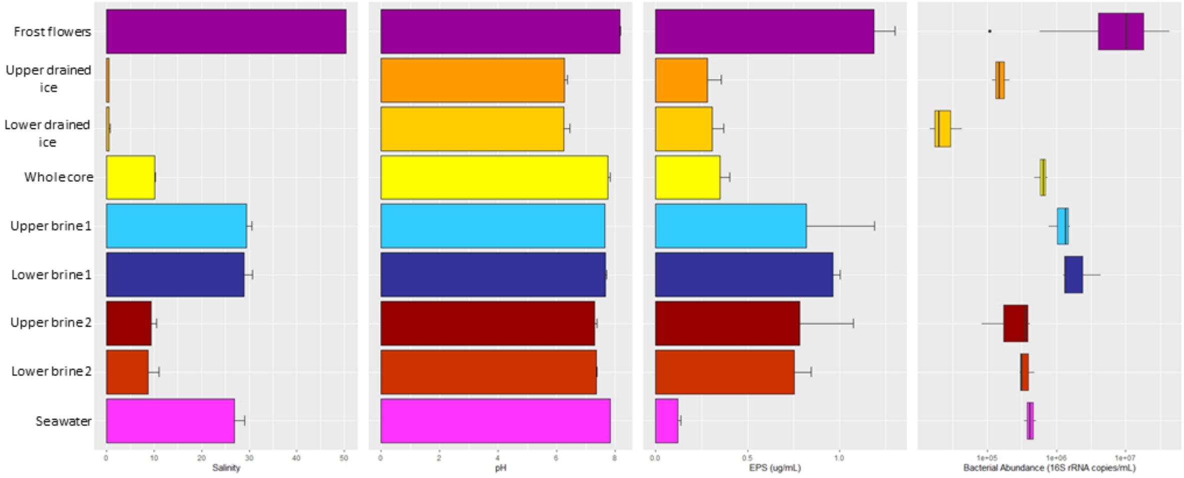


Figure 2.2: Sample characteristics (A) Salinity (pH) (C) pEPS and (D) Bacterial abundance as estimate by qPCR of the 16S rRNA gene

ordination using non-metric multi-dimensional scaling (NMDS). As beta dispersion was shown to be significant between samples (P<0.001) we used ANOSIM to test for differences between the sampling groups with 1000 permutations. To gain insight into whether our measured environmental variables affect the community composition, we performed a constrained ordination on the Bray-Curtis Distances using the functions “capscale” from the Vegan package followed by an ANOVA (1000 permutations) to test for statistical significance. Finally, to identify over-represented taxa between seed and source environments, differential abundance testing was performed using the DeSeq2 package with a p-value < 0.05 (Love, Huber and Anders, 2014).

Prior to correlation analyses, a filtering step was used to remove rare amplicon sequence variants (ASVs) (< 2% relative abundance) present in less than 4 samples from 16S rRNA and 18S rRNA gene ASV tables. Spearman rank correlations and p-values between bacterial and eukaryotic ASVs were computed by concatenating normalised amplicon ASV tables and applying the function rcorr() in the Hmisc R package (Harrell Jr, 2019). P-values were adjusted using the Benjamini and Hochberg (1995) method. Significant correlations (padj < 0.001) were filtered and plotted on scatter graphs resulting in the identification of two major relationship types; linear and logarithmic (data not shown). As such, linear regression curves were fitted to untransformed and log-transformed data and filtered using a p-value and r-squared cutoff <0.001 and >0.5 respectively. Finally, all 18S rRNA ASVs not annotated as microalgae were removed.

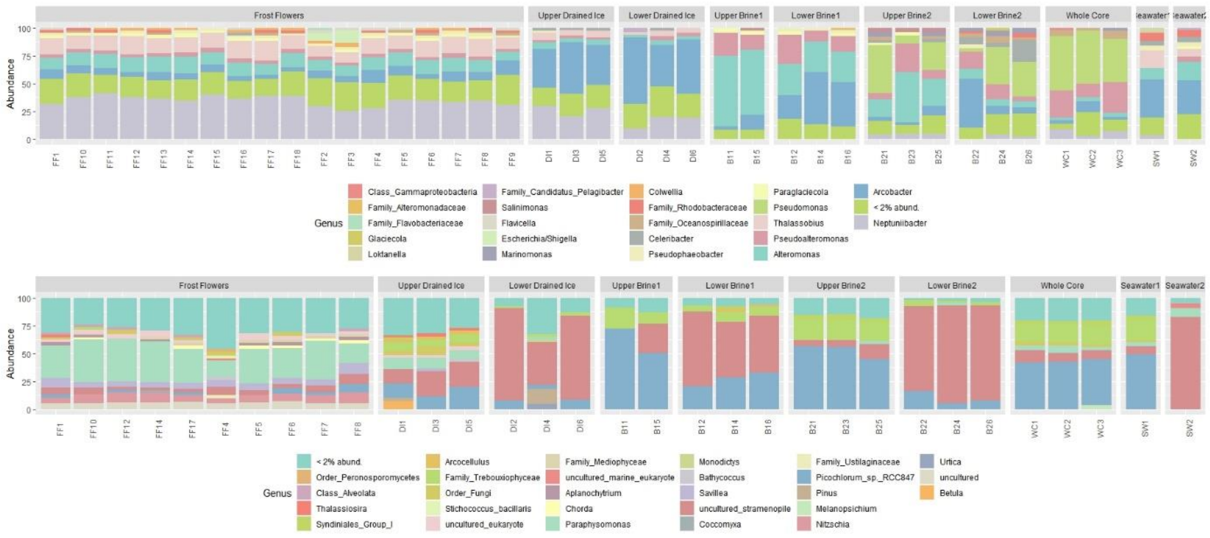


Figure 2.3: Histograms of the relative abundance of the (A) abundant taxa in the 16S rRNA dataset and (B) all taxa in the 18S rRNA dataset at genus level.

2.1.8 Shotgun Metagenomes

FastQC (Andrews *et al.*, 2012) was used to check read quality profiles and Trimmomatic was applied to remove Illumina adapters and dynamically trim and filter sequences based on their quality scores (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36). The resulting sequence files were functionally annotated using EggNOG-Mapper (Huerta-Cepas *et al.*, 2017), based on eggNOG orthology data (Huerta-Cepas *et al.*, 2016), using the default parameters. Following importation into R, the gene count and functional annotation tables were constructed. Uniprot protein names were added to the annotation tables by utilising the “Retrieve/ID mapping” online uniprot function to convert EggNOG string IDs into uniprot protein names. Taxonomic IDs from eggnogmapper were converted to taxonomic descriptions using the Rpackage ‘taxize’(Scott Chamberlain and Eduard Szocs, 2013; Chamberlain *et al.*, 2019).

Gene count and functional annotation tables were further manipulated in R. Bray-Curtis distances on the relative abundance of KEGG numbers were computed and subject to hierarchal clustering to produce dendrograms based on the average distance method. GO-terms of specific interest were mined using the gene count and functional annotation table and tested for differential abundance among sample types using DeSeq2 package (Love, Huber and Anders, 2014).

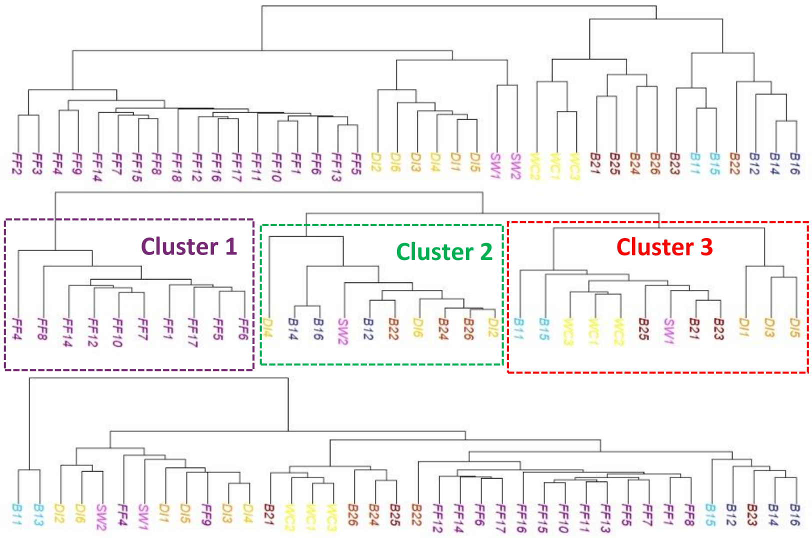


Figure 2.4: Average-linkage dendrograms representing hierarchal clustering of Bray-Curtis distances computed from the relative abundance of (A) 16S ASVs and (B) 18S ASVs and (C) functions (KEGG catergories). In (B) Cluster 1 represents the frost flower samples, Cluster 2 the lower sea ice and post experiment seawater samples and Cluster 3 the upper sea ice and pre-seawater samples

2.2 Results

2.2.1 Sample Characteristics

Salinity corresponded to freshwater (0.5 PSU) in the drained ice sections after complete brine drainage (**Figure 2.2**). On the other side of the spectrum, salinity values peaked in the frost flowers (~50 PSU), followed by the primary brine fractions (UB1 and LB1) at approximately 29 PSU. There was only a small difference between the salinity values of these brine fractions and the seawater (~27 PSU) that could be indicative of over-drainage.

EPS concentrations were approximately ten times higher in the frost flower samples relative seawater (**Figure 2.2**). Within the sea ice, EPS concentrations were high in the liquid phase (1.55 ug/mL) and low (0.11 ug/mL) in the solid phase despite high variability in the primary brine fractions. The frost flowers contained approximately 1.3 x 10^7 16S rRNA gene copies ml^{-1} of melted sample. Mean gene copy counts were roughly an order of magnitude lower in the primary brine fractions (~1.8 x 10^6 16S rRNA gene copies ml^{-1}). In turn, primary brine fractions had approximately five and a half times higher gene counts than the seawater and secondary brine fractions. The lowest counts corresponded to the lower portions of the drained ice.

2.2.2 Bacterial Community Structure and Composition

Post filtering, we classified 2848 taxa across all samples (Tables S2.3A). At class level, the seawater contained a roughly even split of Alpha- (34%), Gamma- (29%) and Epsilon- (33%)

proteobacteria (**Figure S2.3**; Supplementary Materials). Conversely, both the first and second brine fractions, the frost flowers and the whole core all contained over double the abundance of Gammaproteobacteria relative to the seawater at 69%, 75%, 63% and 90% respectively. The largest class representation in the drained ice matrix samples was by the Epsilonproteobacteria (45%) followed by Gammaproteobacteria (37%) (**Figure S2.3**; **Supplementary Materials**).

At the genus level, there are clear differences in the abundances of taxa between samples (**Figure 2.3**). Sequences annotated as *Neptuniibacter* constituted the largest fraction of frost flower samples (35%) and were significantly more abundant in frost flowers relative to all other environments with the exception of the upper drained ice (padj<0.05; **Tables S2.4; Supplementary Materials**). UB1 samples were overwhelmingly dominated by *Alteromonas* (61%), which decreased in relative abundance with decreasing height through the profile. *Pseudomonas* and *Psuedoalteromonas* were prominent taxa in the second brine fractions and the whole cores, yet virtually non-existent in the seawater. Finally, the genus with the highest relative abundance in the seawater was *Arcobacter* (32%), which was also found in high abundance in both the upper (39%) and lower (49%) drained ice samples.

Average-linkage dendrograms representing hierarchal clustering of Bray-Curtis distances computed from the relative abundance of (A) 16S ASVs and (B) 18S ASVs and (C) functions (KEGG catergories). In (B) Cluster 1 represents the frost flower samples, Cluster 2 the lower sea ice and post experiment seawater samples and Cluster 3 the upper sea ice and pre-seawater samples

There were instances where candidate sea ice g enera went undetected in the seawater samples but were detected in one or more of the sea ice profile environments. Examples of this include *Halomonas, Lewinella, Maribacter, Methylophaga, Neptunomonas, Nonlabens, Psychrobacter, Pseudorhodobacter, Psychrobacter, Psychromonas, Rhodococcus, Shewanella, Sphingorhabdus* and *Winogradskyella*. There were also cases where archetypal sea ice bacteria were detectable in the seawater samples but increased significantly in abundance in one or more of the sea ice profile environments. For instance, *Glaciecola* (SW vs LB2 and WC). *Polaribacter* (LDI, WC), *Colewellia* (FF), *Pseudomonas* (UB2, LB2) and *Psudoalteromonas* (all fractions except drained ice) (padj<0.05; **Tables S2.4**).

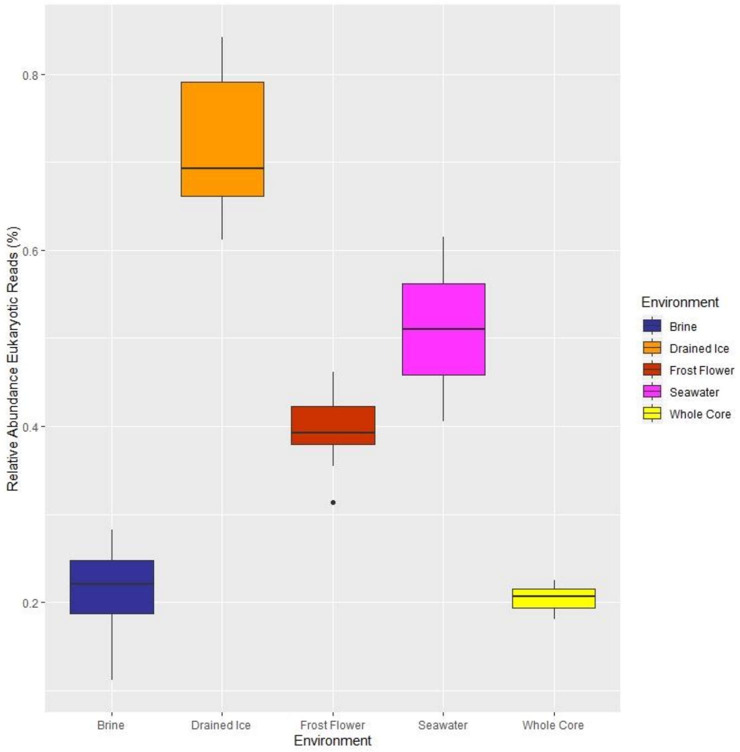


Figure 3.4: Box plots of the percentage of reads assigned as eukaryotic in origin. Differences between environmental niches was significant (p<0.001) as tested by a one-way anova.

An NMDS analyses based on Bray-Curtis Distances of the relative abundances of ASVs (**Figure S2.4; Supplementary Materials**) shows strong habitat segregation. Hierarchal clustering of samples by these distance measures and average clustering (**Figure 2.4**) demonstrate two major groups of samples; (1) the ice dominated environments (the frost flowers and the drained ice matrix) and seawater and (2) the brine fractions and whole core samples.

2.2.3 Eukaryotic Community Structure and Composition

The metagenomic reads classified as eukaryotic in origin composed <1% relative abundance in all samples. However, there were significant differences (p = <2e-16) in the relative abundance of eukaryotic reads between sample types (**Figure 2.5**). The drained ice samples had the greatest proportion of eukaryotic reads followed by the seawater and frost flower samples. Both the brine and whole core samples had the least amount of eukaryotic read assignments.

Over 1500 taxa were classified after the 18S amplicon sequencing data post-filtering (Table S2.3B). Twenty taxonomic groups were resolved to the class level but only 8 of these represented >1% abundance in any sample. The Stramenopiles and Chloroplastida contributed to between 68% and 99% of the total abundance of any sample (**Figure S2.3**). There was a large difference between the class representation in seawater samples pre and post experiment. Although 16 classes were found in the pre-experiment, only 8 of these were detected post experiment seawater. Moreover, while Stramenopiles were moderately abundant in SW1, they composed almost the entirety of SW2 (99%).

While the bacterial communities of the brines showed the strongest segregation by fraction, the eukaryotic communities were vertically segregated **(Figure 2.3)**. This was largely due to either the predominance of Picochlorum or Stramenopiles in upper or lower brine fractions respectively. Hierarchal clustering by the average method produced three distinct groups where the upper drained ice and brine fractions clustered with SW1 and the lower drained ice and brine fractions clustered around SW2 (**Figure 2.4**). Finally, the frost flowers formed a distinct group relative to all other samples in which *Paraphysomonas*, *Nitzschia* and *Savillea* were the only genera to constitute >5% abundance and were significantly enriched relative to the brine **(Figure 2.3)**.

2.2.4 Correlations

Filtering Spearman's rank correlation between 18S rRNA and 16S rRNA ASVs by p-value (padj<0.001) resulted in 304 putative relationships. Further filtering following regression analyses and removal of non-microalgal species resulted in 33 putative correlations. Only 11 bacterial ASVs annotated as 5 unique genera (*Thalassobius*, *Neptuniibacter*, *Celeribacter*, *Pseudoalteromonas*, *Alteromonas* and *Flavicella*) contributed to the total number of co-occurrences (**Table S2.5; Supplementary Materials**). Five of these ASVs were annotated as the genus *Neptuniibacter*. Five 18S rRNA ASVs contributed to the total number of putative co-occurrences between bacteria and microalgae; only 3 could be classified at the genus level (*Thalassiosira*, *Picochlorum* and *Nitzschia*) and two at Family level (Trebouxiophyceae and Mediophyceae). Signficant positive correlations were identified for multiple ASVs annotated as *Neptuniibacter*, *Nitzschia* (euk_9), *Thalassiosira* (euk_16), and Family Mediophyceae. euk_9 (Nitzschia) and euk_22 (Family Mediophyceae) also significantly and positively correlated with *Thalassobius*, *Flavicella* and *Alteromonas*.

2.2.5 Functional Potential of Sea Ice Communities

Grouping of samples based on hierarchal clustering of the relative abundance of functions produced five major groups (**Figure 2.4).** The first was only comprised of two samples; both primary upper brine fractions. Clustering by NMDS (**Figure S2.4**) demonstrated that the final sample clustered near this group. The second cluster was composed of the upper and lower drained ice and the seawater metagenomes (and 2 of 15 frost flower metagenomes; **Figure 2.4)**. All whole core and four of the six secondary brine fractions constituted cluster three

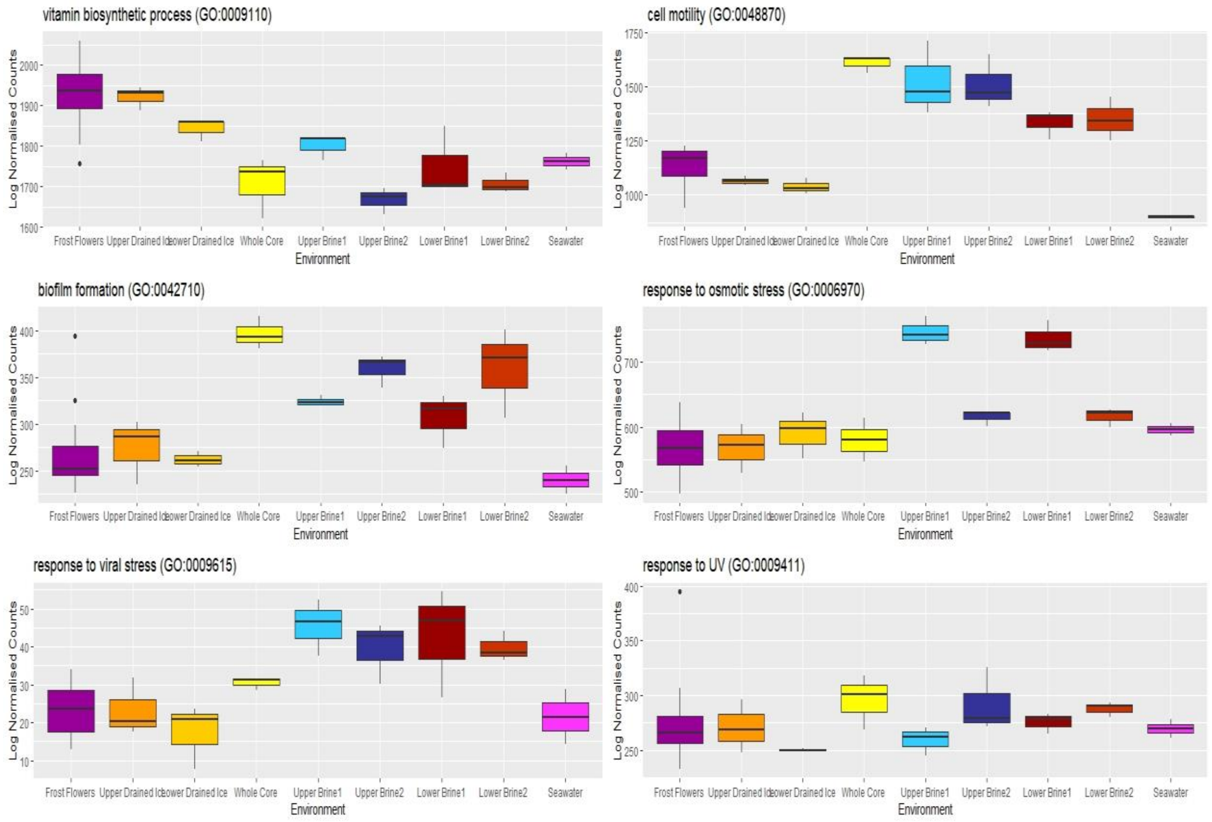


Figure 3.5: Abundances of functional gene groups between environmental niches. Log-normalized counts and differential abundance statistics were determined using DESeq2 ($p < 0.05^{*}$, $p < 0.01^{**}$, $p < 0.001^{***}$).

while the remaining brine samples were the entirety of cluster 5. Finally, the frost flowers aggregated into their own grouping, with one secondary lower brine metagenome. In general, the metagenomes of the sea ice generally clustered by solid and liquid phase with the seawater groups joining the solid phase group.

2.2.6 Genomic features of vitamin biosynthesis, motility and stress responses

The 40 most abundant KEGG categories and GO terms are listed in Table S2.7A-2.7B (Supplementary Materials). In addition to taxonomic shifts in the microbial community, there were significant alterations to functional groups between environments. Genes related to cell motility (GO:0048870), biofilm formation (GO:0042710) and response to viral stress (GO:0009615) were significantly higher in the liquid brine fractions of sea ice relative to the solid phase environments (frost flowers and drained ice) and seawater (**Figure 2.6**). Genes encoding proteins implicated in response to osmotic stress (GO:0006970) were significantly higher in the primary brine fractions relative to all other environments. Conversely, the metagenomes of the solid phase environments harboured significantly higher normalised counts of genes related to vitamin biosynthesis (GO:0009110). Genes related to response to UV (GO) demonstrated no significant difference between environments.

2.3 Discussion

Everything is everywhere but the environment selects: Bacterial community composition of artificial sea ice is consistent with natural sea ice

The physiochemical and molecular parameters we measured during the experiment were generally consistent with similar measurements taken in the field (Figure 2.2) validating its use to study colonisation and ecology of microorganisms. However, the salinity values were generally inconsistent with expected readings. The reason for this is likely two fold; firstly the salinity of the source seawater was much lower than anticipated possibly owing to mis-calibration of instrumentation and secondly, the brine was inevitably over-drained due to unavoidable melt of the ice matrix during collection. Nevertheless, given the consistency of our other post-experiment parameters, those taken during ice formation, and the taxonomic and functional separation evidenced in the sequence space, we consider the model valid for testing our ecological hypotheses in a simplified system.

This type of experiment is conducive to answering some of the big questions in microbial ecology; such as the hotly debated “everything is everywhere but the environment selects”. The seawater used to grow the sea ice was not from an area subject to annual ice formation. In turn, it could be surmised that this rhetoric is supported if the same microorganisms are selected for in our artificial sea ice as has been documented in natural sea ice assemblages. Following this logic, we analysed the genera noted as candidate sea ice bacteria by (Deming and Eric Collins, 2016). We found instances where candidate sea ice genera went undetected in the seawater samples but were detected in one or more of the sea ice profile environments and/or increased significantly in abundance (**Table S2.4A-S2.4B; Supplementary Materials**).

Eukaryotes undergo stochastic selection processes within the sea-ice matrix

Based on the 18S rRNA gene amplicon sequencing, the eukaryotic organisms appeared to be stochastically incorporated into the sea ice as no consistent differences between the seeding seawater and whole cores were observed (**Figure 2.4**). The downward growth of consolidated sea ice means the age is vertically stratified. Accordingly, eukaryotic communities were structured by ice age such that communities were more similar in the upper portion relative to the lower portion, independent of niche. Moreover, the communities of the oldest ice more closely resembled the seawater at the start of the experiment while the newest ice showed a greater similarity in eukaryotic composition to the seawater post-experiment. This supports the hypothesis that the composition of the seeding community is the major driver of eukaryotic community assemblage in newly formed sea ice. Conversely, bacterial communities segregated by niche suggesting that co-transportation with eukaryotic cells are not major drivers of sea ice bacterial communities.

Bacterial colonisation and cell enrichment in sea ice brines is niche-based

The observation of bacterial communities differences in different niches based on their structure and composition supports the notion of environmental selection of underlying genetic traits (**Figure 2.4**). This is consistent with the appearance and increased abundance of archetypal natural sea ice populations in the experimental sea ice relative to the seed seawater (**Table S2.4A-S2.4B; Supplementary Materials**). For example, there is a clear dominance of the ubiquitous sea ice taxa Alteromonadaceae in the most saline brine phases. This family is often a principal taxa in natural sea ice samples and many representatives are halotolerant psychrotrophs (Bowman, 1998; Bian *et al.*, 2012). Reciprocally, there was a predominance of atypical sea ice organisms in the drained ice matrix, for instance, *Arcobacter* (class Epsilonproteobacteria). The partitioning of atypical and typical sea ice bacterial genera into the solid and liquid phases of the profile illustrated the selective conditions of the liquid brine channels and their influence over community structure. Conversely, it suggests that solid phase community structure is driven by either stochastic incorporation or differing processes.

The colonisation source of frost flowers is mainly attributed to the rejection of brine onto the surface of the sea ice (Bowman and Deming, 2010; Bowman *et al.*, 2013). Despite sharing similar physiochemical conditions (*e.g.* high salinity), the upper brine and frost flowers contained distinct bacterial communities (**Figure 2.3**). Frost flower communities were most comparable to those harboured by the ice matrix and again illustrated the partitioning of bacteria into solid and liquid phases. The major genus present in frost flowers was *Neptuniibacter*, a member of the family Oceanospiralles, which are common and abundant across natural sea ice assemblages (Deming and Eric Collins, 2016). In addition, elevated bacterial densities relative to the seawater and selective occurrence of specific taxa has been previously observed in natural frost flower bacterial communities leading to the conclusion that a similar enrichment process may be taking place in our artificial frost flowers (Bowman *et al.*, 2013).

The enrichment of specific genera in frost flowers could be explained by the coupled increase in the abundance of bacteria-consuming protozoa as has been well documented in other environments (Matz and Kjelleberg, 2005). *Paraphysomonas* was the most numerous eukaryotic taxon in our artificial frost flowers and has been identified previously in natural sea ice assemblages. They are also known to graze on bacteria (Zubkov and Sleigh, 1995; Ikävalko, 1998; Tophøj *et al.*, 2018). The high density of bacterial cells found in frost flowers could serve as an attractant to heterotrophic bacterivores that utilise the liquid brine channels to actively move towards prey. Taxon-specific bacterial adaptations against predation could, therefore, have a significant role in shaping frost flower microbial communities.

Co-transportation with algae as a driver for bacterial community composition in the solid phase

Although a clear difference between how bacterial and eukaryotic communities partition in the sea ice profile was observed, the possibility that individual populations exert cross kingdom influence cannot be excluded. We found significant co-abundance profiles in our sequence dataset between *Neptuniibacter* and a number of eukaryotic populations including *Nitzschia* and *Thalassiosira* (**Table S2.5**; Supplementary Materials). *Neptuniibacter* has been linked to the phycosphere of *Nitzschia* in culture (Behringer *et al.*, 2018) and was detected in the cultures of toxigenic diatoms (Guannel, Horner-Devine and Rocap, 2011). In addition, close relatives have been isolated from frazil ice (Celussi *et al.*, 2008) and significantly contribute to cobalamin production in the Southern Ocean sea ice edge where microalgae are abundant (Bertrand *et al.*, 2011). Furthermore, *Neptuniibacter* family members can have close associations with haptophyte blooms where they are also implicated in vitamin biosynthesis (Delmont *et al.*, 2015). Metagenomic analyses revealed a statistically significant increase in reads related to “vitamin biosynthetic process” in the drained ice and frost flower environments, where Neptuniibacter populations dominated, and could be due to their association with ice-scavenged microalgae (**Figure 2.3**). This is consistent with microalgae influence on specific bacterial populations in the solid phase of sea ice and the observation of a higher proportion of eukaryotic reads in the ice matrix metagenomes (**Figure 2.5**).

Environmental and physiological drivers for niche based selection of bacteria

Alternative hypotheses to explain the formation of a distinct sea ice brine bacterial community include the ability of bacteria to actively move towards the newly formed, nutrient-rich environment. Genes related to motility were detected in higher abundance in the liquid brine relative to all other environments (**Figure 2.6**). Previous metagenomic studies have also demonstrated a selection toward genes related to motility in sea ice (Maccario *et al.*, 2019). Although metagenomic analyses only demonstrates functional potential, motility of sea ice microorganisms at very low temperatures has been evidenced (Junge, Eicken and Deming, 2003).

Once bacteria have been passively or actively incorporated into the sea ice, the presence of steep gradients in environmental parameters are likely to exert a selection pressure on bacterial populations. During sea ice formation, salt and other particles are excluded from the ice crystals resulting in hypersaline brine channels with increased cell-cell and cell-particle contact rates. In turn, bacteria face greater osmotic pressures and may be exposed to more interactions and possibility of infection by viruses. This is reflected in the brine metagenomes which encode more functions related to osmotic, viral stress and biofilm formation (**Figure 2.6**). While viral abundance was not quantified in this study, very high virus:bacteria ratios (5.5 to 170 × 107 ml^{-1} VLP scaled brine volume) have been reported in Arctic sea ice (Collins and Deming, 2011).

A critical constraint in sea ice profiles in natural environments is the exposure to high levels of UV radiation during the spring and summer months which has been related to an increase in genes involved in resistance to photo-oxidative stress (Maccario *et al.*, 2019). We were unable to find any statistically significant differences between environments in relation to genes encoding UV protection. Given UV light was not a factor in this study, this result supports the idea that bacteria are selected for by prevailing environmental conditions.

2.4 Conclusion

While it has been shown that algae are scavenged by frazil ice crystals during sea ice formation, it has only been hypothesised that the sea ice-specific bacterial communities observed are driven by their association to diatoms (Grossmann and Gleitz, 1993; Grossmann, 1994; Grossmann and Dieckmann, 1994; Weissenberger and Grossmann, 1998; Riedel *et al.*, 2007). This hypothesis is inconsistent with the high bacterial loads observed in sea ice (even in the absence of eukaryotic cells; Eronen-Rasimus *et al.*, 2014), yet low bacterial colonisation of pelagic diatoms (Droop and Elson, 1966; Kaczmarska *et al.*, 2005; Crenn, Duffieux and Jeanthon, 2018). Our experiment suggested that microalgae may be a driver of the dominant bacterial populations that are entrained in the solid phase of sea ice and frost flowers while the liquid brine bacterial communities are shaped by prevailing environmental conditions, reflected in their metagenomic functional content. These genetic predispositions would allow microbes to detect and migrate toward the nutrient rich sea ice brines, ensure retention during brine rejection and proliferate rapidly in high nutrient conditions.

2.5 Supplementary Figures

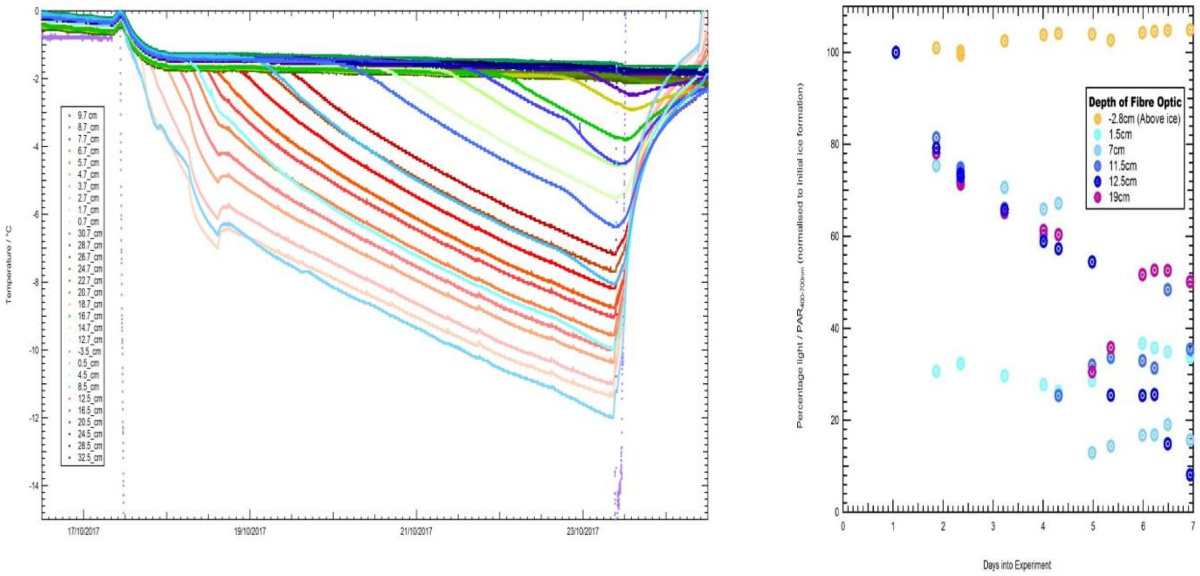


Figure S2.1: (A) Evolution of ice temperature at different depths of the sea ice chamber and (B) Photosynthetically Active Radiation measurements through ice, taken periodically over the course of the experiment

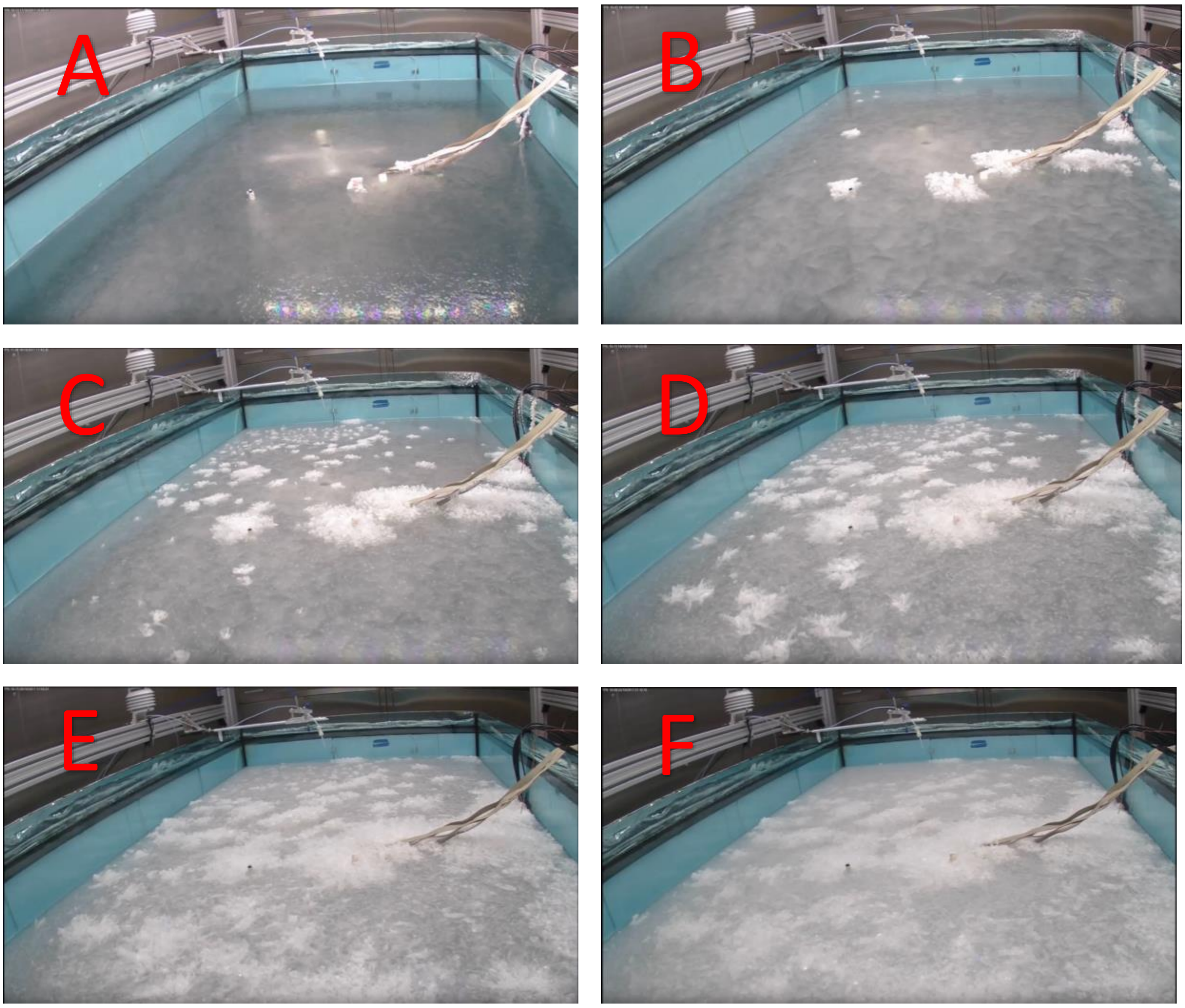


Figure S2.2: Images from the time lapse video taken intermittently to monitor frost flower growth. (A) 17/10/2017 – first stages of ice formation (B) 18/10/2017 – first signs of frost flower growth (C) 18/10/2017 – rapid expansion of frost flower coverage (C) 19/10/2017 (D) 20/10/2017 (E) 21/10/2017 and (F) 22/10/2017 – biological sampling

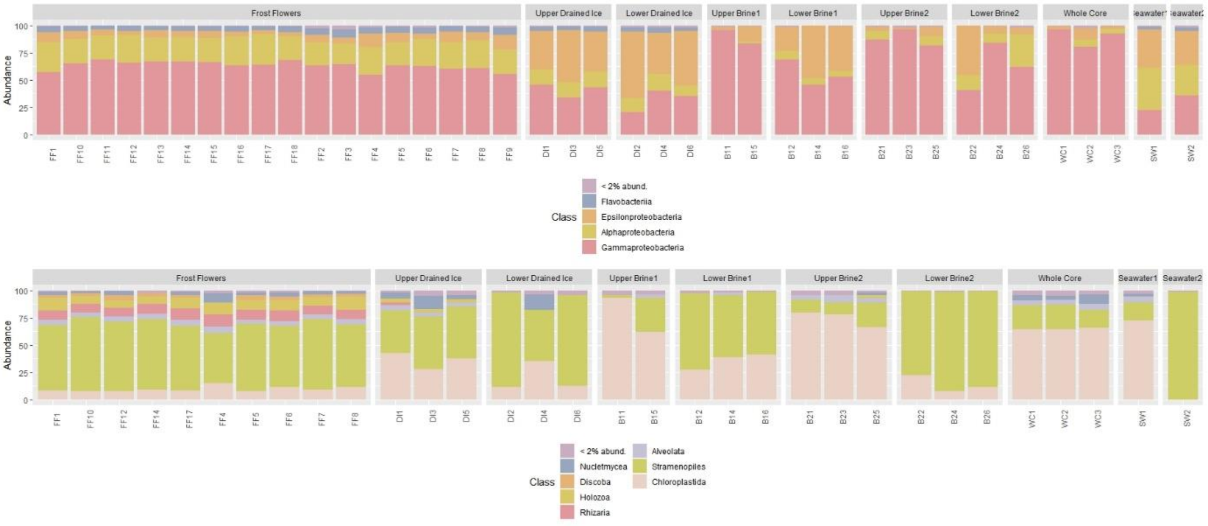


Figure S2.3 Histograms of the relative abundance of the taxa in the (C) 16S rRNA dataset and (D) the 18S rRNA dataset at Class level.

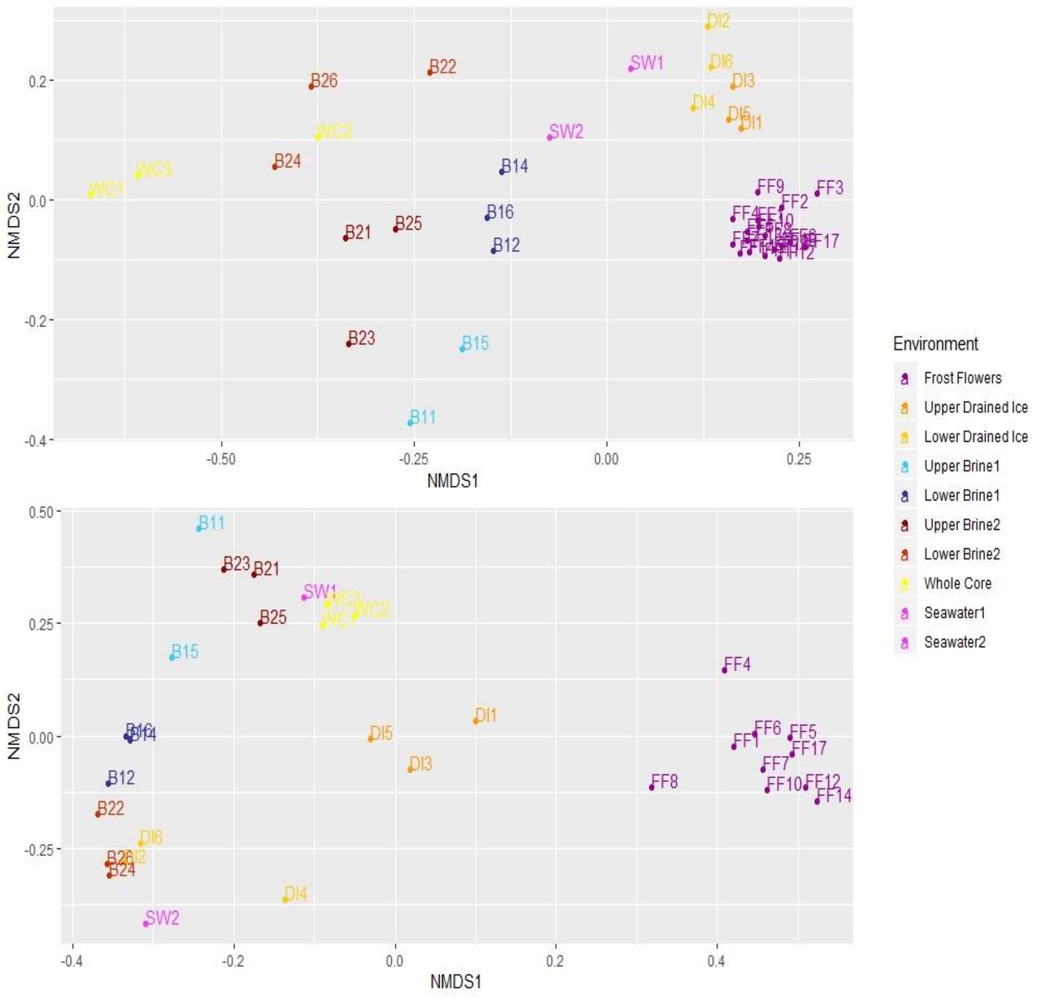


Figure S2.4 NMDS calculated with the Bray-Curtis distances of the relative abundance of ASVs in the (A) 16S rRNA and (B) 18S rRNA ASV table.

2.5 Bibliography

Andrews, S. *et al.* (2012) ‘FastQC’. Babraham, UK.

Behringer, G. *et al.* (2018) ‘Bacterial communities of diatoms display strong conservation across strains and time’, *Frontiers in Microbiology*. Frontiers Media S.A., 9(APR). doi: 10.3389/fmicb.2018.00659.

Benjamini, Y. and Hochberg, Y. (1995) ‘Benjamini-1995.pdf’, *Journal of the Royal Statistical Society B*, pp. 289–300. doi: 10.2307/2346101.

Bertrand, E. M. *et al.* (2011) ‘Vitamin B12 biosynthesis gene diversity in the Ross Sea: the identification of a new group of putative polar B12 biosynthesizers’, *Environmental Microbiology*, 13(5), pp. 1285–1298. doi: 10.1111/j.1462-2920.2011.02428.x.

Bian, F. *et al.* (2012) ‘Genome sequences of six Pseudoalteromonas strains isolated from arctic sea ice’, *Journal of Bacteriology*, pp. 908–909. doi: 10.1128/JB.06427-11.

Bowman, J. P. (1998) ‘Pseudoalteromonas prydzensis sp. nov., a psychrotrophic, halotolerant bacterium from Antarctic sea ice’, *International Journal of Systematic Bacteriology*. Microbiology Society, 48(3), pp. 1037–1041. doi: 10.1099/00207713-48-3-1037.

Bowman, J. S. *et al.* (2013) ‘Selective occurrence of Rhizobiales in frost flowers on the surface of young sea ice near Barrow, Alaska and distribution in the polar marine rare biosphere’, *Environmental Microbiology Reports*, 5(4), pp. 575–582. doi: 10.1111/1758-2229.12047.

Bowman, J. S. *et al.* (2014) ‘The genetic potential for key biogeochemical processes in Arctic frost flowers and young sea ice revealed by metagenomic analysis’, *FEMS Microbiology Ecology*. Blackwell Publishing Ltd, 89(2), pp. 376–387. doi: 10.1111/1574-6941.12331.

Bowman, J. S. and Deming, J. W. (2010) ‘Elevated bacterial abundance and exopolymers in saline frost flowers and implications for atmospheric chemistry and microbial dispersal’, *Geophysical Research Letters*, 37(13), p. n/a-n/a. doi: 10.1029/2010GL043020.

Callahan, B. J. *et al.* (2016) ‘DADA2: High-resolution sample inference from Illumina amplicon data’, *Nature Methods*. Nature Publishing Group, 13(7), pp. 581–583. doi: 10.1038/nmeth.3869.

Celussi, M. *et al.* (2008) ‘Organic-matter degradative potential of Halomonas glaciei isolated from frazil ice in the Ross Sea (Antarctica)’, *FEMS Microbiology Ecology*, 65(3), pp. 504–512. doi: 10.1111/j.1574-6941.2008.00551.x.

Chamberlain, S. *et al.* (2019) ‘taxize: Taxonomic Information from Around the Web’. Available at: https://cran.r-project.org/package=taxize.

Collins, R. E. and Deming, J. W. (2011) ‘Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up’, *Polar Biology*, 34(12), pp. 1831–1841. doi: 10.1007/s00300-011-1008-z.

Crenn, K., Duffieux, D. and Jeanthon, C. (2018) ‘Bacterial Epibiotic Communities of Ubiquitous and Abundant Marine Diatoms Are Distinct in Short- and Long-Term Associations’, *Frontiers in Microbiology*. Frontiers Media SA, 9. doi: 10.3389/fmicb.2018.02879.

Delmont, T. O. *et al.* (2015) ‘Genome reconstructions indicate the partitioning of ecological functions inside a phytoplankton bloom in the Amundsen Sea, Antarctica’, *Frontiers in Microbiology*. Frontiers Media S.A., 6(OCT). doi: 10.3389/fmicb.2015.01090.

Deming, J. W. and Eric Collins, R. (2016) ‘Sea ice as a habitat for Bacteria, Archaea and viruses’, in *Sea Ice: Third Edition*. Wiley Blackwell, pp. 326–351. doi: 10.1002/9781118778371.ch13.

Droop, M. R. and Elson, K. G. R. (1966) ‘Are pelagic diatoms free from bacteria? [31]’, *Nature*, pp. 1096–1097. doi: 10.1038/2111096a0.

Eronen-Rasimus, E. *et al.* (2014) ‘Bacterial community dynamics and activity in relation to dissolved organic matter availability during sea-ice formation in a mesocosm experiment’, *MicrobiologyOpen*, 3(1), pp. 139–156. doi: 10.1002/mbo3.157.

Ewert, M. and Deming, J. W. (2013) ‘Sea ice microorganisms: Environmental constraints and extracellular responses’, *Biology*. MDPI AG, 2(2), pp. 603–628. doi: 10.3390/biology2020603.

Fripiat, F. *et al.* (2007) ‘Diatom-induced silicon isotopic fractionation in Antarctic sea ice’, *Journal of Geophysical Research*, 112(G2), p. G02001. doi: 10.1029/2006JG000244.

Garrison, D. L., Close, A. R. and Reimnitz, E. (1989) ‘Algae concentrated by frazil ice: Evidence from laboratory experiments and field measurements', *Antarctic Science*, 1(4), pp. 313–316. doi: 10.1017/S0954102089000477.

Gradinger, R. and Ikävalko, J. (1998) 'Organism incorporation into newly forming Arctic sea ice in the Greenland Sea', *Journal of Plankton Research*, 20(5), pp. 871–886. doi: 10.1093/plankt/20.5.871.

Grossmann, S. (1994) 'Bacterial activity in sea ice and open water of the Weddell Sea, Antarctica: A microautoradiographic study', *Microbial Ecology*. Springer-Verlag, 28(1), pp. 1–18. doi: 10.1007/BF00170244.

Grossmann, S. and Dieckmann, G. S. (1994) 'Bacterial Standing Stock, Activity, and Carbon Production during Formation and Growth of Sea Ice in the Weddell Sea, Antarctica.', *Applied and environmental microbiology*, 60(8), pp. 2746–53. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16349347 (Accessed: 6 February 2020).

Grossmann, S. and Gleitz, M. (1993) *Microbial responses to experimental sea-ice formation: implications for the establishment of Antarctic sea-ice communities*, *Mar. Biol. Ecol*.

Guannel, M., Horner-Devine, M. and Rocap, G. (2011) 'Bacterial community composition differs with species and toxigenicity of the diatom Pseudo-nitzschia', *Aquatic Microbial Ecology*, 64(2), pp. 117–133. doi: 10.3354/ame01513.

Harrell Jr, F. E., with contributions from Charles Dupont and many others. (2019) 'Hmisc: Harrell Miscellaneous'. Available at: https://cran.r-project.org/package=Hmisc.

Huerta-Cepas, J. *et al.* (2016) 'eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences', *Nucleic Acids Research*, 44(D1), pp. D286–D293. doi: 10.1093/nar/gkv1248.

Huerta-Cepas, J. *et al.* (2017) 'Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper', *Molecular Biology and Evolution*. Oxford University Press, 34(8), pp. 2115–2122. doi: 10.1093/molbev/msx148.

Ikävalko, J. (1998) 'Further observations on flagellates within sea ice in northern Bothnian Bay, the Baltic Sea', *Polar Biology*. Springer, 19(5), pp. 323–329. doi: 10.1007/s003000050253.

Junge, K., Eicken, H. and Deming, J. W. (2003) ‘Motility of Colwellia psychrerythraea strain 34H at subzero temperatures’, *Applied and Environmental Microbiology*, 69(7), pp. 4282–4284. doi: 10.1128/AEM.69.7.4282-4284.2003.

Junge, K., Eicken, H. and Deming, J. W. (2004) ‘Bacterial Activity at -2 to -20°C in Arctic Wintertime Sea Ice’, *Applied and Environmental Microbiology*, 70(1), pp. 550–557. doi: 10.1128/AEM.70.1.550-557.2004.

Kaczmarska, I. *et al.* (2005) ‘Diversity and distribution of epibiotic bacteria on Pseudo-nitzschia multiseries (Bacillariophyceae) in culture, and comparison with those on diatoms in native seawater’, *Harmful Algae*. Elsevier, 4(4), pp. 725–741. doi: 10.1016/j.hal.2004.10.001.

Lebedev, V. V. (1938) ‘Rost l’da v arkticheskikh rekakh i moriakh v zavisimosti ot otritsatel’nykh temperatur vozdukha’, *Probl. Arktiki*, 5, pp. 9–25.

Love, M. I., Huber, W. and Anders, S. (2014) ‘Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2’, *Genome Biology*, 15(12), p. 550. doi: 10.1186/s13059-014-0550-8.

Maccario, L. *et al.* (2019) ‘Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover’, *Scientific Reports*. Nature Publishing Group, 9(1). doi: 10.1038/s41598-019-38744-y.

Martin, M. (2011) ‘Cutadapt removes adapter sequences from high-throughput sequencing reads’, *EMBnet.journal*. EMBnet Stichting, 17(1), p. 10. doi: 10.14806/ej.17.1.200.

Matz, C. and Kjelleberg, S. (2005) ‘Off the hook - How bacteria survive protozoan grazing’, *Trends in Microbiology*. Elsevier Ltd, 13(7), pp. 302–307. doi: 10.1016/j.tim.2005.05.009.

McMurdie, P. J. and Holmes, S. (2013) ‘phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data’, *PLoS ONE*. Edited by M. Watson, 8(4), p. e61217. doi: 10.1371/journal.pone.0061217.

Riedel, A. *et al.* (2007) ‘Enrichment of nutrients, exopolymeric substances and microorganisms in newly formed sea ice on the Mackenzie shelf’, *Marine Ecology Progress Series*, 342, pp. 55–67. doi: 10.3354/meps342055.

Rózańska, M., Poulin, M. and Gosselin, M. (2008) ‘Protist entrapment in newly formed sea ice in the Coastal Arctic Ocean', *Journal of Marine Systems*, 74(3–4), pp. 887–901. doi: 10.1016/j.jmarsys.2007.11.009.

Scott Chamberlain and Eduard Szocs (2013) 'taxize - taxonomic search and retrieval in R', *F1000Research*. Available at: http://f1000research.com/articles/2-191/v2.

Style, R. W. and Worster, M. G. (2009) 'Frost flower formation on sea ice and lake ice', *Geophysical Research Letters*, 36(11), p. L11501. doi: 10.1029/2009GL037304.

Tophøj, J. *et al.* (2018) 'Feeding and growth of the marine heterotrophic nanoflagellates, procryptobia sorokini and paraphysomonas imperforata on a bacterium, pseudoalteromonas sp. With an inducible defence against grazing', *PLoS ONE*. Public Library of Science, 13(4). doi: 10.1371/journal.pone.0195935.

Weissenberger, J. and Grossmann, S. (1998) 'Experimental formation of sea ice: Importance of water circulation and wave action for incorporation of phytoplankton and bacteria', *Polar Biology*, 20(3), pp. 178–188. doi: 10.1007/s003000050294.

Wickham, H. (2016) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. Available at: https://ggplot2.tidyverse.org.

Zubkov, M. V. and Sleigh, M. A. (1995) 'Bacterivory by starved marine heterotrophic nanoflagellates of two species which feed differently, estimated by uptake of dual radioactive-labelled bacteria', *FEMS Microbiology Ecology*. No longer published by Elsevier, 17(1), pp. 57–65. doi: 10.1016/0168-6496(95)00011-X.

Chapter 3: Exploring the taxonomic and metagenomic content of sea ice and saline snow

Layton R, Vogel T.M and Larose, C

3. 0 Introduction

Climate change has accelerated Arctic warming by 0.75°C over the last decade and resulted in unprecedented environmental changes across the sea ice landscape. A decline of approximately 50% in sea ice cover since the beginning of the satellite record and a general shift from multiyear ice (MYI) to new and young sea ice (FYI) is influencing the Arctic ecosystem and global processes (Vihma, 2014). A thorough knowledge of sea ice microbial communities and their biogeochemical importance underscores the ability to discern ecosystem function and how this will be affected by the change in the Arctic ice landscape. Initial studies have characterised microbial biota of sea ice and shown it's variability in relation to space (Brinkmeyer et al., 2003; Yergeau et al., 2017), time (Eric Collins, Rocap, & Deming, 2010) and ice type (Hatam, Lange, Beckers, Haas, & Lanoil, 2016). However, the specific physio-chemical parameters and biological interactions that shape taxonomic and functional structure of sea ice communities are not well understood. In addition, other environmental habitats, such as saline snow cover and frost flowers, are likely to increase across the Arctic in line with the increasing extent of FYI and yet our understanding of their ecology is in its infancy.

The study of sea ice bacterial communities began with the application of culture based techniques. More recently, taxonomic profiling using 16S and 18S rRNA gene sequencing has been used in numerous sea ice ecology papers and has established the presence of distinct communities in which certain rRNA phylotypes are repetitively selected for independent of geography (Junge, Cameron, & Nunn, 2019). As such, shared environmental constraints in sea ice environments appear to drive the appearance of specific taxa. Although eukaryotes and prokaryotes are intimately connected through biogeochemical and food web networks (Cirri & Pohnert, 2019; Ramanan, Kim, Cho, Oh, & Kim, 2016), no studies have combined amplicon sequencing that targets both of these communities in sea ice. Therefore, the significance of their relationships remains largely unexplored. In addition, while taxonomic profiling is recently prevalent in sea ice ecology, metagenomic profiling is limited to a handful of studies despite the increase in its availability (Bowman, Berthiaume, Armbrust, & Deming, 2014; Maccario, Carpenter, Deming, Vogel, & Larose, 2019; Yergeau et al., 2017). Shotgun metagenomics provides insight into the functional potential of microbial communities and could, therefore, provide a means to explore the ecological processes that shape microbial communities in sea ice. Sea ice microbial ecology studies would further benefit from the exploration of connected environments such as the snowpack and the atmosphere. They are often disregarded in studies and yet provide potential sources of microbial communities and can affect the physiochemical conditions (*e.g*. snowpack and UV extinction) and in turn alter the prevailing selective conditions.

Sea ice community microbial population function (*i.e*., role) has mostly been gleamed from targeted experiments and genome sequencing. Archetypal sea ice microorganisms have the potential to survive and even thrive under conditions of perceived severe environmental stress and to respond rapidly to available sources of carbon and in turn to contribute significantly to carbon cycling. For example, members of the class Gammaproteobacteria from sea ice have been shown to respond rapidly to carbon inputs consistent with their putative copiotrophic nature and potential metabolic versatility (Eronen-Rasimus et al., 2014). Other studies have illustrated the potential for cold active catalysis in sea ice community populations and the role of genomic flexibility in successful colonisation and persistence under the selective conditions (Bowman & Deming, 2014; Vollmers et al., 2013; Wang et al., 2018). However, sea ice bacteria are genetically undersampled and a bias toward readily culturable members of the community exists. On the other hand, shotgun metagenomics provides a subsample of entire communities and in turn their overall metabolic potential.

Few studies have applied metagenomic based techniques in studies of sea ice microbial communities however, those that have consistently found differences between the functional repertoire of sea ice, frost flowers, saline snow and the underlying water (Bowman et al., 2014; Maccario et al., 2019; Yergeau et al., 2017). Understanding the projected environmental changes on community dynamics necessitates a fundamental knowledge on which bacterial and eukaryotic communities and functional profiles occur in light of changing physio-chemical conditions. This is particularly important for environments of emerging importance, such as FYI and the overlying snow. In light of this, we applied rRNA amplicon and shotgun metagenomic sequencing to a Greenland ocean-sea ice-snow profile with the following objectives (1) characterize the bacterial and eukaryotic community taxonomy in seawater and the overlying ice and snow horizons and (2) profile the functional potential of these communities to infer underlying selection mechanisms

3.1 Methods

3.1.1 Field site

The spring field study was carried out between March 9th-17th, 2014 on board the Argelvor research sailboat that was frozen into the ice in Amdrup Havn (70° 28' 28.3"N, 21° 54' 56.4"W), located near Ittoqqortoormiit on the north east coast of Greenland. The research vessel was equipped with two windmills, 3 solar panels and a 6.5 kW diesel generator from SDMO to provide enough power to run the on-board laboratory and scientific instruments.

3.1.2 Snow and ice sampling

A large ice section (50 cm x 50 cm x 140 cm) was extracted from the sea-ice and divided into three vertical sections. 4 sections were sampled: surface ice (0-10 cm), a middle section (50-60 cm), an algae rich section (60-70 cm) and the bottom (130-140 cm). The snowpack consisted of seasonal snow that accumulated above the sea-ice. Shallow pits (45 cm) were dug and both surface (top 5 cm) and basal samples (30 cm) were collected. Samples were collected in sterile cuvettes and stored at -20°C until analysis. Samples for microbial analyses were collected in three 3-L sterile sampling bags using a sterilized Teflon shovel for a total of 9 L of snow equivalent to about 3 L of water depending on snow density. To minimize contamination, Tyvex® body suits and latex gloves were worn during sampling and gloves were worn during all subsequent sample handling. The pH of melted snow samples was monitored at 20°C (Heito pH meter, Paris). The electrode was initially calibrated automatically, but was manually calibrated prior to analysis with three different pH buffers (pH 4, 7 and 10, Heito).

3.1.3 Chemistry

For cation/anion chemistry measurements (major, minor ions and organic acids), Inorganic ions (F^- Fl, NO_3^-, Cl^-, SO_4^{2-}, NH_4^+, Ca^{2+}, Na^+, K^+ and Mg^{2+}) and organic acid (methylsulfonic acid (MSA), glutaric acid (Glut), oxalic acid (Ox), acetate with a possible contribution of glycolate (Ace.Glyc), and formate concentrations were measured at the Institut de Géophysique de l'Environnement by conductivity-suppressed Ion Chromatography using a Dionex ICS 3000. Due to the high chloride and sodium concentrations from the fjord, samples were diluted 10 fold for organic acid and minor ion analyses and 100-1000 times for major ions analyses.

3.1.4 DNA isolation and sequencing

Samples were processed in the laboratory located on Argelvor. Snow samples were left to melt for 6 hours at room temperature and the resultant water was filtered with sterile 0.22-µm pore size, 47-mm diameter filters (Millipore) using a sterile filtration unit (Nalge Nunc International Corporation). Meltwater samples were filtered immediately. Filters were stored in sterile bead-beating tubes at -20°C until analysed at the EMG laboratory in France. Procedural blanks were performed by filtering Nanopure water (Siemens) using the same procedure. DNA was extracted using using the Power Water DNA isolation kit (Machery-Nagel), following the manufacturer's instructions.

Preparation of both 16S rRNA and 18S rRNA amplicon libraries was performed according to the standard instructions of the 16S Metagenomic Sequencing Library Preparation protocol (Illumina, Inc., San Diego, CA, United States). The primers used to amplify the (V3-V4) region of the rRNA genes can be found in **Table S3.1**. Sequences were obtained on the Illumina MiSeq platform in a 2 × 250 bp paired-end run (in-house) following the standard instructions of the 16S Metagenomic Sequencing Library Preparation protocol (Illumina, Inc., San Diego, CA, United States). When sufficient DNA concentrations were available from samples, metagenomic sequencing was performed. Metagenomic libraries were prepared using the Nextera XT Library Prep Kit and Workflow (Illumina) according to manufacturer's instructions. All sequencing libraries were paired end sequenced on a MiSeq sequencer (Illumina) at our laboratory in Lyon.

3.1.5 Bioinformatics and Statistical Analyses

Both bacterial and eukaryotic library sequence datasets were treated by similar bioinformatic pipelines. The raw paired-end reads were primer-trimmed using 'cutadapt' (Martin, 2011), and subject to DADA2 analysis processed following the DADA2 (version 1.12) pipeline (Callahan et al., 2016). Briefly, sequence read quality profiles were inspected and sequences filtered with the filterAndTrim function. Based on their quality scores, sequences were truncated (16S: 231 and 227 bp; 18S: 220 and 200 bp) and the maximum error rate set to 3 and 5 for forward and reverse sequence reads, respectively. The resulting reads were denoised based on an error model learned from the data and paired ends were merged. Chimeras were removed using the "consensus" method. The 16S Ribosomal Database Project's (trainset 16) and 18S Silva (v132) databases were used for taxonomy assignment by RDP's naive Bayesian classifier within DADA2. The amplicon sequence variants (ASVs) which were not taxonomically assigned to Bacteria/Eukarya or occurred with only a single sequence in the whole dataset were excluded from further analysis. All statistical analyses were conducted using R (v3.4.1)1 in RStudio (v1.0.153; RStudio Team, 2015). Sample data was managed using the R package 'phyloseq' (v1.20.0; McMurdie & Holmes, 2013) and plots were generated using the R package 'ggplot2' (v2.2.1; Wickham, 2016).

Presence-absence analysis was used to identify overlaps in 16S rRNA gene amplicons (ASVs) between the horizons in the ocean-sea ice-snow profile. ASVs were excluded if they did not occurr in at least 2 samples with a count of at least 2. Lists were used to identify intersects and produce Venn diagrams. To account for proportional abundance Sourcetracker2 (v2.0.1; Knights et al., 2011) was used as a Bayesian approach to estimate the proportion of ASVs in saline snow attributed to possible sources. All saline snow samples were designated as sinks and all other horizons as sources. SourceTracker2 was run with default settings and an alpha of 0.001.

Non-metric multi-dimensional scaling (NMDS) analysis was conducted on Bray-Curtis transformed ASV abundance matrices. The significance of the clustering was tested using a permutational multivariate analysis of variance. The fold-change in abundance of each ASV between horizons was calculated using the R package 'DESeq2' (Love, Huber, & Anders, 2014). Non-metric multi-dimensional scaling (NMDS) analysis was conducted on Bray-Curtis transformed metagenomics abundance matrices of KEGG categories. The significance of the clustering was tested using a permutational multivariate analysis of variance. The 30% most abundant KEGG categories with 60% best environmental fit were selected using ‘ordiselect’ of the goeveg R package (*Package “goeveg” Type Package Title Functions for Community Data and Ordinations*, 2018). Prior to differential abundance analyses of GO Terms by DESeq2, counts of less than 10 were eliminated from the count matrix. GO terms were considered differentially abundant if they had a log fold change > 2 and an adjusted p-values < 0.05. Terms enriched in the ice and snow horizons relative to the underlying seawater were subject to overlap analyses.

In combination with sequencing data, chemical data was evaluated using co-inertia analysis. Briefly, with the exception of pH, chemical data was log transformed and ASV tables were Hellinger transformed. A PCA was calculated for each dataset type using the ade4 package in R and co-inertia analysis was implemented to test the influence of chemistry on bacterial communities. The significance of the co-inertia was tested using a permutation test (1000 permutations).

3.2 Results

3.2.1 Differences in Taxonomic Community Composition Across the ocean-sea ice-snow profile

3.2.1.1 Bacterial Community Composition of the Marine Profile

The surface snow and the atmosphere samples were more diverse relative to the marine profile. Dominant classes in the atmosphere include Actinobacteria (~23%), Gammaproteobacteria (21%), Bacilli (14%) and Alphaproteobacteria (~12%). The surface snow was also characterised by a dominance of Actinobacteria (~11%), Alphaproteobacteria (~21%), Gammaproteobacteria (~12%) but also contained an abundance of Acidobacteria_Gp1 (~20%) and Cyanobacteria (~11%) (Figure S3.1).

Comparison of bacterial community composition between the different regions and fractions was conducted based on the presence/absence of an ASV. Only 32 ASVs were shared across the ocean-sea-ice-basal snow profile representing just 0.32% of the total ASVs (10,054). However, this core community represented 46% of all the sequences. All horizons had a large number of unique ASVs but their proportional abundance within their respective sampling sites varied (**Table S3.2**). Hence, the differences between bacterial communities is likely a result of both the variations in the proportional abundances of the core community as well as the differences in taxa.

At the class level, Gammaproteobacteria dominated the algal ice band and saline snow horizons while the seawater contained a larger proportion of Alphaproteobacteria. Conversely, Flavobacteriia was dominant in the upper ice horizon and to a lesser extent in the middle ice (**Figure 3.1**). In order to further investigate the differences in community composition between the different horizons, we performed differential abundance tests for all shared ASVs using DESeq2. Given the seawater acts as the main seeding source for the sea ice profile, all environments were compared to seawater. The ASV that had a fold change of absolute value higher than 1 and an adjusted *p*-value < 0.05 was defined as differentially abundant (daASVs). There were a total of 70 unique ASVs that were identified as differentially abundant between seawater and any of the ice horizons and basal snow layer. The algal ice band (n=33) and the basal snow layer (n=30) had the largest number of differentially abundant ASVs relative to the underlying seawater. In the basal snow, most enriched daASVs were annotated as the genera *Colwellia* (n = 20) and all but one increased in abundance in the basal snow. Other significantly enriched genera in the basal snow relative to the seawater were *Glaciecola* and *Polaribacter* (**Table S3.3A-F**). The algal ice band was also enriched in strains annotated as the genera *Colwellia* and *Polaribacter* and to a lesser degree by members belonging to the genera *Paraglaciecola, Octadecabacter, Illumatobacter* and *Marinomonas*. Similar enrichment patterns were observed between the ocean and surface and middle ice horizons. *Sulfitobacter* was also significantly enriched between the ocean and surface and middle ice horizons. Conversely, no ASVs were differentially abundant between the ocean and the frazil ice communities. Interestingly, two ASVs annotated as Colwellia (Argelvor16S_271 and Argelvor16S_1308), an archetypal sea ice microorganism, had an increased abundance in the ocean relative to the basal snow or middle and bottom ice horizons. Therefore, we investigated *Colwellia* ASV diversity and

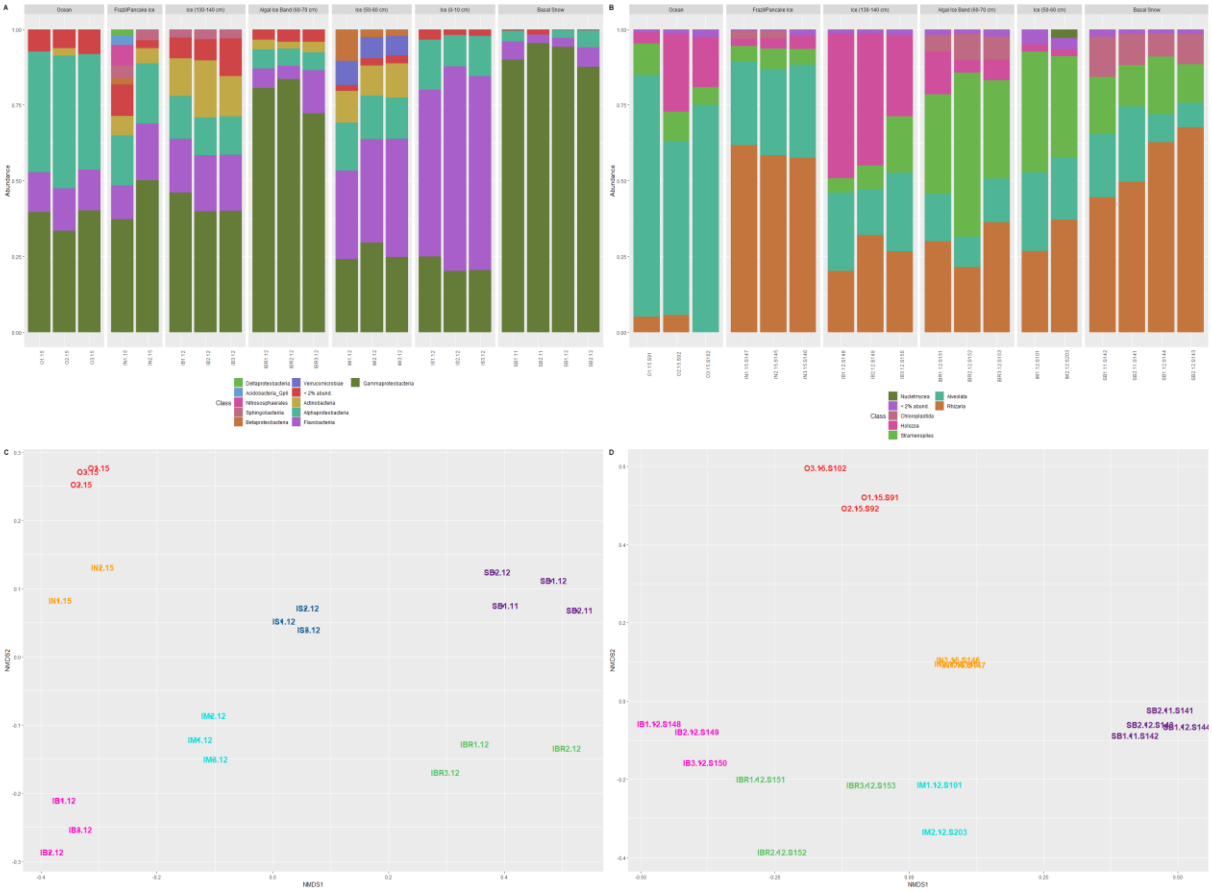


Figure 3.1: Histograms of the relative abundance of classes representing >2% abundance in the (A) 16S rRNA and (B) 18S rRNA datasets. (B) and (C) are NMDS plots based on Bray-Curtis dissimilarity of bacterial and eukaryotic relative abundance, respectively (stress 0.06897909 and 0.09999102).

separation across the sea ice profile. Samples that had at least 100 sequences annotated as *Colwellia* were rarefied to the minimum number present in any sample (Figure S2) and alpha diversity indices calculated. These indices were consistently lower in the saline snow horizons than in all other samples despite a larger initial sequencing effort (**Figure 3.2**).

Overlap analysis of 16S rRNA gene ASVs of source and sink environments demonstrated a small overlap between the seawater, surface snow and saline snow (n = 18) with a higher number of uniquely shared ASVs between the seawater and saline snow (n = 35) (**Figure S3.3**). The uniquely shared ASVs between surface snow and saline snow (n = 19) were mainly annotated as *Colwellia* and *Glacieocola* (n = 15). A similar result was obtained when comparing the surface ice, surface snow and basal snow (n = 60 and 3, respectively) although more genera annotated as *Colwellia* and *Glacieocola* belonged to the core group (**Table S3.4**). Transfer of microorganisms from ocean, sea ice and surface snow samples to saline snow samples was inferred using the software SourceTracker2. Total estimated

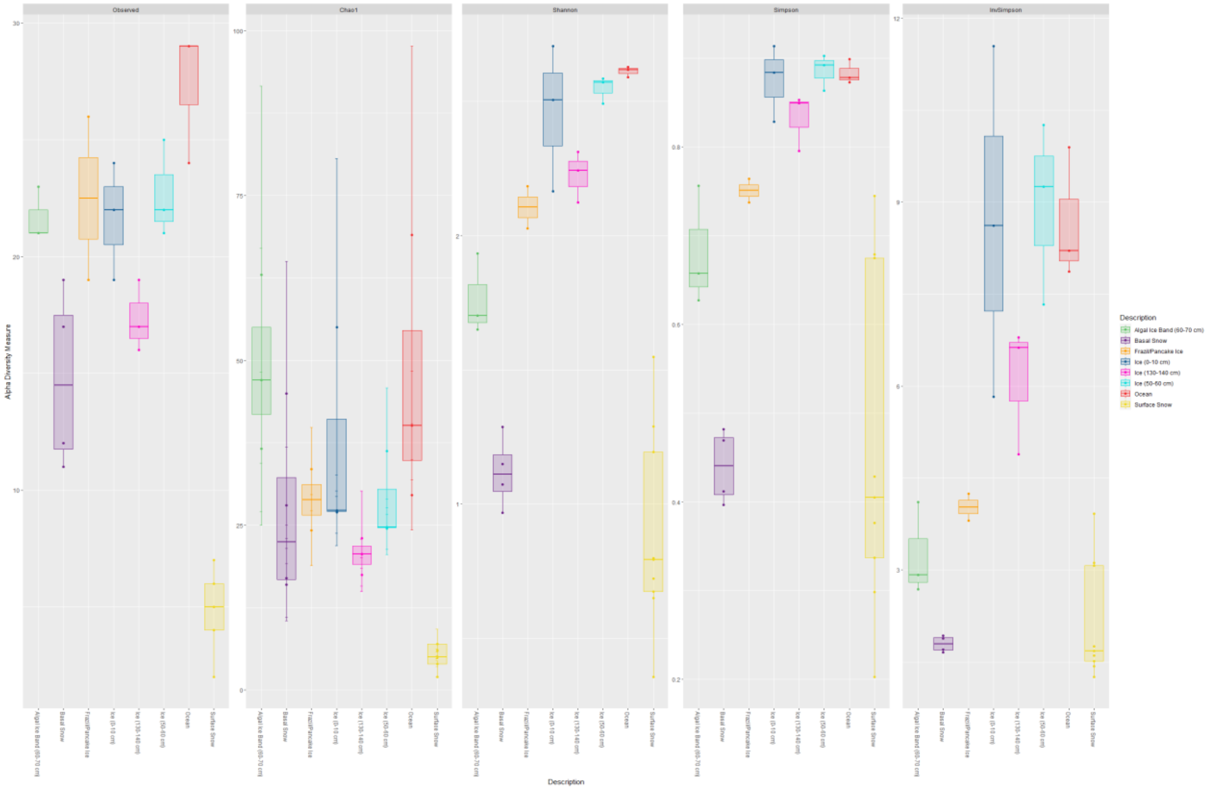


Figure 3.2: Box plots showing alpha diversity variation across samples on rarefied data annotated as the genera *Colwellia*. Only samples >100 ASVs annotated as *Colwellia* were used in determining alpha diversities.

contributions from the tested sources suggests the largest contribution was from the algal ice band (~72%). There was a significant unknown contribution factor (~13%). The next largest predicted source of saline snow microorganisms was the surface ice (~8%) followed by the surface snow (~3%) and the seawater (~2%). All other source environments were predicted to contribute to less than 1% of the saline snow community (**Figure S4)**.

3.2.1.2 Eukaryotic Community Composition of the Marine Profile

A comparable workflow was used to investigate the microbial eukayotic communities. The eukaryotic composition of the atmosphere and surface snow was variable but generally characterised by a predominance of fungi, particularly the Nucletmycea (**Figure S3.1**). In some atmosphere samples, the classes Stramenopiles and Chloroplastida constituted greater than 50% abundance but was completely absent in others. Chloroplastida were relatively consistently abundant in surface snow samples (**Figure S3.1**).

Presence/absence overlap analyses identified a core community of just 9 ASVs representing just 0.66% of the total ASVs (1372) which contributed to 18.7% of all the sequences. All horizons had a large number of unique ASVs but their proportional abundance within their respective sampling sites varied (**Table S3.2**). 18S rRNA ASVs unique to the seawater made up over half the relative abundance of the seawater samples suggesting that there are organisms selectively excluded from the ice profile.

DeSeq2 analyses identified a total of 24 unique ASVs that were differentially abundant between the seeding seawater and any of the overlying horizons. The differentially abundant ASVs belonged to just 4 classes (Alveolata, Holozoa, Rhizaria and Stramenopiles; **Figure 3.1B**; **Table S3.4A-F**). The largest number of differentially abundant ASVs were seen between the seawater and algal ice band (n=x) and all were enriched in the ice horizon. There was only one ASV identified as significantly enriched between the seawater and frazil ice (Argelvor18S_169) annotated as the genera *Gyrodinium*.

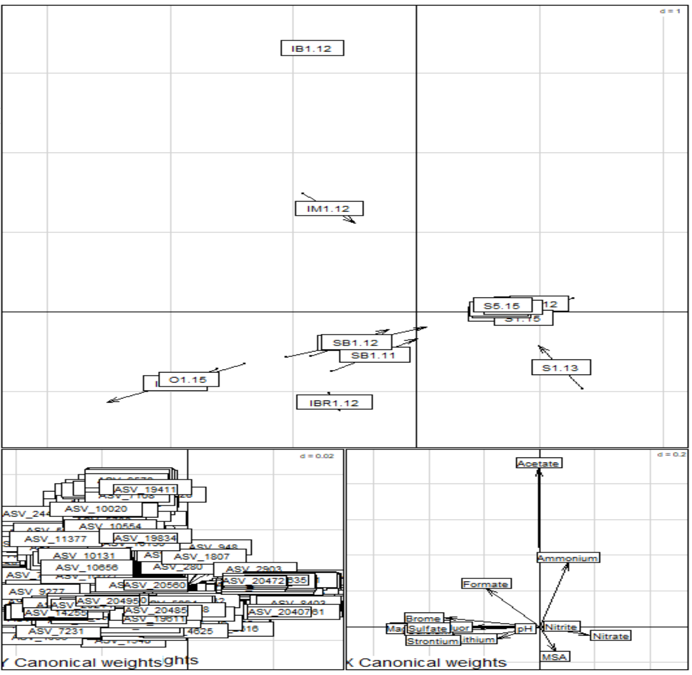


Figure 3.3: Co-inertia analysis (CIA) of the relationship between microbiota at the ASV level and chemistry. The top panels show the CIA of the ASV and OTU principal component analysis; arrows indicate where samples in the chemistry parameter dataset are relative to the ASV dataset. The lower right and left panel shows ASV and chemistry loading data, respectively.

3.2.2 Chemical Drivers of OSISS-SS Microbial Communities

Co-inertia analysis was applied to the OSISS-SS profile chemistry and microbial community structure (**Figure 3.3**). The permutation test demonstrated a significant relationship between the chemistry and microbial community structure (p=0.001, RV=0.304). The RV co-efficient represents the correlation between the microbial community structure and the chemistry data. An RV-coefficient of 1 represents a strong correlation whereas 0 represents no correlation. The first two eigen vectors of the co-inertia analyses explain 94.64 % and 3.06 % of the explained variance, respectively. The chemical parameters that had the most influence on the co-structure as observed by the lengths of the vector arrows in (Figure 3.3) formed six major axes; 1) ions 2) Nitrate 3) Formate 4) Ammonium 5) MSA and 6) Acetate.

3.2.3 Functional differences across the profile

The 30% most abundant KEGG categories with 60% environmental fit, selected using ‘ordiselect’ of the goveg package included a number of metabolism related genes; such as biosynthesis of amino acids, oxocarboxylic acid metabolism and lipopolysaccharide biosynthesis (**Figure 3.4**). Categories that were more abundant in the saline snow relative to the other horizons included ‘biofilm formation’, ‘bacterial secretion system’, ‘RNA’ degradation’, ‘Bacterial chemotaxis’, ‘two-component system’, ‘folate biosynthesis’, ‘thiamine biosynthesis’ and ‘glutathione metabolism’. Additionally, genes involved in ‘starch and sucrose metabolism’ were enriched in both the algal ice band and saline snow relative to other ice horizons and seawater. Enriched GO categories identified using DeSeq2 in all horizons relative to the seawater were evaluated by overlap analyses (**Figure S3.5**). Saline snow harboured both the largest number of enriched terms (n = 202) and unique enriched terms (n = 103). Unique enriched GO terms included those relating to motility (GO:0009426, GO:0009428, GO:0071977, GO:0071975), pilus assembly (GO:0009297, GO:0043711, GO:0030031), phage shock (GO:0009271), biofilm formation (GO:0044010, GO:1900190), response to UV and anoxia (GO:0034644, GO:0071454) hydrogen peroxide (GO:0042542, GO:0008785), responses to metal ions (GO:0071294, GO:0071292, GO:0010288, GO:0055073) and acid (GO:00712) and sporulation (GO:0030436). There were also a wide range of terms associated with metabolism, particularly mixed acid (mixed-acid) fermentation, amino acid metabolism and various enzymes involved in carbohydrate metabolism.The differentially abundant GO:Terms shared between basal snow and algal ice band also pertain to carbohydrate and amino acid metabolism and those involved in iron acquisition. Those unique to the algal band were involved in processes involving nitrogen compounds (GO:0042126, GO:0042128, GO:0005337, GO:2001057), arabinose (GO:0046373, GO:0019568, GO:0019569, GO:0019572), xylulose (GO:0051167, GO:1901159), pyoverdine (GO:0002048, GO:0002049), phosphonates (GO:0019700, GO:0019634), amylase (GO:0004556, GO:0016160), inosine (GO:0046102) and DNA strand ligase activity (GO:0003910, GO:0000733) and response to nitric acid (GO:0071731). There are 18 GO terms that are uniquely enriched in the surface ice relative to the underlying seawater. These include cellular components relating to viruses (GO:0019038; GO:0019012; GO:0019015; GO:0044423; GO:0019586), leucine transport (GO:0015820) biological processes involved in galacturonate , (GO:0019698, GO:0046396, GO:0046397, GO:0019586), adenine (GO:0008988, GO:0000034, GO:0006146, GO:0052907) and diols (GO:0051143, GO:0034311). There are 6 GO:Terms that are enriched in all ice and snow horizons relative to the underlying seawater; TAM protein secretion complex (GO:0097347), maltodextrin phosphorylase activity (GO: 0031220), glutaminase activity (GO:0004359), glutathionylspermidine amidase activity (GO:0008884), glutathionylspermidine synthase activity (GO:0008885) and glutamine catabolic process (GO:0006543). Finally, there are no significantly enriched GO:Terms in the frazil ice metagenomes relative to the ocean metagenomes.

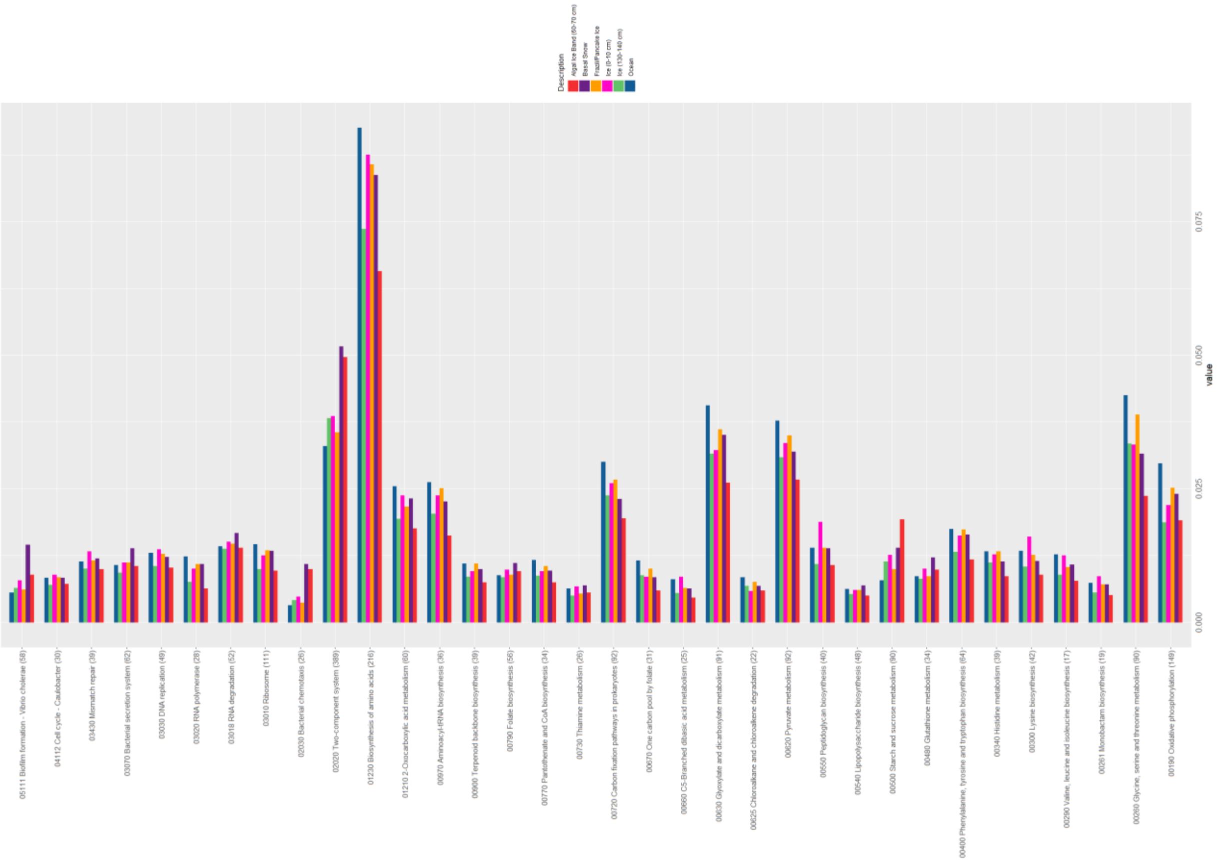


Figure 4: The 30% most abundant KEGG categories with 60% best environmental fit as determined by NMDS ordination on Bray-Curtis distances using ‘ordiselect’ of the goveg package.

3.3 Discussion

The spatial and genomic structuring of microbial communities in snow and ice were investigated to understand the importance of physico-chemical and community pressure on selecting different microbial strains in different compartments of the frozen Arctic seawater. In this chapter, we examined the diversity and range of microbial community structure in snow, ice and seawater.

Bacterial dynamics across the sea ice profile

Selection of bacterial taxa was not evident in the newly formed pancake ice, which mirrored the source seawater and showed no differentially abundant ASVs. Our data suggests that bacterial community selection forms post sea ice consolidation in agreement with some previous work (Collins et al., 2010; Eronen-Rasimus et al., 2014). However, the appearance of specific taxa in the most recently formed (bottom) ice layer still highlights the importance of bacterial-eukaryotic interactions in sea ice and potential for co-transport post consolidation. For example, the genera *Leucothrix* is enriched in the bottom ice horizon and is a widespread epiphyte of marine algae (Brock, 2006).

Not only did the sea ice bacterial community structure differ relative to the underlying seawater, but it was also clearly stratified. This is visible at Class level where Gammaproteobacteria are particularly abundant in the algal ice and basal snow horizons. Gammaproteobacteria are frequently characterized as r-strategists (Weinbauer, Christen, & Höfle, 2006), defined as having high growth rates (r) and are able rapidly colonise unexploited environments with high resource availability (Pianka, 1970). Conversely, K-selected species succeed where bacterial densities are close to carrying capacity (K) by competitive specialism and high affinity for resources and thus tend to dominate mature communities. At class level Flavobacteiia have been identified as K-strategists (Vadstein, Attramadal, Bakke, & Olsen, 2018) and they dominate our surface ice and (to a lesser extent) middle ice communities. Flavobacteriia are highly diverse, and are capable of degrading recalcitrant organics (e.g., cellulose, chitin, and pectin) (Kirchman, 2002). Although taxonomy, especially at class resolution, cannot determine life strategies, it does indicate a potential metabolic difference dependent on the age and physiochemical conditions of the ice. We therefore propose that different life strategies prevail at different ice horizons depending on age and resource availability. Opportunistic, pioneering species are frequent where declined algal blooms provide a diverse array of exudates that can be utilised as metabolic substrates (Fadeev et al., 2018). In contrast, the most mature surface ice layer has likely been depleted of metabolic substrates and in turn selects for K-strategists. Interestingly, the saline snow appears to be a developing community that has not reached its carrying capacity as evidenced by the overwhelming dominance of Gammaproteobacteria. As such, we suggest that substrate diversity and availability is a strong driver of microbial communities of sea ice.

Despite spatially segregated, the algal ice horizon and the basal snow layer are dominated by the genera *Colwellia* and to a lesser extent, *Paraglaciocola* and *Glaciecola* respectively. While few studies have profiled the taxonomy of saline snow overlying sea ice, Maccario, Carpenter, Deming, Vogel, & Larose (2019) also saw a striking enrichment of *Glaciecola* in the saline snow. Further consistency with this study is the presence of halotolerant psychrophiles and psychrotolerant organisms in the sea ice which were less frequent or undetected in the basal snow despite evidence for genomic features beneficial to this environment. Moreover, differential abundance analyses revealed that despite the dominance of *Colwellia* in the basal snow and algal ice band, some Colwellia ASVs were enriched in the ocean, suggesting underlying genomic differences between species or strains which account for their selection. Despite a much larger sequencing depth, the ASV sequences annotated as *Colwellia* consistently showed lower alpha diversity indices relative to all other layers supporting the hypothesis that specific genus variants are strongly selected for in the saline snow. Conversely, the algal ice band harboured a large number of sequence variants and a larger diversity relative to the saline snow which could be explained by the presumed increase in the diversity of niches available in terms of organic substrate abundance and diversity. Given the taxonomic parallelism between the basal snow and algal bloom layer, we hypothesised that the availability of organic substrates and in turn the ability of microorganisms to utilise them is a major driver of community structure. To investigate this hypothesis, as well as the contribution of other environmental drivers, to the perceived enrichment of specific taxa in the algal ice layer and saline basal snow layers, we explored the functional profiles of the ocean, sea ice and snow.

Saline Snow as an interface

Although initially comprised of organisms deposited during snowfall, snow will be subject to post-depositional selection processes (Larose et al., 2010, 2013). Saline snow forms when brine from the underlying sea ice is rejected to the surface delivering salt, nutrients and microorganisms to the snow. A taxa-specific enrichment has been evidenced in saline snow however, it is still not clear whether these organisms are enriched by the process of brine rejection or whether nutrient influx from the brine stimulates the growth of existing snow microbes (Maccario et al., 2019). To investigate the potential contributions of brine and surface snow to the saline snow community we compared the ASV overlap of source (O and SS) and sink (BS) environments. Nearly all the ASVs uniquely shared between the surface and saline snow layers were annotated as the saline snow dominant genera *Colwellia*, *Glaciecola* and *Paraglacieocola*. Although a larger proportion of ASVs were shared between the ocean and saline snow, they were mostly annotated as taxa with low relative abundance in the saline snow. We then considered the overlap between the surface ice and the surface and saline snow layers. A significantly larger proportion of ASVs were shared between the surface ice and the saline snow layers suggesting a larger influence of brine rejection on the saline snow community. However, presence/absence analyses do not take into account the abundances of the genera. The SourceTracker analysis, which accounts for sequence abundance, revealed that the algal ice band had by far the most influence on the saline snow community despite spatial segregation. Given the lack of direct connectivity, we suggest that the parallelism between the algal ice band and the saline snow are governed by similar physiochemical conditions. Moreover, similar taxonomic enrichment patterns have been seen in saline snow without obvious underlying algal bands (Maccario et al., 2019). Therefore, we suggest that while the primary source of saline snow microorganisms is from the brine, there is active selection and likely growth of pioneering microorganisms, although this hypothesis will require further testing.

Functional selection across the sea ice profile

There are a number of specific constraints in the sea ice and saline snow environment that may act as selection mechanisms and in turn shape bacterial communities. By exploring the functional potential of sea ice and saline snow communities, we identified a number of genes with elevated abundance which may be indicative of the underlying selection mechanisms at play. In parallel with the bacterial taxonomic analyses, there were no differentially abundant functional groups in the frazil ice relative to the seawater. Given the metagenomes were dominated by bacterial sequences (~x%), this suggests that selection of bacteria is stochastic during the initial stages of ice formation. By comparing all sea ice and saline snow horizons to the seawater, we were able to identify a small number of terms that were enriched in all ice based environments. Among these, genes associated with the translocation and assembly (TAM) are enriched and involved in the secretion of autotransporters. Autotransporters are often associated with functions such as adhesion, aggregation and biofilm formation which could facilitate attachment and retention in the ice and snow as well as promote survival and growth (T. J. Wells, Tree, Ulett, & Schembri, 2007). There was also a blanket enrichment in the ice and snow layers of genes pertaining to glutathionylspermidine activity, specifically glutathionylspermidine amidase and synthase, which are thought to regulate the redox environment and protect the cell against oxidative damage (Stewart, Dunston, Woster, & Casero, 2018). Although heterogenous across the sea ice matrix, a combination of high oxygen, DOM concentrations, high salinity and periodically high irradiance results in elevated production of reactive oxygen species (ROS) (Chattopadhyay et al., 2011; Hou et al., 2019). Since glutathionylspermidine activity has been identified in polar bacterial genomes (Mocali et al., 2017), it may provide an overall response of sea ice and saline snow microorganisms to oxidative stress.

Aside from the frazil ice, the basal ice had the least number of differentially enriched functional categories relative to the seawater, possibly explained by their direct contact resulting in continued exchange of microorganisms. The surface ice contained the next least number of uniquely enriched functional categories relative to the seawater. Virion components were significantly enriched across the ice environments relative to the seawater but only a logfold change >2 was observed in the upper ice. Viruses are considered a major component of sea ice ecosystems, exhibit high virus:bacteria ratios in brines and are thought to be a key determinant in the control of bacterial populations and nutrient turnover (Collins & Deming, 2011; L. E. Wells & Deming, 2006). Previous metagenomic analysis did not evidence transfer of viruses from sea ice to the overlying snow (Maccario et al., 2019) however, our data highlights that viruses could have a role in this environment.

The basal snow contained the highest number of enriched GO terms relative to the seawater and also shared the most with the algal ice band, suggestive of both similar and unique selective mechanisms across the two horizons. Uniquely enriched genes in the saline snow pertained to stress responses including osmotic stress, UV response and phage shock. Post algal bloom, macromolecules and small molecules (e.g., polysaccharides, proteins, lipids, and other material from cell lysis) are liberated by the decline and lysis of the algae and serve as chemoattractants to capable bacteria (Buchan, LeCleir, Gulvik, & González, 2014). Upward brine expulsion from the ice matrix delivers these substances to the overlying snow and may create a similar chemical environment. These algal derived substrates may provide ecological niches for specific populations and explain the parallelism between the snow and algal ice band horizons. Marine microorganisms utilise two-component systems (TCS) to respond to variation in their physical, chemical and biological surroundings and therefore can mediate microbe-environment interactions (Held, McIlvin, Moran, Laub, & Saito, 2019). Additionally, their abundance in microbial genomes shows a direct association to their lifestyle where a greater occurrence relates to copiotrophy as they permit rapid adaptation to changes in nutrient availability (Held et al., 2019). Although TCS are the most common regulatory systems in bacteria, they are thought to be particularly abundant in marine settings. Interestingly, we saw a “x fold” increase in the number of TCS genes in the ice algal band and basal snow layers relative to any other horizon. This could be an indication of the predominance of copiotrophs and is consistent with a higher abundance of generally copiotrophic *Proteobacteria* in SS and IAB. Coupled with an increase in the abundance of genes related to motility and taxis, we suggest microorganisms abundant in the SS and IAB actively sense and seek out these presumably nutrient rich environments. Gas vacuolated bacteria have been consistently identified in sea ice bacteria (Auman, Breeze, Gosink, Kämpfer, & Staley, 2006; Auman et al., 2010; J. J. Gosink & Staley, 1995; J.J. Gosink, Herwig, & Staley, 1997; John J. Gosink, Woese, & Staley, 1998; Irgens, Gosink, & Staley, 1996) and could play a role in the stratification of communities across sea ice horizons. Indeed, we identified a significant enrichment in genes related to gas vesicles in saline snow metagenomes which may provide an additional strategy to colonising saline snow.

Outside of algal bloom stages, sea ice is dominated by heterotrophic processes which appears to extend to the overlying snow (Maccario et al., 2019). Despite the low taxonomic diversity in our basal snow samples, we observed an increase in the gene abundance for carbohydrate metabolic processes relative to all other horizons except the algal ice band layer for which numbers were approximately equivalent. These included genes involved in starch, sucrose, cellulose and chitin metabolism reflecting the potential for metabolic versatility in saline snow microorganisms. Although genome sizes inferred from our metagenomic dataset were similar across the profiles, larger genomes sizes were previously shown to correlate with the abundance of genes related to metabolism, biosynthesis and two-component systems; all of which were more abundant in our snow metagenomes (Nayfach & Pollard, 2015). Considering the low taxonomic diversity, we suggest that the saline snow is a highly selective environment where microorganisms that can sense and utilise diverse metabolic substrates flourish. Coupled with the higher relative abundance of genes related to genetic transfer, broad substrate specificity, detection and motility and genome flexibility may contribute to colonisation of saline snow environments and post-algal blooms entrapped in the ice.

Chemical drivers in the profile

A significant co-structure based on a Monte Carlo test with 10,000 permutations (p=0.001) linked to the ocean-sea ice-snow profile chemistry with the microbial community structure. Algal ice and saline snow were similarly structured and appear to be driven by inorganic ions. Both the algal ice band and the saline snow were dominated by Gammaproteobacteria, specifically *Colwellia*, which had a large effect on the ordination. Although differently clustered, the ocean and frazil ice samples were similarly driven by inorganic ion content. Acetate had the strongest effect on the ordination and was linked to the bottom ice layer and to a lesser extent the middle ice layer. Although acetate concentrations were relatively high in the bottom ice layer, we were unable to identify any acetic acid bacteria (AAB) in either the bottom or middle ice layers. Although we did not measure EPS in our samples, EPS is known to concentrate during sea ice formation where the highest concentrations are typically found in the lowest ice portions (Underwood et al., 2013). As EPS generally consists of monosaccharides and some non-carbohydrate substituents, such as acetate, we hypothesise that the high concentration of acetate in our IB samples are tightly coupled to EPS (Lindberg, 1990). Nitrate appears to be the major chemical driver measured in relation to the community structure of surface snow which is consistent with past studies that documented an impact of nitrate addition on snow bacterial community composition (Holland et al., 2017). Genera driving the ordination of surface snow samples include *Sphingomonas*, *Corynebacterium*, *Clostridium*; species known to be capable of nitrogen fixation as well as *Staphylococcus* which has been show to increase nitrogen fixation potential in some bacteria (Berndt et al., 1976; Chen, 2006; Holguin, Guzman, & Bashan, 1992; Videira, de Araujo, da Silva Rodrigues, Baldani, & Baldani, 2009). As well as nitrate and nitrite reducing bacteria such as Rhodopila and Micrococcus (Imhoff, Rahn, Künzel, & Neulinger, 2018; Kocur, Kloos, & SCHLEIFER, 2006). While pH is known to be a major factor in community structure of soils, its effect on sea ice communities appears to be minimal despite quite a large variation (pH 5.53 – 7.89) (Lauber, Hamady, Knight, & Fierer, 2009). It should be emphasised that other environmental variables not measured in this study may participate in community dynamics however, overall the OSIS community structure appears to be differentially effected by chemical parameters.

Eukaryotic dynamics across the sea ice profile

The taxonomic affiliation of eukaryotic sequences associated to the surface snow was strikingly different from the underlying saline snow, sea ice and seawater. Surface snow was characterised by a dominance of fungal sequences, reflective of the atmospheric samples. Therefore, surface snow is largely shaped by wind-dispersed spores; as has been suggested in other studies of High Arctic snowpacks (Harding, Jungblut, Lovejoy, & Vincent, 2011). Conversely, the saline snow is dominated by marine protists, particularly *Protapsis*, suggesting transfer from the sea ice via brine rejection. *Protapsis* species are Cercozoan flagellates which are a ubiquitous group across Antarctic and Arctic sea ice containing members particularly well adapted for survival and growth at extreme temperatures (Caron & Gast, 2010). Research suggests that *Protapsis* are dominant at the initial stages of declining blooms but are rapidly replaced in the latter stages (Berdjeb, Parada, Needham, & Fuhrman, 2018); this would be consistent with their minimal abundance in the algal band layer (which is in post-bloom). Although diatom abundances are comparatively small relative to *Protapsis*, there are a number of diatom genera that make up >2% abundance in at least one of the basal snow samples (Carteria, Chlamydomonas and Pyrarimonas), which could support a transient *Protapsis* population. Alternatively, the high abundances of these heterotophic protists provides circumstantial evidence that significant bacterivory is occurring. Although grazing in snow is yet to be addressed, studies suggest that defence against grazing is an important survival strategy (Jurgens' & Gude2, 1994; Ballen-Segura, Felip, & Catalan, 2017; Sanders & Gast, 2012) and in turn, may provide a major constraint on bacterial communities of the saline snow layer and lack of diversity therein.

Surprisingly, the prevalence of diatoms was extremely low across the profile. The lack of diatom abundance or frequency contrasts other studies of High Arctic sea ice (e.g. Harding et al., 2011; Lovejoy et al., 2007) but reflects a taxonomic survey of the sea ice and seawater eukaryotes at the end of the polar night (Bachy, López-García, Vereshchaka, & Moreira, 2011). However, our sampling was undertaken during long periods of light where evidence of a previous diatom bloom was entrapped in the ice. As such, we suggest a boom and bust cyclic dynamic with regards to diatoms, where blooms result in the dominance of diatoms often evidenced in the literature. However, even during the light season, decline of diatom blooms supports populations of heterotrophic dinoflagellates and parasitic populations which completely eliminate diatom populations; as seen in this study. This is supported by the abundance of genera related to archetypal diatom predators such as *Cryothecomonas* and *Gyrodinium* which dominate across the sea ice and seawater profile (Sherr & Sherr, 2007). Despite the lack of diatom related sequences in our study, we were able to identify a stratification in the eukaryotic community assemblage suggesting selection of eukaryotes is different at varying horizons. For instance, in agreement with previous research, selection of eukaryotes appears to occur at the initial stages of ice formation as evidenced by the drastic differences in community composition between seawater and newly formed frazil ice (Gradinger & Ikävalko, 1998; Rózańska, Poulin, & Gosselin, 2008). We identified a significant loss of abundance in ASVs associated with *Gyrodinium* and an enrichment in *Polarella* and *Cryothecomonas* during the transition from seawater to frazil ice. Strains belonging to the former are known ice nucleators which could explain their initial incorporation (Tesson & Šantl-Temkiv, 2018). The bottom ice layer showed some resemblance to the underlying ocean in that it contained a significant relative abundance of *Gyrodinium*, a dinoflagellate frequent of Arctic waters, suggesting continued colonisation post-ice crystal formation (Kubiszyn & Wiktor, 2016). The dominant heterotrophic flagellate in the seawater and bottom layer were replaced by other heterotophic flagellates in the ice algal layer. For instance, *Ebria* and *Oblongichytrium* belonging to the family Thraustochrytids were enriched and are known to display a negative relationship with phytoplankton cells but a positive one with their exudates and detritus; reflective of post-bloom dynamics (Damare et al., 2020).

3.4 Conclusion

We identified potential drivers of microbial selection in the increasingly important first-year ice (FYI) and saline snow (SS) environments. Brine rejection appears to seed basal snow environments with both nutrients and microorganisms, yet, there is a striking taxa-specific enrichment of bacteria that results in a saline snow specific community. Based on previous research by Maccario et al. (2019) and within this thesis (Chapter 3), this phenomenon appears to be independent of geography and as such is likely due to corresponding environmental constraints. Within this study, parallels could be drawn between the bacterial communities of saline snow and post-bloom algae trapped in the ice suggesting similar structuring variables. Through taxonomic and metagenomic analyses, selective forces may be linked to the ability to detect, move toward and utilise organic substrates. The lack of diatom sequences and high abundance of potential eukaryotic grazers may provide an additional constraint that is seen across the entirety of the profile. Preliminary analyses of sequence variation amongst abundant saline snow taxa suggests selection at a minimal resolution of genus and possibly strain. Given the strain diversity of saline snow dominant organisms and their stratification by presence and abundance across the horizons, we hypothesised that these are pioneering versatile species that occupy diverse niches. While they may share core genes related to primary colonisation, strain-variable genes will reflect the adaptations to the habitats in which the strains are abundantly present and to their affinity for different organic substances. Further work utilising genome-genome comparison analysis and identification of environmental constraints is required to refine the Arctic ice and snow strain selection.

3.5 Supplementary figures

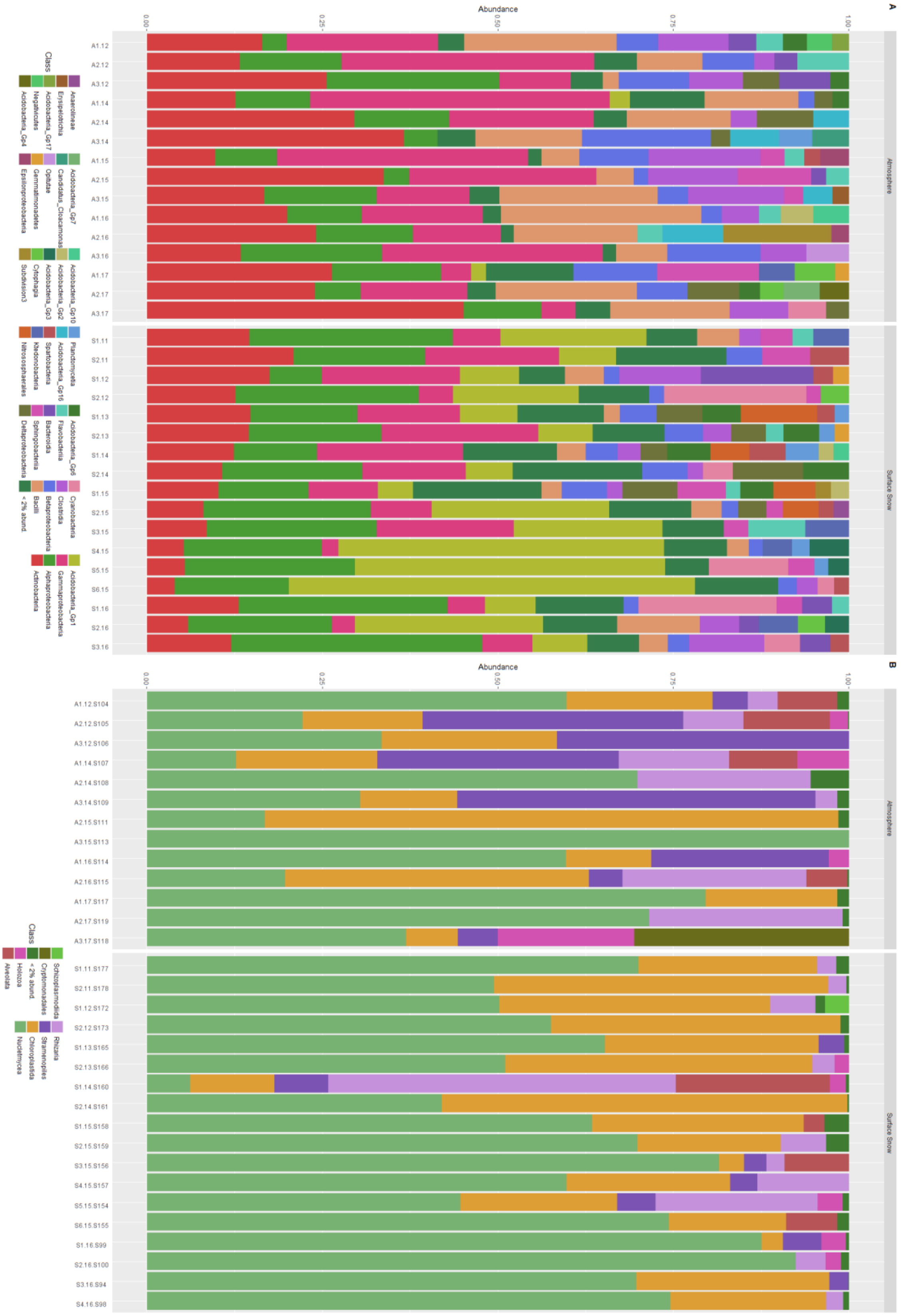


Figure S3.1: Histograms of the relative abundance of classes representing >2% abundance in the (A) 16S rRNA and (B) 18S rRNA datasets of the atmosphere and surface snow samples

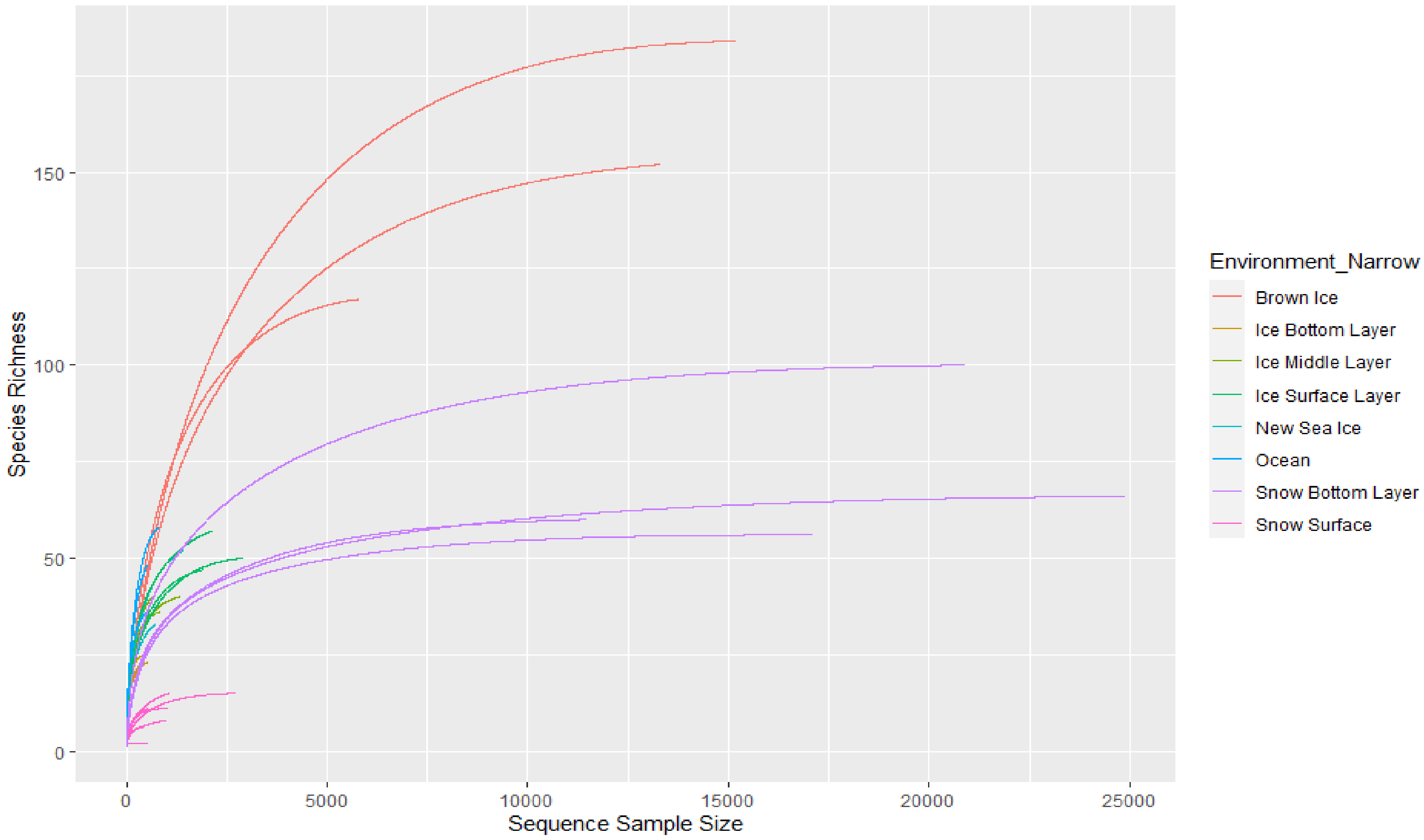


Figure S3.2: Rarefaction curves of 16S rRNA ASVs annotated as the genera *Colwellia*.

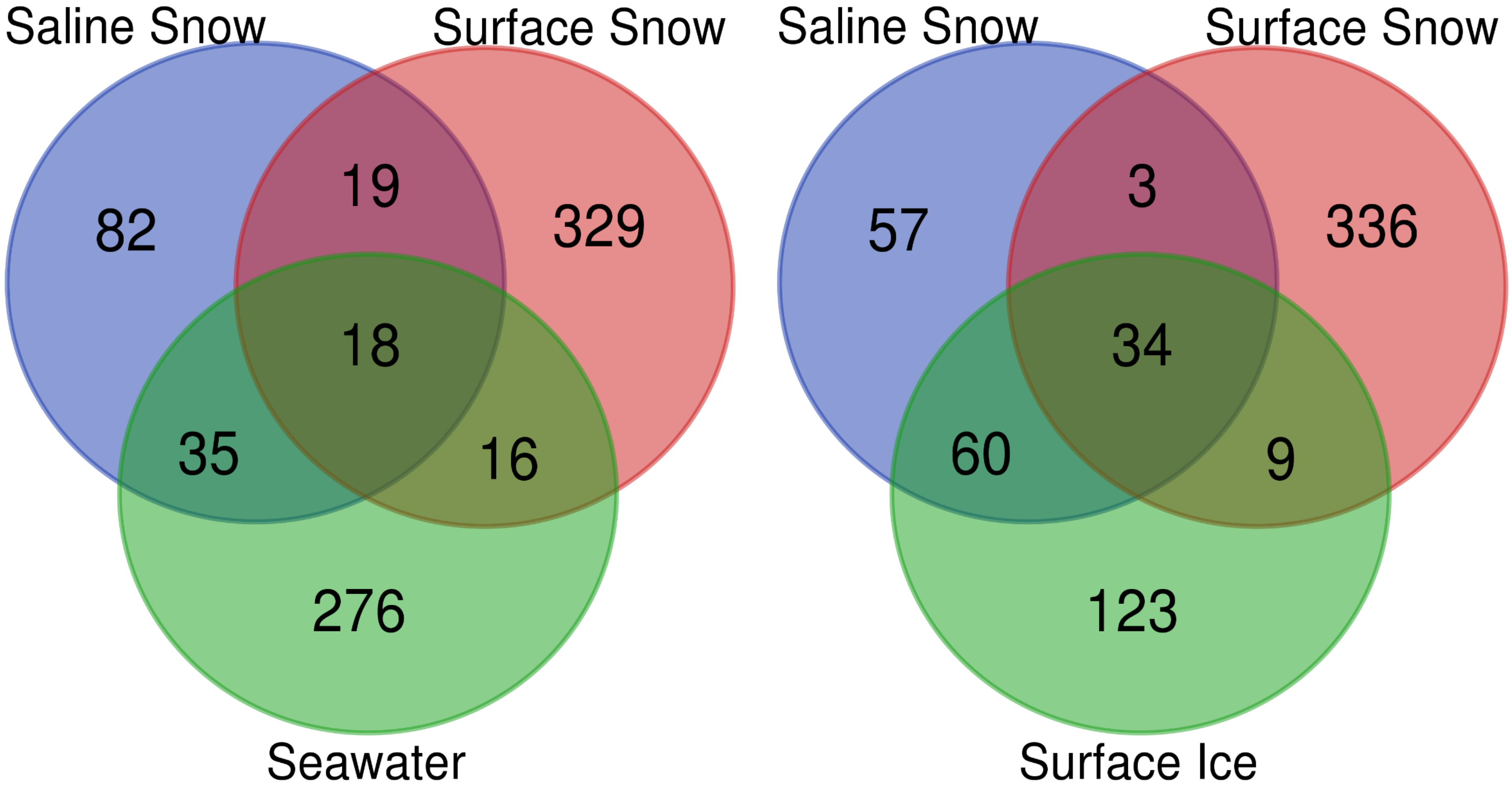


Figure S3.3: Overlap of 16S rRNA ASV's that appear in at least 2 replicates ad have a count of at least 2.

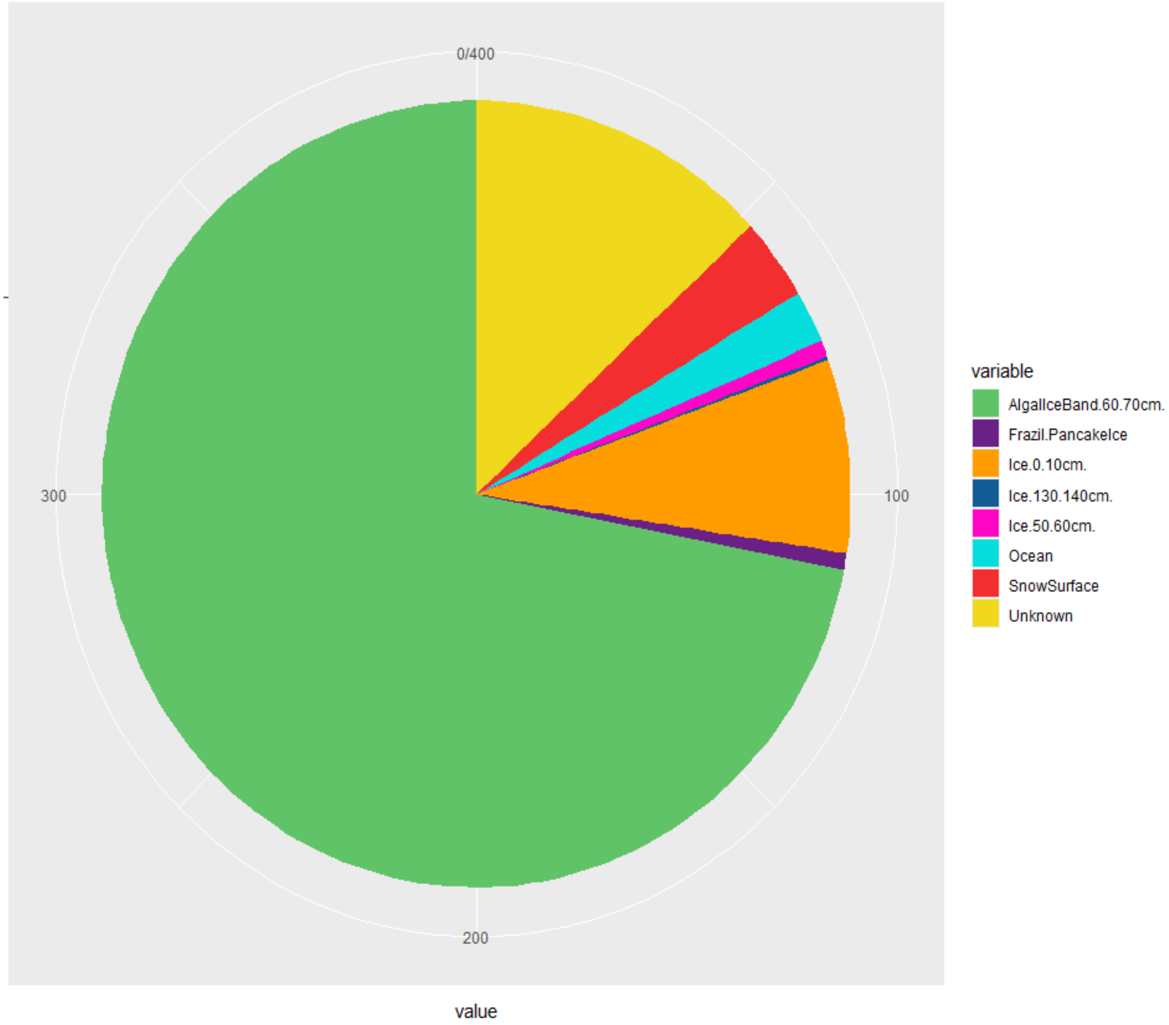


Figure S3.4: The source proportions using SourceTracker estimations of source contribution (seawater, sea ice and surface snow) to the saline snow horizon

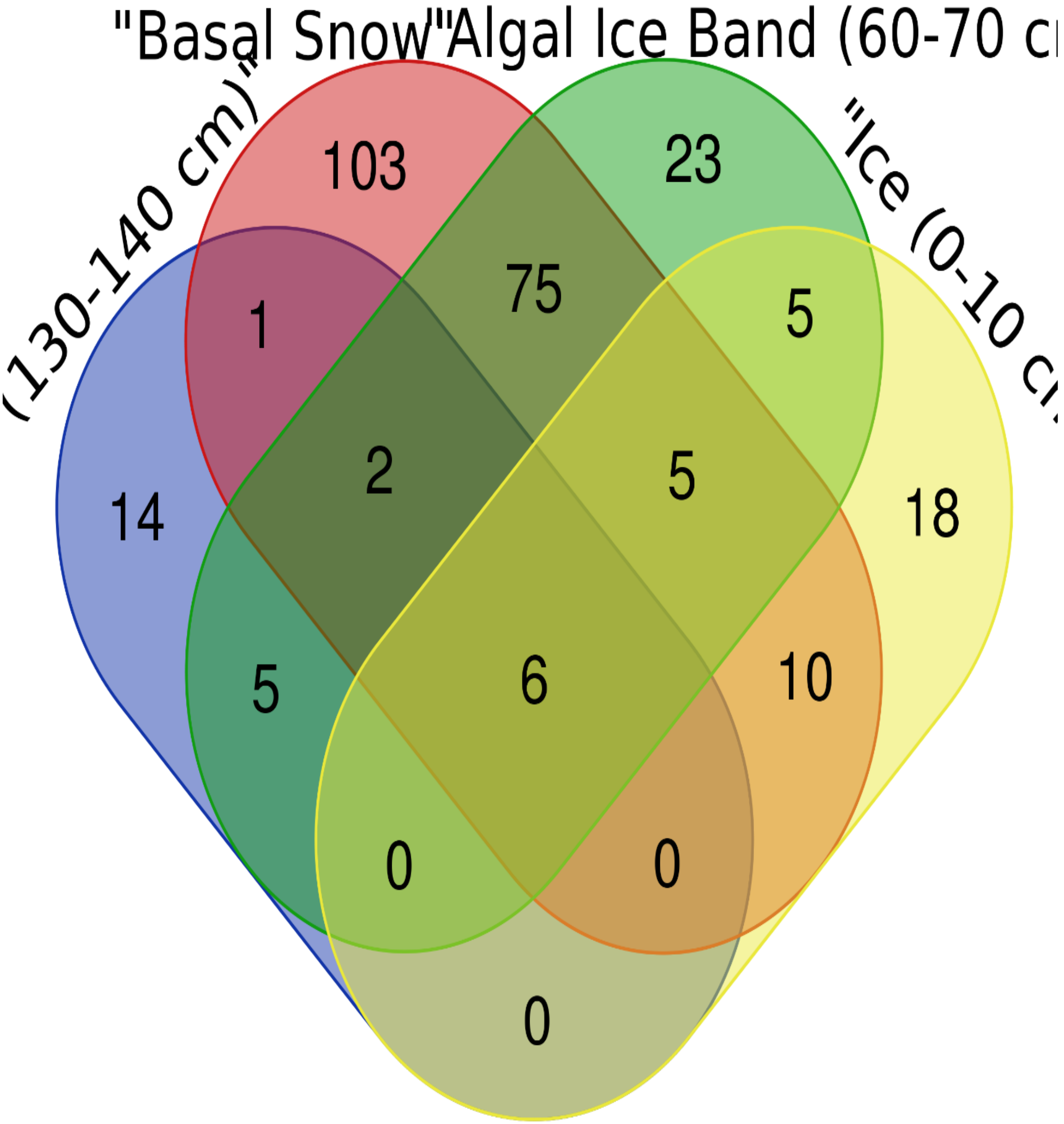


Figure S3.5: Venn diagrams demonstrating differentially abundant GO term overlap of ice and snow layers relative to the seawater. Numbers correspond to unique GO terms within a subset.

3.6 References

~urgens’, K., & Gude2, H. (1994). The potential importance of grazing-resistant bacteria in planktonic systems. In *MARINE ECOLOGY PROGRESS SERIES Mar. Ecol. Prog. Ser* (Vol. 112).

Auman, A. J., Breeze, J. L., Gosink, J. J., Kämpfer, P., & Staley, J. T. (2006). Psychromonas ingrahamii sp. nov., a novel gas vacuolate, psychrophilic bacterium isolated from Arctic polar sea ice. *International Journal of Systematic and Evolutionary Microbiology*, *56*(5), 1001–1007. https://doi.org/10.1099/ijs.0.64068-0

Auman, A. J., Breezee, J. L., Gosink, J. J., Schumann, P., Barnes, C. R., Kämpfer, P., & Staley, J. T. (2010). Psychromonas boydii sp. nov., a gas-vacuolate, psychrophilic bacterium isolated from an Arctic sea-ice core. *International Journal of Systematic and Evolutionary Microbiology*, *60*(1), 84–92. https://doi.org/10.1099/ijs.0.007773-0

Bachy, C., López-García, P., Vereshchaka, A., & Moreira, D. (2011). Diversity and Vertical Distribution of Microbial Eukaryotes in the Snow, Sea Ice and Seawater Near the North Pole at the End of the Polar Night. *Frontiers in Microbiology*, *2*(MAY), 106. https://doi.org/10.3389/fmicb.2011.00106

Ballen-Segura, M., Felip, M., & Catalan, J. (2017). Some mixotrophic flagellate species selectively graze on archaea. *Applied and Environmental Microbiology*, *83*(2). https://doi.org/10.1128/AEM.02317-16

Berdjeb, L., Parada, A., Needham, D. M., & Fuhrman, J. A. (2018). Short-term dynamics and interactions of marine protist communities during the spring-summer transition. *ISME Journal*, *12*(8), 1907–1917. https://doi.org/10.1038/s41396-018-0097-x

Berndt, H., Ostwal, K. P., Lalucat, J., Schumann, C., Mayer, F., & Schlegel, H. G. (1976). Identification and physiological characterization of the nitrogen fixing bacterium Corynebacterium autotrophicum GZ 29. *Archives of Microbiology*, *108*(1), 17–26. https://doi.org/10.1007/BF00425088

Bowman, J. S., Berthiaume, C. T., Armbrust, E. V., & Deming, J. W. (2014). The genetic potential for key biogeochemical processes in Arctic frost flowers and young sea ice revealed by metagenomic analysis. *FEMS Microbiology Ecology*, *89*(2), 376–387.

https://doi.org/10.1111/1574-6941.12331

Bowman, J. S., & Deming, J. W. (2014). Alkane hydroxylase genes in psychrophile genomes and the potential for cold active catalysis. *BMC Genomics*, *15*(1), 1120. https://doi.org/10.1186/1471-2164-15-1120

Brinkmeyer, R., Knittel, K., Jurgens, J., Weyland, H., Amann, R., & Helmke, E. (2003). Diversity and structure of bacterial communities in Arctic versus Antarctic pack ice. *Applied and Environmental Microbiology*, *69*(11), 6610–6619.

Brock, T. D. (2006). The Genus Leucothrix. In *The Prokaryotes* (pp. 931–938). https://doi.org/10.1007/0-387-30746-x_36

Buchan, A., LeCleir, G. R., Gulvik, C. A., & González, J. M. (2014, August 19). Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nature Reviews. Microbiology*, Vol. 12, pp. 686–698. https://doi.org/10.1038/nrmicro3326

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. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Caron, D. A., & Gast, R. J. (2010). Heterotrophic Protists Associated with Sea Ice. In *Sea Ice: Second Edition* (pp. 327–356). https://doi.org/10.1002/9781444317145.ch9

Chattopadhyay, M. K., Raghu, G., Sharma, Y. V. R. K., Biju, A. R., Rajasekharan, M. V., & Shivaji, S. (2011). Increase in oxidative stress at low temperature in an antarctic Bacterium. *Current Microbiology*, *62*(2), 544–546. https://doi.org/10.1007/s00284-010-9742-y

Chen, J.-S. (2006). Nitrogen Fixation in the Clostridia. In *Genetics and Regulation of Nitrogen Fixation in Free-Living Bacteria* (pp. 53–64). https://doi.org/10.1007/1-4020-2179-8_3

Cirri, E., & Pohnert, G. (2019). Algae−bacteria interactions that balance the planktonic microbiome. *New Phytologist*, *223*(1), 100–106. https://doi.org/10.1111/nph.15765

Collins, R. E., & Deming, J. W. (2011). Abundant dissolved genetic material in Arctic sea ice Part II: Viral dynamics during autumn freeze-up. *Polar Biology*, *34*(12), 1831–1841. https://doi.org/10.1007/s00300-011-1008-z

Damare, V. S., D'Costa, P. M., Shivaramu, M. S., Borges, V., Fernandes, M., Fernandes, C., & Cardozo, S. (2020). Preliminary study on the response of marine fungoid protists, the thraustochytrids, to lipid extracts of diatoms. *Aquatic Ecology*, *54*(1), 355–367. https://doi.org/10.1007/s10452-020-09747-z

Eric Collins, R., Rocap, G., & Deming, J. W. (2010). Persistence of bacterial and archaeal communities in sea ice through an Arctic winter. *Environmental Microbiology*, *12*(7), 1828–1841. https://doi.org/10.1111/j.1462-2920.2010.02179.x

Eronen-Rasimus, E., Kaartokallio, H., Lyra, C., Autio, R., Kuosa, H., Dieckmann, G. S., & Thomas, D. N. (2014). Bacterial community dynamics and activity in relation to dissolved organic matter availability during sea-ice formation in a mesocosm experiment. *MicrobiologyOpen*, *3*(1), 139–156. https://doi.org/10.1002/mbo3.157

Fadeev, E., Salter, I., Schourup-Kristensen, V., Nöthig, E. M., Metfies, K., Engel, A., … Bienhold, C. (2018). Microbial communities in the east and west fram strait during sea ice melting season. *Frontiers in Marine Science*, *5*(NOV), 429. https://doi.org/10.3389/fmars.2018.00429

Gosink, J. J., & Staley, J. T. (1995). Biodiversity of gas vacuolate bacteria from antarctic sea ice and water. *Applied and Environmental Microbiology*, *61*(9), 3486–3489. https://doi.org/10.1128/aem.61.9.3486-3489.1995

Gosink, J.J., Herwig, R. P., & Staley, J. T. (1997). Octadecabacter arcticus gen. nov., sp. nov., and O. antarcticus, sp. nov., Nonpigmented, Psychrophilic Gas Vacuolate Bacteria from Polar Sea Ice and Water. *Systematic and Applied Microbiology*, *20*(3), 356–365. https://doi.org/10.1016/S0723-2020(97)80003-3

Gosink, John J., Woese, C. R., & Staley, J. T. (1998). Polaribacter gen. nov., with three new species, P. irgensii sp. nov., P. franzmannii sp. nov. and P. filamentus sp. nov., gas vacuolate polar marine bacteria of the Cytophaga-Flavobacterium-Bacteroides group and reclassification of "Flectobacillus glomeratus" as Polaribacter glomeratus comb. nov. *International Journal of Systematic Bacteriology*, *48*(1), 223–235. https://doi.org/10.1099/00207713-48-1-223

Gradinger, R., & Ikävalko, J. (1998). Organism incorporation into newly forming Arctic sea ice in the Greenland Sea. *Journal of Plankton Research*, *20*(5), 871–886. https://doi.org/10.1093/plankt/20.5.871

Harding, T., Jungblut, A. D., Lovejoy, C., & Vincent, W. F. (2011). Microbes in high arctic snow and implications for the cold biosphere. *Applied and Environmental Microbiology*, *77*(10), 3234–3243. https://doi.org/10.1128/AEM.02611-10

Hatam, I., Lange, B., Beckers, J., Haas, C., & Lanoil, B. (2016). Bacterial communities from Arctic seasonal sea ice are more compositionally variable than those from multi-year sea ice. *ISME Journal*, *10*(10), 2543–2552. https://doi.org/10.1038/ismej.2016.4

Held, N. A., McIlvin, M. R., Moran, D. M., Laub, M. T., & Saito, M. A. (2019). Unique Patterns and Biogeochemical Relevance of Two-Component Sensing in Marine Bacteria. *MSystems*, *4*(1). https://doi.org/10.1128/msystems.00317-18

Holguin, G., Guzman, M. A., & Bashan, Y. (1992). Two new nitrogen-fixing bacteria from the rhizosphere of mangrove trees: Their isolation, identification and in vitro interaction with rhizosphere Staphylococcus sp. *FEMS Microbiology Letters*, *101*(3), 207–216. https://doi.org/10.1016/0378-1097(92)90817-8

Holland, A. T., Pinto, B. B., Layton, R., Williamson, C. J., Anesio, A. M., Vogel, T. M., … Tranter, M. (2017). *Over winter microbial processes in a Svalbard snow pack : an experimental approach*. (2012), 1–14.

Hou, Y., Qiao, C., Wang, Y., Wang, Y., Ren, X., Wei, Q., & Wang, Q. (2019). Cold-adapted glutathione S-transferases from Antarctic psychrophilic bacterium Halomonas sp. ANT108: Heterologous expression, characterization, and oxidative resistance. *Marine Drugs*, *17*(3). https://doi.org/10.3390/md17030147

Imhoff, J. F., Rahn, T., Künzel, S., & Neulinger, S. C. (2018). New insights into the metabolic potential of the phototrophic purple bacterium Rhodopila globiformis DSM 161T from its draft genome sequence and evidence for a vanadium-dependent nitrogenase. *Archives of Microbiology*, *200*(6), 847–857. https://doi.org/10.1007/s00203-018-1489-z

Irgens, R. L., Gosink, J. J., & Staley, J. T. (1996). Polaromonas vacuolata gen. nov., sp. nov., a psychrophilic, marine, gas vacuolate bacterium from Antarctica. *International Journal of Systematic Bacteriology*, *46*(3), 822–826. https://doi.org/10.1099/00207713-46-3-822

Junge, K., Cameron, K., & Nunn, B. (2019). Diversity of Psychrophilic Bacteria in Sea and Glacier Ice Environments—Insights Through Genomics, Metagenomics, and Proteomics Approaches. In *Microbial Diversity in the Genomic Era* (pp. 197–216). https://doi.org/10.1016/b978-0-12-814849-5.00012-5

Kirchman, D. L. (2002). The ecology of Cytophagaâ€“Flavobacteria in aquatic environments. *FEMS Microbiology Ecology*, *39*(2), 91–100. https://doi.org/10.1111/j.1574-6941.2002.tb00910.x

Knights, D., Kuczynski, J., Charlson, E. S., Zaneveld, J., Mozer, M. C., Collman, R. G., … Kelley, S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature Methods*, *8*(9), 761–765. https://doi.org/10.1038/nmeth.1650

Kocur, M., Kloos, W. E., & SCHLEIFER, K.-H. (2006). The Genus Micrococcus. In *The Prokaryotes* (pp. 961–971). https://doi.org/10.1007/0-387-30743-5_37

Kubiszyn, A. M., & Wiktor, J. M. (2016). The Gymnodinium and Gyrodinium (Dinoflagellata: Gymnodiniaceae) of the West Spitsbergen waters (1999–2010): biodiversity and morphological description of unidentified species. *Polar Biology*, *39*(10), 1739–1747. https://doi.org/10.1007/s00300-015-1764-2

Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D., & Vogel, T. M. (2010). Microbial sequences retrieved from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway. *Extremophiles*, *14*(2), 205–212. https://doi.org/10.1007/s00792-009-0299-2

Larose, C., Prestat, E., Cecillon, S., Berger, S., Malandain, C., Lyon, D., … Vogel, T. M. (2013). Interactions between snow chemistry, mercury inputs and microbial population dynamics in an arctic snowpack. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079972

Lauber, C. L., Hamady, M., Knight, R., & Fierer, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and Environmental Microbiology*, *75*(15), 5111–5120. https://doi.org/10.1128/AEM.00335-09

Lindberg, B. (1990). Components of Bacterial Polysaccharides. *Advances in Carbohydrate*

Chemistry and Biochemistry, *48*(C), 279–318. https://doi.org/10.1016/S0065-2318(08)60033-5

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

Lovejoy, C., Vincent, W. F., Bonilla, S., Roy, S., Martineau, M.-J., Terrado, R., … Pedrós-Alió, C. (2007). DISTRIBUTION, PHYLOGENY, AND GROWTH OF COLD-ADAPTED PICOPRASINOPHYTES IN ARCTIC SEAS. *Journal of Phycology*, *43*(1), 78–89. https://doi.org/10.1111/j.1529-8817.2006.00310.x

Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M., & Larose, C. (2019). Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-38744-y

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Mocali, S., Chiellini, C., Fabiani, A., Decuzzi, S., Pascale, D., Parrilli, E., … Fani, R. (2017). Ecology of cold environments: New insights of bacterial metabolic adaptation through an integrated genomic-phenomic approach. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-00876-4

Nayfach, S., & Pollard, K. S. (2015). Average genome size estimation improves comparative metagenomics and sheds light on the functional ecology of the human microbiome. *Genome Biology*, *16*(1), 51. https://doi.org/10.1186/s13059-015-0611-7

Package "goeveg" Type Package Title Functions for Community Data and Ordinations. (2018).

Pianka, E. R. (1970). On r- and K-Selection. *The American Naturalist*, *104*(940), 592–597. https://doi.org/10.1086/282697

Ramanan, R., Kim, B. H., Cho, D. H., Oh, H. M., & Kim, H. S. (2016, January 1). Algae-bacteria interactions: Evolution, ecology and emerging applications. *Biotechnology Advances*, Vol. 34, pp. 14–29. https://doi.org/10.1016/j.biotechadv.2015.12.003

Rózańska, M., Poulin, M., & Gosselin, M. (2008). Protist entrapment in newly formed sea ice in the Coastal Arctic Ocean. *Journal of Marine Systems*, *74*(3–4), 887–901. https://doi.org/10.1016/j.jmarsys.2007.11.009

Sanders, R. W., & Gast, R. J. (2012). Bacterivory by phototrophic picoplankton and nanoplankton in Arctic waters. *FEMS Microbiology Ecology*, *82*(2), 242–253. https://doi.org/10.1111/j.1574-6941.2011.01253.x

Sherr, E., & Sherr, B. (2007). Heterotrophic dinoflagellates: a significant component of microzooplankton biomass and major grazers of diatoms in the sea. *Marine Ecology Progress Series*, *352*, 187–197. https://doi.org/10.3354/meps07161

Stewart, T. M., Dunston, T. T., Woster, P. M., & Casero, R. A. (2018, November 30). Polyamine catabolism and oxidative damage. *Journal of Biological Chemistry*, Vol. 293, pp. 18736–18745. https://doi.org/10.1074/jbc.TM118.003337

Tesson, S. V. M., & Šantl-Temkiv, T. (2018). Ice Nucleation Activity and Aeolian Dispersal Success in Airborne and Aquatic Microalgae. *Frontiers in Microbiology*, *9*(NOV), 2681. https://doi.org/10.3389/fmicb.2018.02681

Underwood, G. J. C., Aslam, S. N., Michel, C., Niemi, A., Norman, L., Meiners, K. M., … Thoma, D. N. (2013). Broad-scale predictability of carbohydrates and exopolymers in Antarctic and Arctic sea ice. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(39), 15734–15739. https://doi.org/10.1073/pnas.1302870110

Vadstein, O., Attramadal, K. J. K., Bakke, I., & Olsen, Y. (2018, November 14). K-selection as microbial community management strategy: A method for improved viability of larvae in aquaculture. *Frontiers in Microbiology*, Vol. 9. https://doi.org/10.3389/fmicb.2018.02730

Videira, S. S., de Araujo, J. L. S., da Silva Rodrigues, L., Baldani, V. L. D., & Baldani, J. I. (2009). Occurrence and diversity of nitrogen-fixing *Sphingomonas* bacteria associated with rice plants grown in Brazil. *FEMS Microbiology Letters*, *293*(1), 11–19.

https://doi.org/10.1111/j.1574-6968.2008.01475.x

Vihma, T. (2014). Effects of Arctic Sea Ice Decline on Weather and Climate: A Review. *Surveys in Geophysics*, *35*(5), 1175–1214. https://doi.org/10.1007/s10712-014-9284-0

Vollmers, J., Voget, S., Dietrich, S., Gollnow, K., Smits, M., Meyer, K., … Daniel, R. (2013). Poles apart: Arctic and Antarctic Octadecabacter strains share high genome plasticity and a new type of xanthorhodopsin. *PloS One*, *8*(5), e63422. https://doi.org/10.1371/journal.pone.0063422

Wang, Y., Hou, Y., Wang, Y., Zheng, L., Xu, X., Pan, K., … Wang, Q. (2018). A novel cold-adapted leucine dehydrogenase from antarctic sea-ice bacterium pseudoalteromonas sp. ANT178. *Marine Drugs*, *16*(10). https://doi.org/10.3390/md16100359

Weinbauer, M. G., Christen, R., & Höfle, M. G. (2006). The response of Vibrio- and Rhodobacter-related populations of the NW Mediterranean Sea to additions of dissolved organic matter, phages, or dilution. *Microbial Ecology*, *51*(3), 336–344. https://doi.org/10.1007/s00248-006-9028-x

Wells, L. E., & Deming, J. W. (2006). Modelled and measured dynamics of viruses in Arctic winter sea-ice brines. *Environmental Microbiology*, *8*(6), 1115–1121. https://doi.org/10.1111/j.1462-2920.2006.00984.x

Wells, T. J., Tree, J. J., Ulett, G. C., & Schembri, M. A. (2007). Autotransporter proteins: novel targets at the bacterial cell surface. *FEMS Microbiology Letters*, *274*(2), 163–172. https://doi.org/10.1111/j.1574-6968.2007.00833.x

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Retrieved from https://ggplot2.tidyverse.org

Yergeau, E., Michel, C., Tremblay, J., Niemi, A., King, T. L., Wyglinski, J., … Greer, C. W. (2017). Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Scientific Reports*, *7*(1), 42242. https://doi.org/10.1038/srep42242

CHAPTER 4: Taxonomic and functional differences of microbial communities of first year ice and multiyear ice and their influence on the post-depositional selection in snow

Layton R, Dommergue A, Jaffrezo J.L, Keuschnig, C, Vogel T.M and Larose, C

4.0 Introduction

Sea ice has a fundamental role in polar ecosystems and its melting has major consequences on the global climate. Sea ice can be classified into two types; (1) first year ice (FYI) which is defined by a growth season of no more than one year and (2) multiyear ice (MYI) which has survived at least one melt season (Vonnahme, Dietrich, & Hassett, 2020; Weeks & Ackley, 1986). The transition between a dominance of MYI in the arctic to FYI has accelerated, resulting in an ice cover decline of approximately 70% to 30% in the last three decades (Maslanik, Stroeve, Fowler, & Emery, 2011). Predicting the consequences of this continued loss of MYI will require a fundamental understanding of the biological and physiochemical differences between MYI and FYI; an opportunity which is rapidly disappearing.

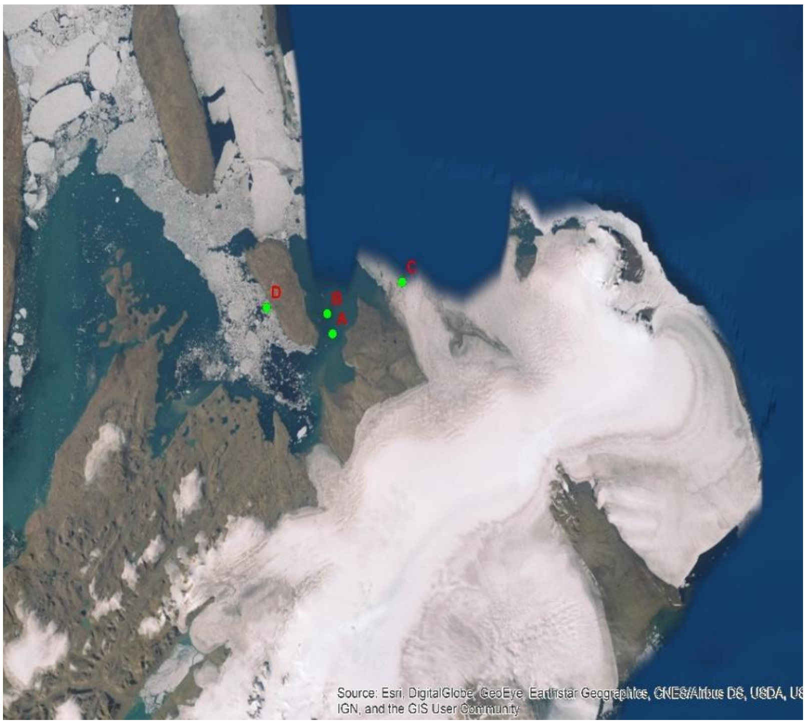
Arctic sea ice serves as an important habitat for a vast diversity of microorganisms that actively contribute to the biogeochemical cycling of oceans (Bowman, 2015)**.** The physiochemical gradients that characterise FYI and MYI are distinct and are thought to translate to differences in microbial assembly (Hatam, Lange, Beckers, Haas, & Lanoil, 2016). For instance, the relative thinness of FYI results in greater light exposure and steeper gradients of temperature, salinity and nutrient concentrations (Vancoppenolle *et al.*, 2013; Weeks & Ackley, 1986). FYI is permeated by liquid brine channels that fluctuate in volume as a function of temperature and harbour most of the microbial biomass. Conversely, MYI is mostly devoid of brine due to the process of gravitational drainage and snowmelt percolation (Vancoppenolle *et al.*, 2013). The few comparative studies of the microbial communities of FYI and MYI have highlighted differences in taxonomic composition where the latter represents more stable but less diverse communities (Hatam *et al.*, 2016). Stratification of taxonomic assemblages reflective of the distinct horizons in sea ice are also evident. For instance, Hatam *et al.* (2014) demonstrated that the bacterial communities represented the surface refrozen meltpond ice, > 2 year old middle ice or newly grown bottom ice (Hatam *et al.*, 2014).

Although a handful of studies have highlighted taxonomic incongruity between MYI and FYI, the differences in functional potential have yet to be identified. Moreover, much emphasis has been placed on the biologically productive sea ice bottom horizon that plays host to the spring and summer diatom blooms (*e.g.* Arrigo *et al.*, 2014; Assmy *et al.*, 2017; Yergeau *et al.*, 2017). However, the surface horizons of sea ice harbour a diverse community of heterotrophic bacteria that have a key role in the oceanographic microbial loop (Gradinger & Zhang, 1997; Rysgaard & Glud, 2004; Rysgaard, Glud, Sejr, Blicher, & Stahl, 2008). The surface of sea ice also represents the interface with the atmosphere and/or any overlying snow layers. While microbial communities of snow appear to undergo post-depositional selection, the underlying ice type might contribute to these processes (Larose, Dommergue, & Vogel, 2013; Maccario, Carpenter, Deming, Vogel, & Larose, 2019). In light of this, we conducted a field campaign at the Villum Research Station, Greenland. Four geographically distinct but neighbouring sites representing both FYI and MYI were selected for sampling. High throughput next generation sequences (NGS) of 16S rRNA gene amplicons and shotgun metagenomics was used to (1) explore the taxonomic variance of MYI and FYI manifest in the functional potential of these microbial communities (2) determine whether FYI and the overlying brine fed snow formed synonymously across geographically distinct sites and (3) evaluate the influence of surface ice type on the post-depositional selection processes in snow.

4.1 Methods

4.1.1 Site description and sampling procedures

Samples were collected between the 26/04/2017 and the 02/05/2017 from snow-covered sea ice in the vicinity of Station Nord, Greenland (81°43'N, 17°47'W) with no close source of human activity; with the exception of the research station. A total of 4 different sites were sampled (see Figure 4.1 for co-ordinates and site descriptions). All FYI sites were land-fast ice (LFI), but for simplicity and given its proximity to the land, site C is referred to as the LFI site and site A and B are referred to as the FYI sites.



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, US
IGN, and the GIS User Community

Site A: First year ice, 1.2 m thick
N81°34.401’
W017°02.062’

Site B: First year ice
N81°36.346’
W017°06.154’

Site C: Multi-year ice depth 2.2 m
N81°39.393’
W016°08.721’

Site D: First year ice
N81°36.946’

Figure 4.1: Satellite image of the sampling sites at a warmer time of year.

At each site, a 10 m^2 pristine snowfield with a visually homogeneous snow cover was selected for triplicate snow pit sampling (see Figure S4.1 for sampling details) for surface snow layers (referred to as ISX), basal snow layers (referred to as ISSX), brine (when available, referred to as IBX) and sea-ice cores. The top five centimeters of the snowpack were sampled for surface snow and the basal saline snow layer was identified based on physical structure and collected in 3-L sterile bags (whirlpack sampling bags). Snow-cleared sea ice was drilled with a 9-cm diameter ice corer (Mark II Coring System, Kovac Enterprise) to form a 50-cm deep sackhole and left to fill with brine for one hour and recovered (when possible) using a 1-L syringe. Ice cores were recovered for biological and chemical analysis. All sampling and storage materials were rinsed with ethanol and a subset of the sampled media (snow, sea ice) prior to collection. Subsamples from each layer were collected in sterile Acuvettes for chemical analysis. Temperature was measured for each horizon using an electronic thermometer. Salinity and pH were measured from samples transported directly to the laboratory. Samples were processed at the Station Nord research facility. Snow samples were melted at room temperature under constant agitation to avoid warming and filtered immediately after melting. All samples were filtered using 0.2 µm filters (Merck Millipore). Filters were stored at –20°C and shipped to the laboratory in France for further processing.

4.1.2 Chemistry

Anion and cation chemistry was measured at the Laboratoire de Glaciologie et Géophysique de l'Environnement by conductivity-suppressed Ion Chromatography on a Dionex ICS 3000.The following chemical variables were measured; Fl, NO_3^-, Cl^-, SO_4^{2-}, NH_4^+, Ca^{2+}, Na^+, K^+ and Mg^{2+} and oxalic acid.

4.1.3 DNA Extraction, Sequencing and qPCR

DNA extraction and sequencing was carried out as outline in Chapter 2. Briefly, DNeasy PowerWater Kit (Qiagen) was used to extract DNA from filters and quantified on the QubitTM dsDNA HS Assay Kit (Thermo Fisher Scientific). For amplicon sequencing, the V3–V4 regions of the 16S rRNA gene were amplified using the universal primers Bac_341F and Bac_805R modified with the Illumina overhang adapter sequences. All primers are listed in Table S4.1 (Supplementary Materials). PCRs were performed using the Platinum PCR SuperMix (Invitrogen) using the following cycles and conditions: 35 cycles at 92°C 30 s, 55°C 30 s and 72°C. The 16S rRNA gene sequencing libraries were prepared using the 16S rRNA gene Library Preparation Workflow recommended by Illumina. Metagenomic sequencing was conducted on samples with DNA concentrations that permitted. Metagenomic libraries were prepared using the Nextera XT Library Prep Kit and Workflow (Illumina) according to manufacturer's instructions. All sequencing libraries were subject to paired end sequencing on a MiSeq sequencer (Illumina) at the laboratory in Lyon.

16S gene copy numbers were estimated by qPCR using the primer sets listed in Table S4.2 (Supplementary Materials). Standards for 16S and 18S rRNA gene quantification were obtained from PCR products amplified from *E.coli*. Following the cloning of the 16S rRNA sequence fragments in a plasmid, DNA quantification was performed with the Broad-Range Qubit Fluorometric Quantification (Thermo Fisher Scientific) and standards diluted to the applicable qPCR concentrations. qPCR was carried out using the SensiFast SYBR No-Rox kit (Bioline) on a Rotorgene 3000 machine (Qiagen). The reaction mixture contained 10 µL of SYBR master mix, 2 µL of DNA and RNAse-free water to a final volume of 20 µL. The 2-step qPCR program consisted of; enzyme activation (95°C; 2 min), then 35 cycles of 5 s at 95°C and 20 s at 60°C for hybridization and elongation respectively. Subsequently, melt curve analyses were performed at of 1°C increments between 55 to 95°C. Gene copy numbers were calculated from the standard curves of 16S and 18S rRNA genes by using the 1Ct (cycle threshold) method.

4.1.4 Chemical data analysis

A PCA of log transformed data was used to investigate differences between the chemistry of different sample groups and tested using the Adonis function in R. Hierarchical clustering was applied to the resulting PCA to determine partitions in the samples by chemical variation.

4.1.5 Taxonomic and metagenomic data analysis

Raw paired-end reads of the 16S rRNA libraries were primer-trimmed using 'cutadapt' (Martin, 2011), and processed using DADA2 analysis following the DADA2 (version 1.12) pipeline (Callahan *et al.*, 2016) and as detailed in Chapter 2. Following annotation, singletons and ASVs which were not taxonomically assigned to Bacteria/Eukarya were excluded from further analysis. All statistical analyses were conducted using R (v3.4.1)1 in RStudio (v1.0.153; RStudio Team, 2015). Sample data was managed using the R package 'phyloseq' (v1.20.0; McMurdie & Holmes, 2013). Dynamic trimming and primer removal from metagenomic libraries was conducted using Trimmomatic based on their quality scores (PARAMETERS) and read quality profiles were checked in FastQC. Subsequently, functional annotations were assigned using EggNOG-Mapper using default parameters. Annotations were imported into R and gene counts and functional annotation tables were generated using in-house scripts. Plots were generated using the R package 'ggplot2' unless stated otherwise (v2.2.1; Wickham, 2016).

Both the ASV and metagenomic datasets were explored as a whole and partitioned to explore differences in similar sample types at different sites. Given the similarity between ice cores and brine samples at sites A and B, these sample types were analysed together. Hierarchal clustering of the ASV relative abundance table using the ward clustering method and correlation as a distance measure was performed to explore taxonomic differences in bacterial communities. ASV count data was rarefied prior to calculating alpha diversity measures. Significant differences in the alpha diversities was performed using a one way ANOVAa. Overall functional community differences were explored by examining square root transformed abundance data to minimize the influence of the most abundant groups. Bray-

Curtis dissimilarity matrices of square root transformed functional abundances were analysed using Adonis tests in R with 9999 permutations. Assumptions of the Adonis test were confirmed using the *betadisper* function with vegan, which tests the multivariate homogeneity of group dispersions (variances). The 30% most abundant KEGG pathways with 60% best environmental fit were selected using 'ordiselect' of the goeveg R package and plotted using the R package gplots. Presence/absence lists of GO terms in sample types were generated in base R and Venn diagrams were created. Subsequent to the removal of read counts of less than 10, the log fold change of raw abundance matrices of ASVs and GO terms were calculated using the R package 'DESeq2'. Differentially abundant ASVs (daASVs) and GO annotations (daGOAs) were defined by a log-fold change of >1 and a padj < 0.05.

Log transformed chemistry data (excluding pH) and Hellinger transformed taxonomic and functional relative abundance data was evaluated using a co-inertia analysis. Principal component analyses were applied to each dataset using the ade4 R package. Subsequently, co-inertia analyses (ade4) was used to test the impact co-variance of biological data with chemistry. The significance was tested using a permutation test with 1000 permutations.

4.2 Results

4.2.1 Chemical site and sample differences

The physical properties and chemical composition of snow, sea ice and brine were analysed for all samples (Table S2). The chemical composition was significantly different between both sites and sample type (Adonis, p-value = 0.021 and 0.001, respectively), although sample type explained a larger proportion of the variation than site (R2 = 0.64 and 0.25, respectively; Figure S4.2). Hierarchal clustering resolved three distinct groups based on chemical composition; (1) all surface snow samples and the basal snow samples of site C (2) all ice core samples and two basal snow samples of site D (3) saline snow and brine samples of site A and B and one basal snow sample of site D (Figure S4.2). The first axis explained 76.8% of the variation and separated cluster 1 from clusters 2 and 3. The differences were attributed to high concentrations in most major ions (Cl, NO_3, SO_4, Ox, Na, NH_4, K, Mg and Ca) in clusters 2 and 3 and low concentrations in cluster 1. Relative to cluster 2, cluster 1 contained elevated levels of nitrate, although high concentrations were found in two samples from other environments

(Table S4.2; I1B, I1SS1). Relatively high concentrations of oxalate and an increase in pH and salinity characterised cluster 3, although within group oxalate concentrations were patchy (Table S4.2).

4.2.2 Chemistry, taxonomic assemblage and functional profiles in snow and sea ice communities

Co-inertia analysis was used to explore the effect of chemistry on taxonomic assemblage in snow and sea ice communities. The co-variance of taxonomy and chemistry was not significant when using the abundance of ASVs or genera (RV = 0.167 and 0.162, pvalue = 0.247, 0.103 respectively), but did explain some of the variation at higher taxonomic levels. At the family level, 24.6% of the variation in the abundance was predicted to be explained by the chemistry ($p = 0.01$). The first axis generally split the surface snow from the ice cores, brine and saline snow and the second axis separated the ice cores and the saline snow (Figure S4.3). The differences in the saline snow were explained by a dominance of the families Alteromonadaceae, Colwelliaceae and Oceanospirillaceae as determined by the size of the vectors. The family Cryptomonadaceae had the largest effect on the ordination in relation to the ice core samples, while the surface snow cluster was characterised by members of Cyanobacteria (GPI, GPIV, GPXIII), Cytophagaceae and Bradyrhizobiaceae. Overall taxonomic profiles of sea ice, brine and snow communities based on ASVs were significantly different based on site (ADONIS; R2 = 0.19, pvalue = 0.001) and on sample type (ADONIS: R2 = 0.31; pvalue = 0.001). Diversity was generally highest in surface snow and brine, and lowest in saline snow (Table S4.3).

The co-variance of the chemistry and function data was significant for all levels of annotation (KEGG pathways, KO ids and GO terms) and chemistry was predicted to explain approximately 40% of the functions (pvalues = 0.001). A similar split of samples by axis was seen as in the taxonomy. Overall functional potential of sea ice, brine and snow communities based on KEGG pathways were not significantly different based on site (ADONIS; R2 = 0.16, pvalue = 0.58) but were based on sample type (R2 = 0.50; pvalue = < 0.01). The KEGG pathways representing the 30% most abundant with the 60% best environmental fit based on an NMDS of the square root relative abundances using Bray-Curtis distances are plotted in Figure S4.4. Dominant KEGG pathways that explained the variation across samples included those related to fatty

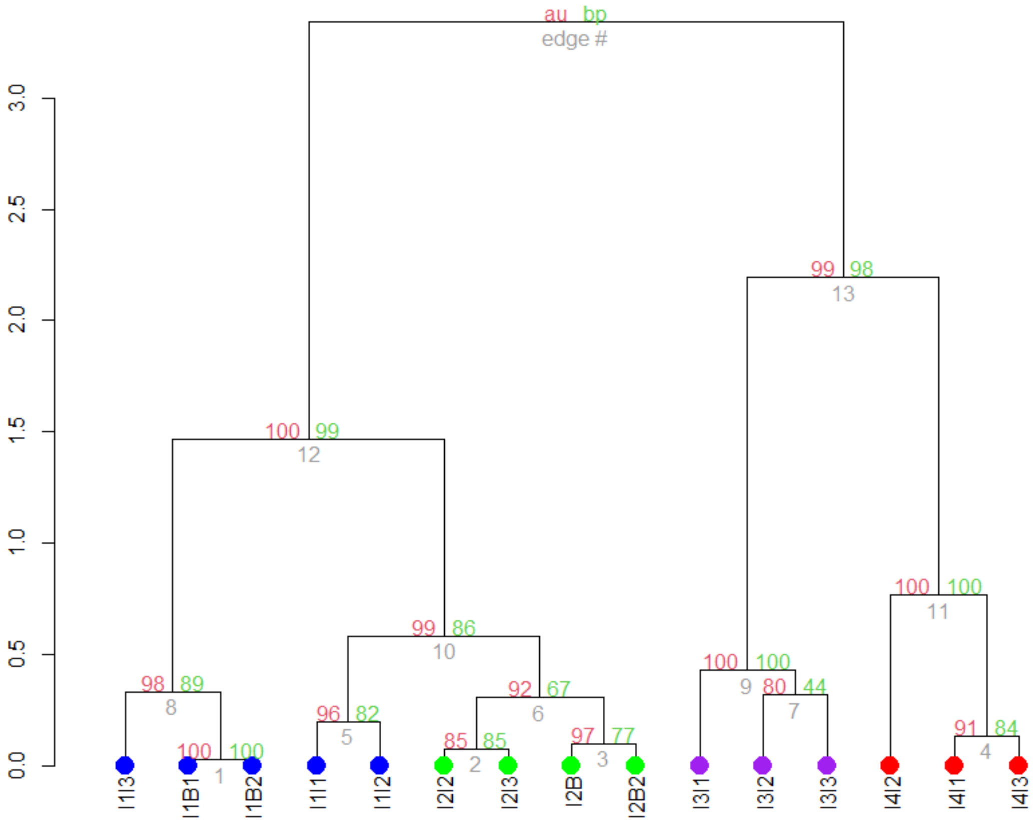


Figure 4.2: Hierarchical cluster using ward clustering method and correlation as distance measure and 1000 bootstrap resampling based on relative abundance of all ASVs detected. Bootstrap probability (BP) values in percentage are indicated in green and AU pvalues are indicated in green. AU >= 95% are considered to be strongly supported by the data. Distance between clusters is indicated by height bar. Colours represent surface (purple) or basal snow from each sample site (blue=site A, green=site B, purple=site C and red=site D).

acid, amino acid (valine, leucine, isoleucine, tryptophan, phenylalanine and tyrosine), sulfur, nitrogen, aromatic compounds (benzoate/aminobenzoate degradation) and chlorinated alkane/alkene metabolism/catabolism. ABC transporters, two-component systems, quorum sensing, bacterial secretion systems, photosynthetic pathways and peroxisome processes were also abundant and explained some of the variation seen between sample types (Figure S4.4).

Figure 4.3: Stacked bar charts of the relative abundance of ASVs at class level. Samples are grouped by site and type.

4.2.3 Bacterial community composition and functional potential of ice types

Bacterial community composition of ice cores

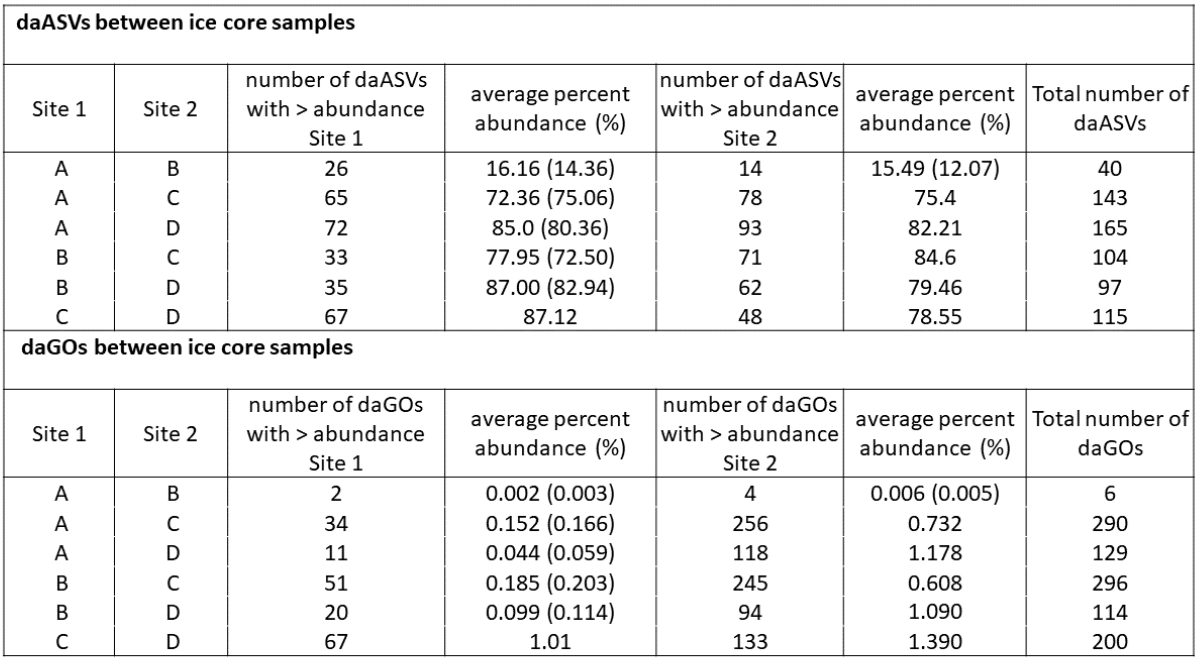
To understand whether sea ice type influences taxonomic assemblage of bacterial communities, we compared taxonomic profiles of ice cores and their associated brines samples (where possible) from four distinct but neighbouring sites representing FYI (sites A and B), MYI (site C) and LFI (site D). There was no significant difference in the alpha diversity of ice cores from the different sites (Table S4.3). Abundance estimates for bacteria ranged from 8.3x10^3-2.3x10^5 16S rRNA copies/mL, with MYI harbouring a 10-fold increase in copy numbers relative to the other ice types (Table S4.2). Hierarchal clustering using the ward method and Bray-Curtis distance as a measure on the relative abundance of ASV's demonstrated that samples generally clustered by sampling site (Figure 4.2).

At class level, two of the most abundant taxa in FYI cores and brine were Gammaproteobacteria and Alphaproteobacteria, constituting between 10-63% and 11-41% of the community, respectively (Figure 4.3). Conversely, these classes were absent or represented very small abundances in the LFI and MYI communities (<5% on average). The MYI was dominated by Betaproteobacteria (average 32%) and Actinobacteria (average 35.5%) with smaller contributions from Cytophagia (average 7.3%), Sphingobacteria (average 4.9%) and Verrucomicrobiae (average 2.9%). The LFI also contained variable and large proportions of Betaproteobacteria (1-77%), but was dominated by sequences annotated with chloroplast origin (14-68%).

Clostridia and Bacilli are also present at relative abundances >2% (Figure 4.2). The number of differentially abundant ASVs (daASVs) between ice cores and their associated brines at different sites and their average contribution to the relative abundance can be found in Table 4.1.

A similar pattern of enrichment at site A and B, the FYI sites, was evident relative to MYI (site C) and LFI (site D). More daASVs annotated at the class Actinobacteria were enriched in site C and D relative to the FYI sites, specifically of the family Microbacteriaceae and including the genera Frondihabitans, Cryobacterium and Rhodoluna (Table S4A-F). There were also greater counts of daASVs of the class Betaproteobacteria enriched in site C relative to A/B; all of which belonged to the order Burkholderiales and represented the families Oxalobacteraceae, Comamonadaceae and Burkholderiales_incertae_sedis. Additionally, more Sphingobacteriales dASVs were enriched in site C relative to B. Site D had a higher abundance of sequences of chloroplast origin relative to sites A/B and C, particularly Streptophyta (Table S4.4A-F). Conversely, FYI sites harboured a greater number of enriched dASVs annotated as Gammaproteobacteria and Alphaproteobacteria including the genera *Colwellia* and *Glaciecola* relative to the other sites (Table S4.4A-F).

Table 4.1: Differentially abundant ASVs (daASVs) and GO terms (daGOs) as identified by DESeq2 with a padj <0.05 and a log2foldChange > 1. Average proportional abundances of all enriched daASVs and daGOs are given in percentages. Numbers in parenthesis represent values for associated brine samples, when collected.



daASVs between ice core samples

Site 1	Site 2	number of daASVs with > abundance Site 1	average percent abundance (%)	number of daASVs with > abundance Site 2	average percent abundance (%)	Total number of daASVs
A	B	26	16.16 (14.36)	14	15.49 (12.07)	40
A	C	65	72.36 (75.06)	78	75.4	143
A	D	72	85.0 (80.36)	93	82.21	165
B	C	33	77.95 (72.50)	71	84.6	104
B	D	35	87.00 (82.94)	62	79.46	97
C	D	67	87.12	48	78.55	115

daGOs between ice core samples

Site 1	Site 2	number of daGOs with > abundance Site 1	average percent abundance (%)	number of daGOs with > abundance Site 2	average percent abundance (%)	Total number of daGOs
A	B	2	0.002 (0.003)	4	0.006 (0.005)	6
A	C	34	0.152 (0.166)	256	0.732	290
A	D	11	0.044 (0.059)	118	1.178	129
B	C	51	0.185 (0.203)	245	0.608	296
B	D	20	0.099 (0.114)	94	1.090	114
C	D	67	1.01	133	1.390	200

Functional community structure of ice cores

To understand whether functional content varied between different ice core communities of FYI, LFI and MYI, we compared their metagenomes using Bray-Curtis distances. All levels of annotation (KEGG pathways, GO terms and KO numbers) were significantly different between ice cores and their associated brine samples (where available) from different sites (Table 4.2).

Venn diagrams were used to investigate the overlap between the presence/absence of GO terms in ice cores (Figure 4.4). A more diverse number of genes were annotated in sites A and B. Not only did site A and B share the highest number of annotations, but they also contained the most unique annotations. The largest intersect was annotations common to all ice types (Figure 4.3), but there were also a number of annotations uniquely attributed to site.

Table 4.2: Results of significance testing of r-squared transformed relative abundance data annotated at different levels using the "adonis" function from the R-package Vegan and grouping factor was either "site" or "sample type".

	Square root transformation				No transformation			
Ice cores and brine	**Site**		**Sample Type**		**Site**		**Sample Type**	
Annotation level	**R-sq**	**p value**	**R-sq**	**p value**	**R-sq**	**p value**	**R-sq**	**p value**
KEGG pathways	0.4092	0.0169	0.0976	0.297	0.3885	0.0781	0.0891	0.3526
KO numbers	0.4567	0.0002	0.0627	0.706	0.4125	0.001	0.0466	0.9251
GO terms	0.4545	0.0015	0.0653	0.577	0.3235	0.238	0.0259	0.7987

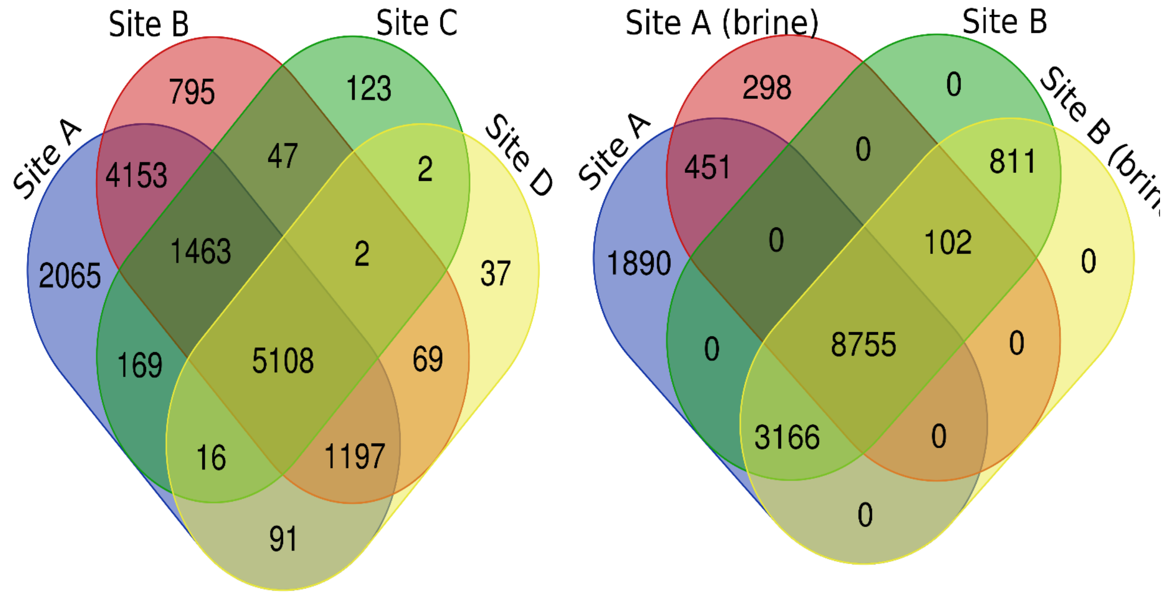


Figure 4.4: Venn diagram summarizing the number of shared and unique GO annotations. The left-most Venn diagram includes all ice core samples, the right-most compares site A and B ice core and brine samples.

Differentially abundant GO terms (daGOs) were identified among the ice samples (Table 4.1), with the highest amount found between the FYI sites and MYI (290 and 296), followed by MYI and LFI (200). However, these terms represented an abundance of 1% or less in the metagenomes. The GO categories that were more abundant in site A (FYI) relative to site C (MYI) were related to related to benzoate, glycolate, cysteine and lipopolysaccharide metabolism/catabolism, protein secretion (type IV secretion system), and transmembrane transport (specifically bacteriocin and phosphate; Tables S4.6A-F). These annotations were also abundant in site B relative to C, with the addition of biofilm formation. On the other hand, more abundant annotations in the MYI relative to the FYI sites included sporulation, motility and cellular components pertaining to eukaryotes (*e.g.* plastids, nucleus, thylakoid components). Annotations relating to the metabolism of nucleic acids and acetoin and gene silencing were also enriched in MYI. Comparisons of FYI and LFI cores also showed an enrichment in annotations relating to protein secretion systems in FYI (sites A and B). Enriched terms in LFI cores relative to FYI included terms related to the catabolism of pyrimidine, siderophores and those associated to eukaryotic structures (chlorophyll). There were very few (n=6) GO categories that were differentially abundant between site A and B ice core and brine samples. Most were more abundant in site B and were related to motility and biofilm formation (Table S4.6A).

4.2.4 Bacterial composition and functional profiles of the snow horizons

Bacterial community composition of the snow

To explore the effect of post-depositional processes on snow and whether the underlying ice type influenced these processes, we compared the taxonomic profiles of surface snow and basal snow from all four sites. The analysis demonstrated the tight clustering of saline snow samples from sites A and B. Overall, basal snow samples and surface snow samples formed a separate cluster from the saline snow (Figure 4.5).

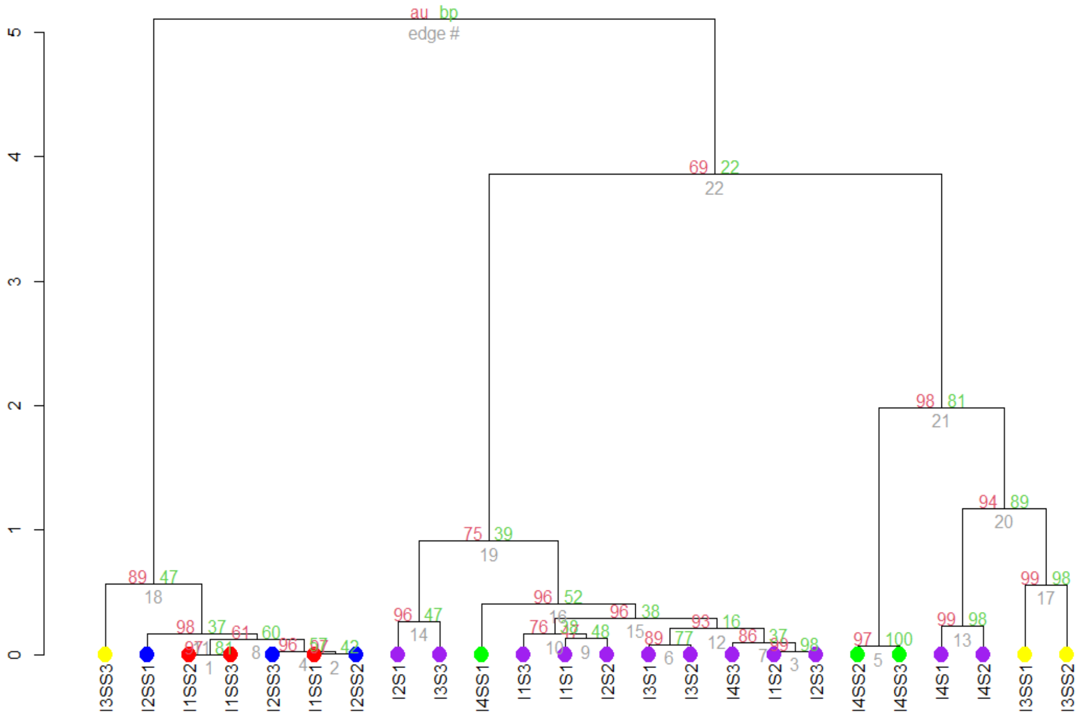


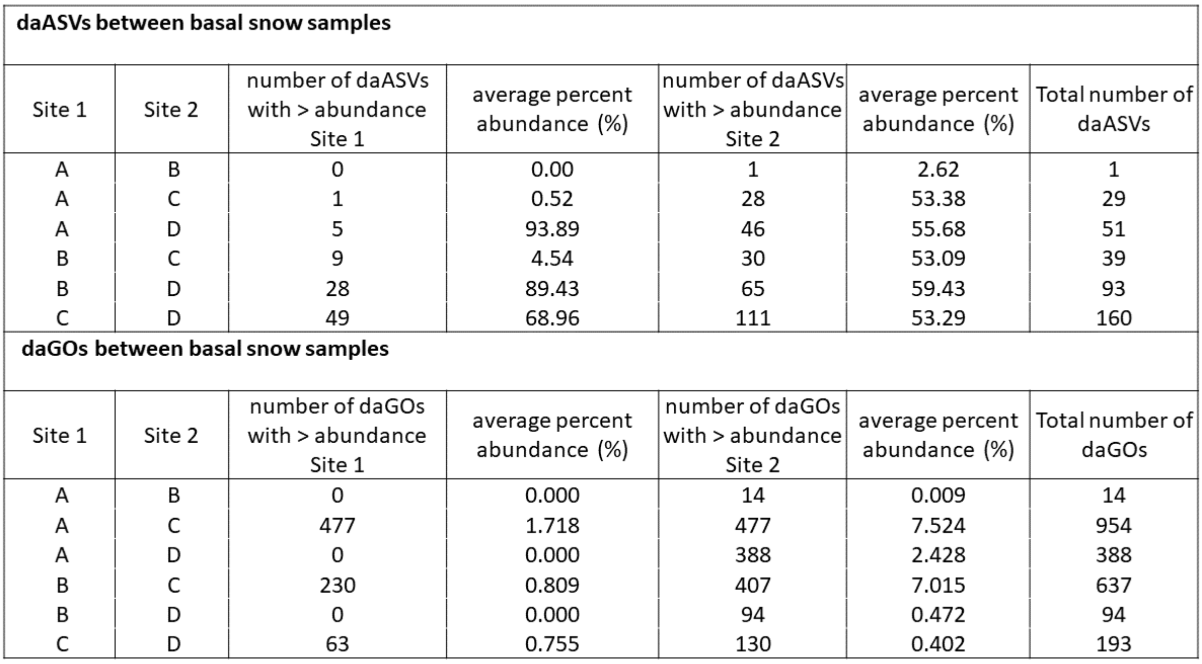
Figure 4.5: Hierarchical cluster using ward clustering method and correlation as distance measure and 1000 bootstrap resampling based on relative abundance of all ASVs detected. Bootstrap probability (BP) values in percentage are indicated in green and AU pvaues are indicated in green. AU >= 95% are considered to be strongly supported by the data. Distance between clusters is indicated by height bar. Colours represent surface (purple) or basal snow from each sample site (red=site A, blue=site B, yellow=site C and green=site D).

The basal snow and surface snow cluster could be approximately divided by horizon, although two site D surface snow samples (14S1 and I4S2) clustered closely to the basal snow horizons. Similarly, one of the basal snow samples of site D (I4SS1) clustered amongst the surface snow. Abundance estimates for bacteria ranged from $1.34*10^1$-$3.51*10^2$ 16S rRNA gene copies/mL in surface snow, from $6.74*10^0$-$4.36*10^3$ 16S rRNA copies/mL in MYI basal snow and from $3.50*10^2$-$1.37*10^3$ 16S rRNA gene copies/mL in LFI basal snow. Saline snow from FYI samples had 10-100 fold higher 16S rRNA gene copy numbers relative to the other ice types (ranging from $1.01*10^4$-$8.96*10^4$ 16S rRNA gene copies/mL) (Table S4.1). The diversity of the basal snow at sites A and B were significantly lower relative to any other environment or site (Table S4.3). Gammaproteobacteria overwhelmingly dominated the basal snow layer at sites A and B (59.4-98%) with small contributions of Alphaproteobacteria (<5% on average). One sample (I2SS1) harboured a slightly more diverse community, including Alphaproteobacteria (14.2%), Actinobacteria (8.1%), Betaproteobacteria (8.1%) and Flavobacteriia (6.4%), but was still dominated by Gammaproteobacteria (59.4%). Comparatively, the basal snow layers at site C and D harboured greater proportional contributions of Actinobacteria (average 29.4%), Alpha- (average 12.2%) and Beta- proteobacteria (average 5.6%) and sequences of chloroplast origin (average 16.2%). Cytophagia and Cyanobacteria were dominant classes in two of the three basal snow samples at site C, but not detected in the third replicate or any of the basal snow samples at Site D. Surface snow was highly diverse (Table S4.3) but relatively consistent across samples at class level (Figure 4.3). Abundant classes include Alphaproteobacteria (average 41.7%), Cyanobacteria (average 7.3%) and Actinobacteria (average 12.3%).

enriched daASVs and daGOs are given in percentages.

To understand the differences in taxonomic presence and abundance in basal snow samples collected from different sites, we used DESeq2 analyses to identify differentially abundant ASVs (Table 4.3). Nearly all of the daASVs in site B and C comparisons were identified in site A and C comparisons. daASVs enriched at site C were annotated as either Actinobacteria,

Table 4.3: Differentially abundant ASVs (daASVs) and GO terms (daGOs) as identified by DESeq2 with a padj <0.05 and a log2foldChange > 1. Average proportional abundances of all



daASVs between basal snow samples

Site 1	Site 2	number of daASVs with > abundance Site 1	average percent abundance (%)	number of daASVs with > abundance Site 2	average percent abundance (%)	Total number of daASVs
A	B	0	0.00	1	2.62	1
A	C	1	0.52	28	53.38	29
A	D	5	93.89	46	55.68	51
B	C	9	4.54	30	53.09	39
B	D	28	89.43	65	59.43	93
C	D	49	68.96	111	53.29	160

daGOs between basal snow samples

Site 1	Site 2	number of daGOs with > abundance Site 1	average percent abundance (%)	number of daGOs with > abundance Site 2	average percent abundance (%)	Total number of daGOs
A	B	0	0.000	14	0.009	14
A	C	477	1.718	477	7.524	954
A	D	0	0.000	388	2.428	388
B	C	230	0.809	407	7.015	637
B	D	0	0.000	94	0.472	94
C	D	63	0.755	130	0.402	193

Alphaproteobacteria, Chloroplast, Cyanobacteria, Cytophagia, Gammaproteobacteria, Opitutae or Sphingobacteriia. The only ASV that was more abundant at site A was annotated as *Colwellia* (Villum_363). Conversely, ten daASVs were enriched in site A relative to site C representing the genera *Glaciecola*, *Reinekea* and *Colwellia*. Comparisons between sites A/B and D yielded similar results, but with the addition of *Octadecabacter*. Notable daASVs between the basal snow layers of site C and D include an enrichment of cyanobacterial ASVs in C (Table S4.5A-F).

To determine whether FYI and the overlying brine fed snow formed synonymously across geographically distinct sites, we implemented hierarchal clustering using the Ward method and Bray-Curtis as a distance measure on the relative abundance of ASVs. Ice and brine samples generally clustered based on sampling site at sites A and B, while the saline snow from both sites clustered together (excluding I2SS1; Figure 4.6). Bacterial abundance in brine samples was slightly higher than that of saline snow samples, ranging from 7.38x10^4-1.59x10^5 16S rRNA gene copies/mL (Table S4.1).

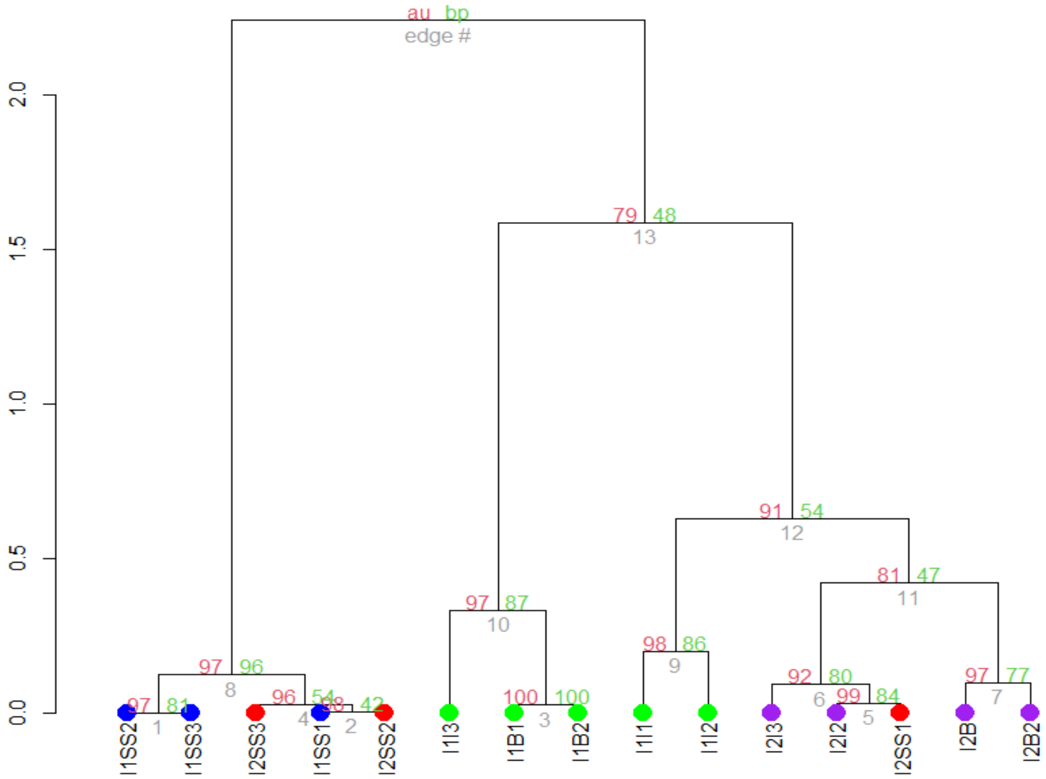


Figure 4.6: Hierarchical cluster using ward clustering method and correlation as distance measure and 1000 bootstrap resampling based on relative abundance of all ASVs detected. Bootstrap probability (BP) values in percentage are indicated in green and AU pvalues are indicated in green. AU >= 95% are considered to be strongly supported by the data. Distance between clusters is indicated by height bar. Colours represent ice (purple), brine (green) or saline snow from each sample site (blue=site A, red=site B).

There was only one daASV between the saline snow horizons at site A and B and it corresponded to *Comamonadaceae* (Table S4.5A). There were a total of 40 daASVs identified between ice core and brine samples of site A and B. In the ice core and brine communities of site B, all the daASVs were annotated as Gammaproteobacteria, specifically *Reinekea*, *Colwellia* and *Glaciecola*. Conversely, among the annotations of the enriched daASVs at site A were *Actinobacteria, Alphaproteobacteria, Deltaproteobacteria, Flavobacteriia* and *Sphingobacteria* (Table S4.4A).

Functional community structure of snow horizons

Table 4.4: Results of significance testing of r-squared transformed relative abundance data annotated at different levels using the “adonis” function from the R-package Vegan and grouping factor was either “site” or “sample type”

	Square root transformation				**No transformation**			
Surface, saline and basal sno	**Site**		**Sample Type**		**Site**		**Sample Type**	
Annotation level	**R-sq**	**p value**	**R-sq**	**p value**	**R-sq**	**p value**	**R-sq**	**p value**
KEGG pathways	0.155	0.0548	0.5	1.00E-04	0.23791	0.1643	0.63623	1.00E-04
KO numbers	0.20703	0.1727	0.40711	1.00E-04	0.1962	0.2718	0.50822	1.00E-04
GO terms	0.23021	0.1241	0.54698	1.00E-04	0.10877	0.6537	0.82142	1.00E-04

Snow samples were significantly different from each other dependent on sample type (saline snow, basal snow and surface snow) but were not significantly different by site (Table 4.4).

Overlap analysis indicated the highest amount of annotations were functions shared by all metagenomes (**Figure 5**). Site D had the largest number of unique annotations.

DaGOs were identified among the basal snow samples (Table 4.3), with the highest amount found between the saline snow (sites A and B) and MYI basal snow (954 and 637), followed by saline snow from site A and LFI (637). A total of 193 daGOs were found between MYI and LFI basal snow samples. Overall, these terms represented an abundance of 7.5% or less in the metagenomes. GO terms enriched in the saline snow horizons of site A and B relative to the basal snow layers of sites C include annotations related to biofilm formation, motility, inorganic (sodium, rubidium, magnesium and iron) and organic (sulfate) ion transport, lactate metabolism, glycine betaine and glutathione biosynthesis, plasmid recombination and secretion complexes (Table S4.7A-F). However, no statistically significant GO term enrichment in FYI sites A and B relative to LFI were detected (Table 4.3). Basal snow from sites C and D were enriched in

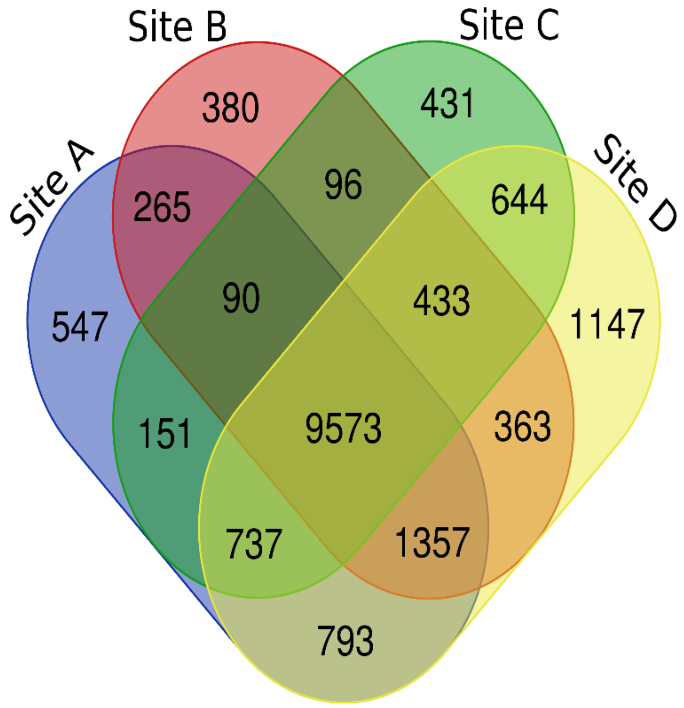


Figure 4.7: Venn diagram summarizing the number of shared and unique GO annotations of all basal snow horizons.

GO terms related to fungal and algal metabolism and cellular processes relative to the FYI sites (Table S4.7A-F). There were 14 GO terms enriched in site B saline snow horizons relative to site A and none in the reverse comparison. These terms were mostly photosynthesis and eukaryotic related (pertaining to thylakoid and mitochondrial components and photosystems; Table S4.7A).

4.3 Discussion

Does sea ice type drive taxonomic and functional community structure?

Differences in ice core and brine taxonomic microbial community were evident between site D (LFI), which represented ice directly anchored to the land, and sites A and B (FYI) with the closest proximity to the open ocean. For example, sites A and B harboured a large abundance of typically marine, copiotriophic Gammaproteobacteria that were absent in site D ice cores. Instead, site D was enriched in ASVs linked to psychrophilly and oligotrophy microorganisms such as *Masillia* and *Variovorax*. The facultatively psychrophilic *Massillia eurypsychrophila and M. glaciei* and various *Variovorax* species have been isolated from glacial ice cores and other species of these genera have been identified in oligotrophic environments. In addition, a larger fraction of sequences were predicted to be of chloroplast origin, such as the green algae *Streptophyta*, that have a preference for freshwater life styles (Becker & Marin, 2009). These results are in line with observations that LFI is influenced more by riverine input and sediment upwelling than pack ice (Wadhams, 2000), but highlights the gradient from land.

The increased thickness of MYI results in more stable conditions and less steep gradients of temperature, light and salinity and a reduction or complete elimination of the brine, where the archetypal copiotrophic sea ice microorganisms typically thrive. We would expect this to result in a loss of diversity and selection of specialist bacteria under the stable conditions. Surprisingly, the abundance and diversity of MYI bacterial communities increased relative to other ice types and is inconsistent with previous observations (Bowman et al., 2012). An important difference with our study is we targeted the upper horizon of sea ice while the Bowman *et* al. (2012) integrated the entire core. However, the low abundance of Gammaproteobacteria in our MYI samples is consistent with the lack of brine and follows the decrease in abundance of this class seen in a previous sectional analysis of MYI cores (Hatam et al., 2014). The taxa that were enriched in the MYI relative to sites A and B belonged to genera primarily found in terrestrial and sediment environments, such as *Frondihabitans*, *Pedobacter* and *Hymenobacter*. As brine was major component of site A and B ice cores, yet, was lacking at site C, the increase in typically copiotrophic marine microorganisms can be explained by brine provision. *Colwellia*, *Glaciecola* and *Octadecabacter* are all major components of brine communities (Junge, Cameron, & Nunn, 2019) and are abundant at site A and B compared to C.

The overall metabolic pathways were similar between ice core samples and a large fraction of the annotations were shared between sites. Statistical analyses of the abundance of functional pathways (KEGG pathways and GO terms) did not separate the communities from all four sites except at high resolution annotations (KO ids). Given the taxonomic analyses demonstrated distinct clustering of samples by location, we suggest a high level of functional redundancy in surface sea ice across time and space. The taxonomic variation and lack of functional variation could be due to the different seeding sources of the ice types and similar forcing physiochemical conditions, respectively. Surface layers of the MYI have undergone longer selection and surface processes, including meltwater infiltration, that will affect the taxonomic seeding of the community. On the other hand, surface FYI is predominantly seeded by the underlying ocean and has undergone less exposure to surface processes while subjected to similar physiochemical conditions as the other ice types (Figure S2). The co-variance of chemistry and ASV datasets was not significant (co-inertia coefficient RV = 0.17, p-value = 0.212) whereas the chemistry explained approximately 40% of the variation of functional content (GO-terms; p-value = 0.001). However, a high level of functional redundancy is a general characteristic of stable bacterial communities. When square root transformations were applied to relative abundance data to reduce the influence of the dominant categories, statistical analysis was able to discern the communities of all four sites for all annotation types. This demonstrated that while major and abundant functions are similar across ice types, less abundant functions show variation in their presence and abundance.

Although a large proportion of GO terms were shared between samples, site A and B had both the highest number of unique and shared annotations demonstrating the potential for ice type governed selection. GO terms that were enriched differentially between sites only represented a small proportion of the total abundance yet pertained to unique characteristics of sea ice types. For instance, annotations associated with protein secretion, transmembrane transport, biofilm formation and quorum sensing were enriched in site A and B (specifically the brine fractions) and have previously been associated with a copiotrophic lifestyle. Comparatively, annotations linked to eukaryotic organisms and photosynthesis were more prevalent in both site C and D, mirroring the taxonomy and previous studies.

The influence of brine on snow communities

Despite the taxonomic differences in the ice cores of site A and B, no differentially abundant ASVs were evident between the two sample sites for saline snow and they clustered independently to the underlying ice environment. In turn, it appears that niche selective conditions exert more control on taxonomic assembly of saline snow than geographic distance. These synonymous communities demonstrated minimal variability, low diversity and a dominance of *Glacieocola* and to a lesser extent *Colwellia*. Interestingly, the dominance of these genera are mirrored in two previous studies of saline snow (Maccario et al., 2019, Chapter 3) suggesting specific and highly selective environmental conditions independent of space and time resulting in the dominance of highly specialised taxa. Brine rejection to the basal snow horizons has been recognised as the major driving force of this taxa specific enrichment. This is supported by the bacterial cell abundance estimates, which are comparable between brine samples and the overlying saline snow and between 10 to 100 fold higher than in the LFI and MYI basal snow. However, it is still not clear whether this phenomenon is related to the delivery of *Colwellia* and *Glaciecola* to the snow or the provision of nutrients which stimulate the growth of populations already present in the snow. A single basal snow (I3SS3) sample from the MYI profile also harboured large relative abundance of these genera which could be the result of localised brine rejection. Although the chemistry only showed a slight increase in salinity, there is likely patchiness within the sample itself. Moreover, this sample contained a large number of reads associated with chlorophyll, specifically diatoms from the family Bacillariophyta. These organisms secrete compounds utilised by many bacteria and could support the large copiotrophic *Colwellia* and *Glacieocola* populations. Finally, bacterial counts are a factor of 100 times higher in this sample relative to the other MYI basal snow layers, indicating biomass enrichment, possibly by brine rejection (Table S4.1). In terms of functional potential, FYI sites A and B were characterised by terms associated with flagellum mediated motility, plasmid recombination and glutathione/glycine betaine biosynthesis relative to MYI. These have also previously been shown to be enriched in the saline snow (Maccario et al., 2019), further supporting the hypothesis of niche-based selection.

While basal snow samples of the brine rich sea ice communities were extremely selective for copiotrophic, archetypal sea ice bacteria and their associated functions, site C and D basal snow samples harboured greater diversity of bacteria which were more reflective of the surface snow. The comparative lack of nutrient input at site C and D could explain the low bacterial cell abundance estimates and the dominance of snow microorganisms that may be adapted to more oligotrophic conditions. Basal snow from sites C and D were enriched in GO terms almost exclusively related to eukaryotes, such as fungal and algal metabolism and cellular processes (i.e. cell wall components), including photosynthesis (Table S4.7A-F). Although we were unable to sequence the eukaryotic community to derive taxonomic information, the dominance of fungal-related genes in LFI and MYI basal snow metagenomes suggests the presence of cold-adapted fungi. Fungi have been shown to be active in oligotrophic environments such as snow and ice (ref). Interestingly, there was no enriched terms in FYI saline snow relative to site D, the LFI site, suggesting that all the GO terms in saline snow were also present in LFI, despite the absence of brine. In addition, when comparing LFI and MYI basal snow, the enriched terms in LFI correspond to terms also found in saline snow. This would suggest that the LFI basal snow contains legacy functions from brine rejection during ice formation, which could also explain the higher bacterial abundance relative to MYI basal snow.

While it is difficult to discern whether the variability seen in the snow layers is a function of differences in the composition of depositional events or post-depositional processes, there is evidence for the evolution of the snowpack even in the absence of brine supplementation. Community assemblage in fresh surface snow samples was generally independent of location, while basal snow samples indicated localised change. Surface snow which had undergone post-depositional transformation at site C, as evidenced by observations in snow crystal structure (Table S1), bared more similarity to the basal snow layer at site C and D than freshly deposited surface snow at other locations. The surface underlying the basal snow samples appeared to exert an influence on the functional signatures of the basal snow communities. For instance, genes related to photosynthesis were more frequent in the basal snow samples of site C and D. The 16S rRNA dataset for MYI basal snow indicated a large relative abundance of diatoms, specifically from the family *Bacilliarophyta* (up to 40% relative abundance). These sequences were found in low abundances in the surface snow (<1.4%) and in the ice core samples (<1.3%) of site C. However, given the bacterial cell abundance estimates that are up to 1000-fold higher in the ice cores relative to the snow samples, MYI might serve as a potential seedbank of these microorganisms. It has been shown that some ice-algae over-winter in glacial environments and rapidly bloom during the onset of increased light availability (Olsen et al., 2017; Williamson et al., 2019). Therefore, we speculate that the coupled increase in abundance of diatom and photosynthesis related sequences in the basal snow of these sites are due to active migration of diatoms from the underlying ice cores, exploiting the light availability. Interestingly, the surface snow samples from all sites do not harbour many diatom sequences but instead carry a comparatively large abundance of the photosynthetic Cyanobacteria, illustrating spatial separation of photosynthesis in the snow pack. Cyanobacteria have been reported as abundant in fresh snow samples in the high Arctic previously (Harding, Jungblut, Lovejoy, & Vincent, 2011; Maccario et al., 2019). Although UV extinction was not measured in this study, this stratification of photosynthetic organisms may be related to differences in UV tolerance. Although not significant, relative to the basal snow at site C, the surface snow harbours a higher abundance of genes related to UV protection (GO:0009650; padj = 0.08). Specific functions that act against oxidative stress are significantly enriched in surface snow such as glutamate-cysteine ligase activity (GO:0004357, padj < 0.02) and spermidine biosynthetic process (GO: 0008295 , padj < 0.02). Glutamate-cysteine ligase catalyses the rate-limiting step in the production of the cellular antioxidant glutathione (GSH) and spermidine can act as a powerful free radical scavenger.

4.4 Supplementary figures

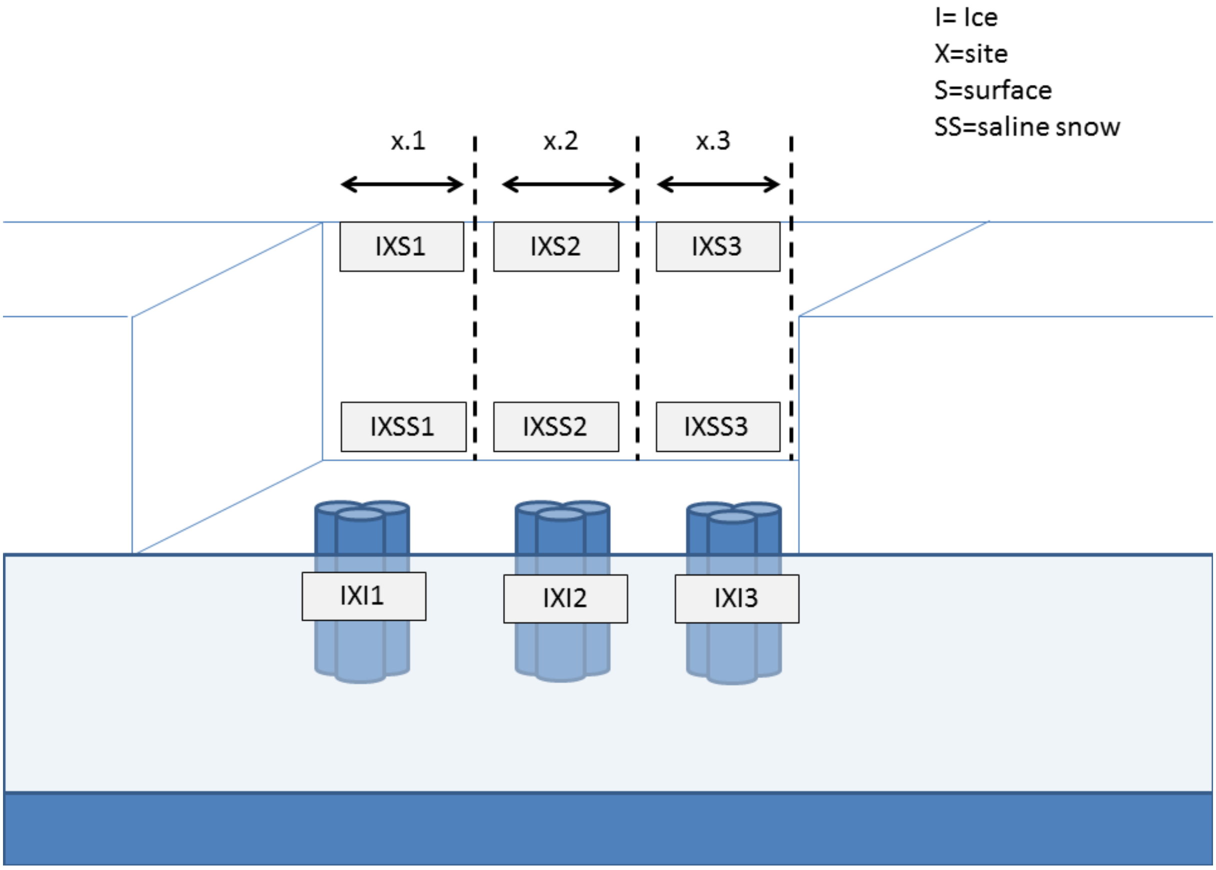


Figure S4.1: Sampling strategy at each site

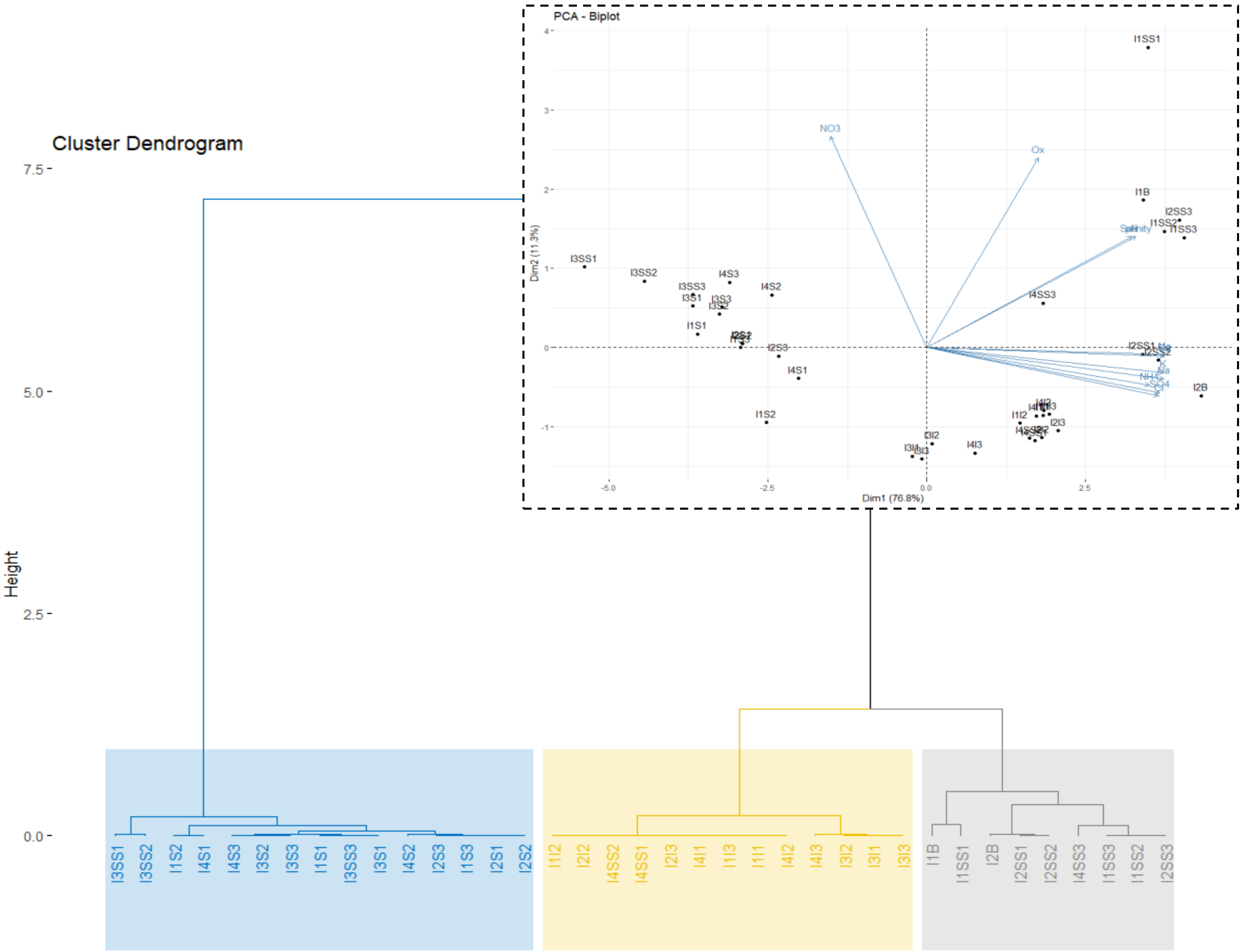


Figure S4.2: (A) Principal component analysis (PCA) biplot of the chemical analyses of samples used in this study. Vectors represent the different chemical variables. Samples are represented by points and labelled appropriately (B) Hierarchal clustering of (A) using the ward method.

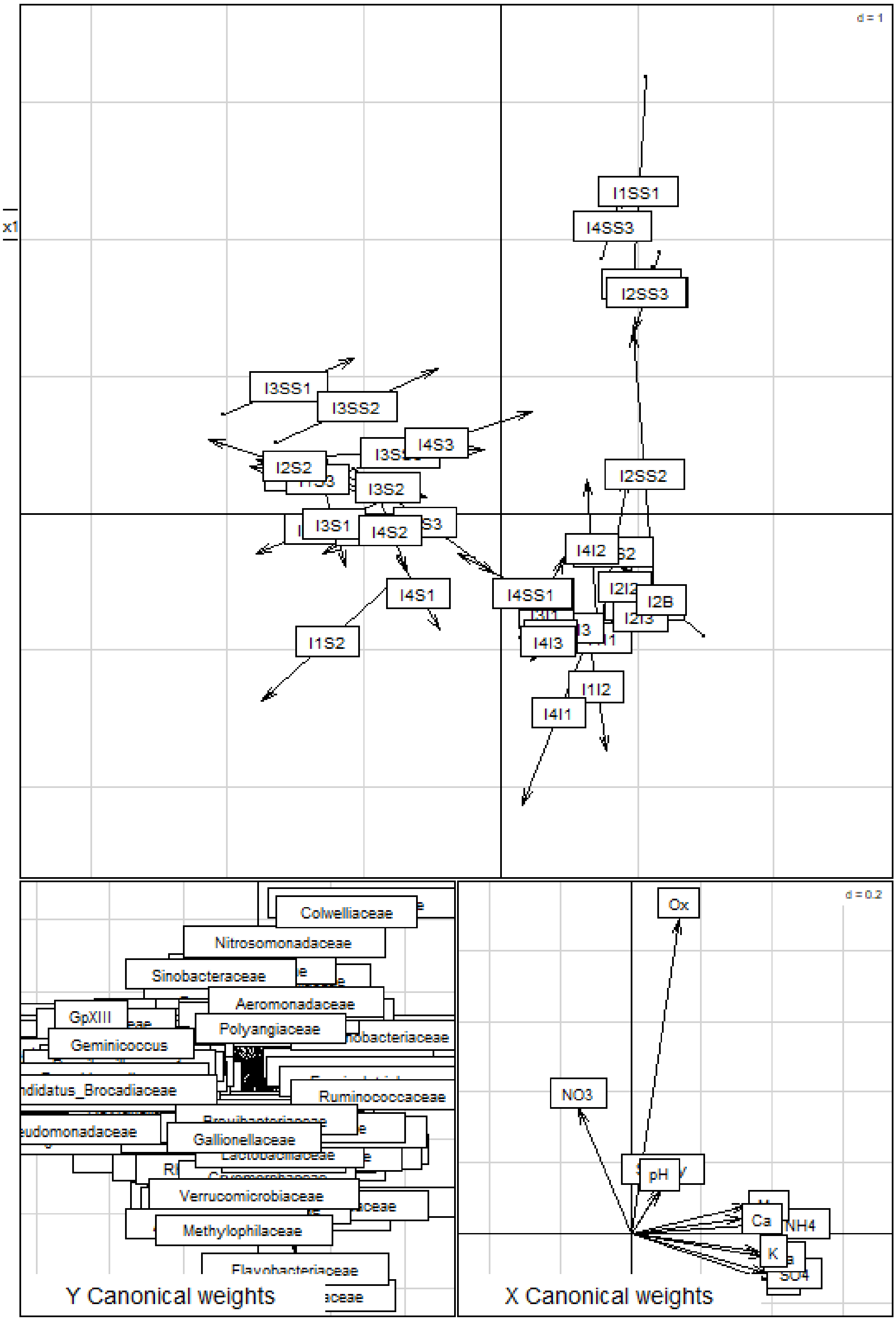


Figure S4.3: (A) Co-inertia graph for the ASV's agglomerated at Family level versus the chemistry data generated by ade4 R package. Arrows represent co-variance scores between the relative. (B) and (C) are the canonical weights (coefficients of the combinations of the variables) of the ASV's and chemical variables to define the coinertia axes, respectively.

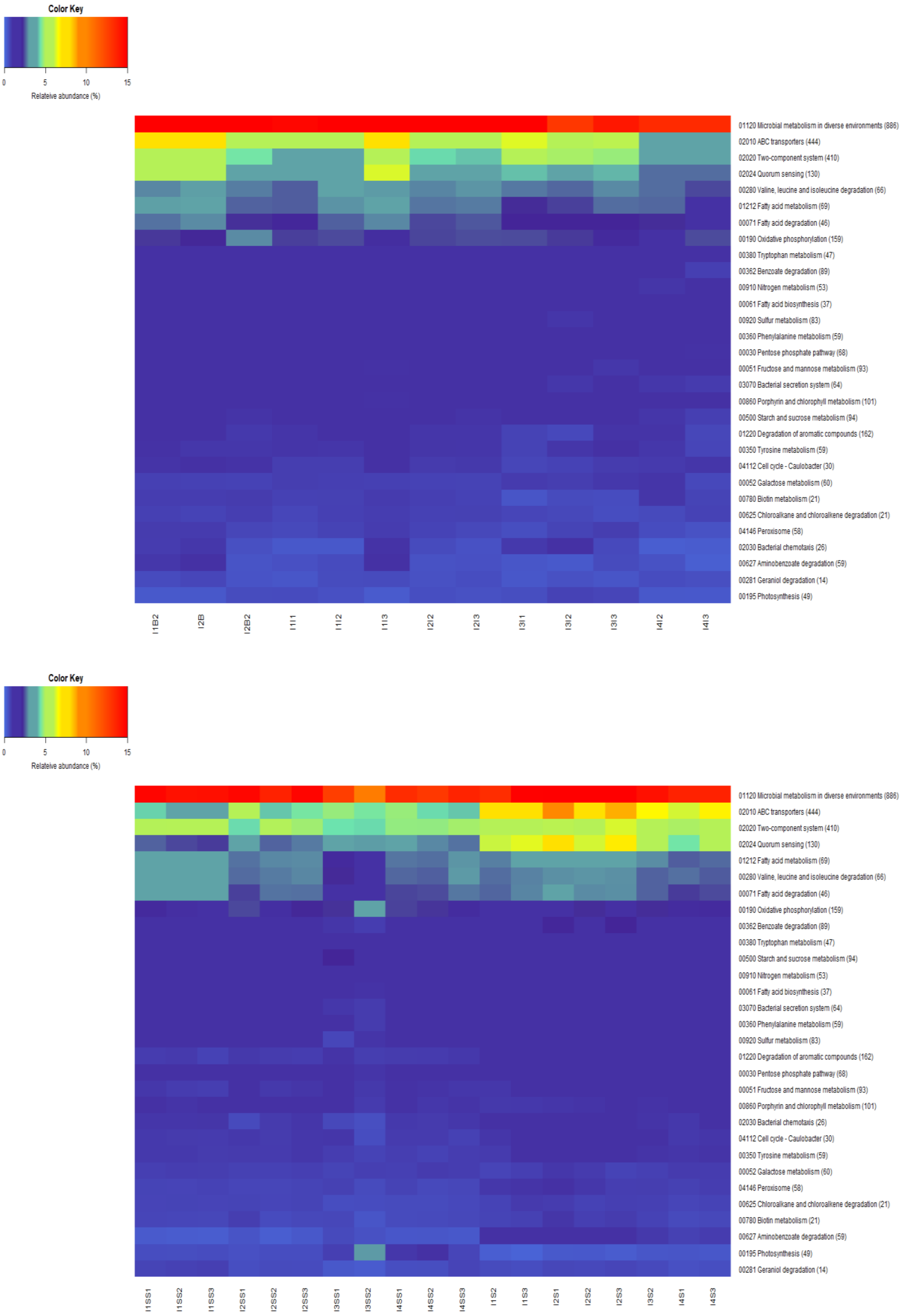


Figure S4.4: Heatmap displaying the KEGG pathways representing the 30% most abundant with the 60% best environmental fit based on an NMDS (using the ordiselect function in the R package vegan) of the square root relative abundances using Bray-Curtis distances. (A) Includes all ice core and brine samples and (B) includes all snow horizons.

4.4 References

Arrigo, K. R., Perovich, D. K., Pickart, R. S., Brown, Z. W., van Dijken, G. L., Lowry, K. E., … Swift, J. H. (2014). Phytoplankton blooms beneath the sea ice in the Chukchi sea. *Deep-Sea Research Part II: Topical Studies in Oceanography*, *105*, 1–16. https://doi.org/10.1016/j.dsr2.2014.03.018

Assmy, P., Fernández-Méndez, M., Duarte, P., Meyer, A., Randelhoff, A., Mundy, C. J., … Granskog, M. A. (2017). Leads in Arctic pack ice enable early phytoplankton blooms below snow-covered sea ice. *Scientific Reports*, *7*(1), 1–9. https://doi.org/10.1038/srep40850

Becker, B., & Marin, B. (2009). Streptophyte algae and the origin of embryophytes. *Annals of Botany*, *103*(7), 999–1004. https://doi.org/10.1093/aob/mcp044

Bowman, J. S. (2015, October 13). The relationship between sea ice bacterial community structure and biogeochemistry: A synthesis of current knowledge and known unknowns. *Elementa*, Vol. 3. https://doi.org/10.12952/journal.elementa.000072

Bowman, J. S., Rasmussen, S., Blom, N., Deming, J. W., Rysgaard, S., & Sicheritz-Ponten, T. (2012). Microbial community structure of Arctic multiyear sea ice and surface seawater by 454 sequencing of the 16S RNA gene. *ISME Journal*, *6*(1), 11–20. https://doi.org/10.1038/ismej.2011.76

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. *Nature Methods*, *13*(7), 581–583. https://doi.org/10.1038/nmeth.3869

Gradinger, R., & Zhang, Q. (1997). Vertical distribution of bacteria in Arctic sea ice from the Barents and Laptev Seas. *Polar Biology*, *17*(5), 448–454. https://doi.org/10.1007/s003000050139

Harding, T., Jungblut, A. D., Lovejoy, C., & Vincent, W. F. (2011). Microbes in high arctic snow and implications for the cold biosphere. *Applied and Environmental Microbiology*, *77*(10), 3234–3243. https://doi.org/10.1128/AEM.02611-10

Hatam, I., Charchuk, R., Lange, B., Beckers, J., Haas, C., & Lanoil, B. (2014). Distinct bacterial assemblages reside at different depths in Arctic multiyear sea ice. *FEMS Microbiology*

Ecology, *90*(1), 115–125. https://doi.org/10.1111/1574-6941.12377

Hatam, I., Lange, B., Beckers, J., Haas, C., & Lanoil, B. (2016). Bacterial communities from Arctic seasonal sea ice are more compositionally variable than those from multi-year sea ice. *ISME Journal*, *10*(10), 2543–2552. https://doi.org/10.1038/ismej.2016.4

Junge, K., Cameron, K., & Nunn, B. (2019). Diversity of Psychrophilic Bacteria in Sea and Glacier Ice Environments—Insights Through Genomics, Metagenomics, and Proteomics Approaches. In *Microbial Diversity in the Genomic Era* (pp. 197–216). https://doi.org/10.1016/b978-0-12-814849-5.00012-5

Larose, C., Dommergue, A., & Vogel, T. M. (2013, February 5). The dynamic Arctic snow pack: An unexplored environment for microbial diversity and activity. *Biology*, Vol. 2, pp. 317–330. https://doi.org/10.3390/biology2010317

Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M., & Larose, C. (2019). Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-38744-y

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, *17*(1), 10. https://doi.org/10.14806/ej.17.1.200

Maslanik, J., Stroeve, J., Fowler, C., & Emery, W. (2011). Distribution and trends in Arctic sea ice age through spring 2011. *Geophysical Research Letters*, *38*(13). https://doi.org/10.1029/2011GL047735

McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS ONE*, *8*(4), e61217. https://doi.org/10.1371/journal.pone.0061217

Olsen, L. M., Laney, S. R., Duarte, P., Kauko, H. M., Fernández-Méndez, M., Mundy, C. J., … Assmy, P. (2017). The seeding of ice algal blooms in Arctic pack ice: The multiyear ice seed repository hypothesis. *Journal of Geophysical Research: Biogeosciences*, *122*(7), 1529–1548. https://doi.org/10.1002/2016JG003668

Rysgaard, S., & Glud, R. N. (2004). Anaerobic N2 production in Arctic sea ice. *Limnology and Oceanography*, *49*(1), 86–94. https://doi.org/10.4319/lo.2004.49.1.0086

Rysgaard, S., Glud, R. N., Sejr, M. K., Blicher, M. E., & Stahl, H. J. (2008). Denitrification activity and oxygen dynamics in Arctic sea ice. *Polar Biology*, *31*(5), 527–537. https://doi.org/10.1007/s00300-007-0384-x

Vancoppenolle, M., Meiners, K. M., Michel, C., Bopp, L., Brabant, F., Carnat, G., … van der Merwe, P. (2013). Role of sea ice in global biogeochemical cycles: Emerging views and challenges. *Quaternary Science Reviews*, *79*, 207–230. https://doi.org/10.1016/j.quascirev.2013.04.011

Vonnahme, T. R., Dietrich, U., & Hassett, B. T. (2020). Progress in Microbial Ecology in Ice-Covered Seas. In *YOUMARES 9 - The Oceans: Our Research, Our Future* (pp. 261–277). https://doi.org/10.1007/978-3-030-20389-4_14

Wadhams, P. (2000). *Ice in the ocean*.

Weeks, W. F., & Ackley, S. F. (1986). The Growth, Structure, and Properties of Sea Ice. In *The Geophysics of Sea Ice* (pp. 9–164). https://doi.org/10.1007/978-1-4899-5352-0_2

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Retrieved from https://ggplot2.tidyverse.org

Williamson, C. J., Cameron, K. A., Cook, J. M., Zarsky, J. D., Stibal, M., & Edwards, A. (2019). Glacier Algae: A Dark Past and a Darker Future. *Frontiers in Microbiology*, *10*(APR), 524. https://doi.org/10.3389/fmicb.2019.00524

Yergeau, E., Michel, C., Tremblay, J., Niemi, A., King, T. L., Wyglinski, J., … Greer, C. W. (2017). Metagenomic survey of the taxonomic and functional microbial communities of seawater and sea ice from the Canadian Arctic. *Scientific Reports*, *7*(1), 42242. https://doi.org/10.1038/srep42242

Chapter 5: Microbial adaptation to saline snow using metagenomic assembled genomes

Layton R, Vogel T.M, Larose C.

5.0 Introduction

One of the largest global interface ecosystems is the snowpack, which extends to over a third of the Earth's land surface, covering up to 47 million km^2. For most of the year, sea ice is typically covered by snow; slow accumulation starts in October, when new sea ice is formed and continues through the winter months. The depth and characteristics of the snowpack influence the physiochemical and biological conditions of the underlying sea ice and are predicted to change as a consequence of global warming. For instance, snow meltpools that form on the surface of sea ice during the Spring increase heat absorption resulting in faster ice melt (Liu, Zhang, Cheng, & Hu, 2019). Additionally, the depth of the snowpack controls the amount of light reaching photosynthetic organisms in the icepack and the underlying seawater and in turn affects primary productivity (Tedesco & Vichi, 2014). Recently, the snowpack itself has been proposed as functional ecosystem for microorganisms, is influenced by interface environments and undergoes post-depositional selection (Larose et al., 2010; Larose, Dommergue, & Vogel, 2013; Larose, Prestat, et al., 2013; Maccario, Carpenter, Deming, Vogel, & Larose, 2019).

We explored the taxonomic and functional profiles of snow overlying sea ice in chapters 3 and 4. The taxonomy of the saline snow directly overlying first year ice (FYI) was consistent between sites and previous findings (Maccario *et al*., 2019). The seeding source for microbes inhabiting this saline snow was presumed to be the upwelling of brine from the underlying sea ice as this brine contains numerous predominant taxa with adaptive qualities suitable to icy and saline habitats. However, saline snow consistently harbours a low diversity community due to the overwhelming dominance of the genera *Colwellia* and *Glacieocola* and independent of sampling time or location (Maccario et al., 2019; Chapters 3 & 4). These two taxa constitute up to 99% combined abundance of saline snow communities despite a seeding community of diverse halotolerant, copiotrophic, psychrophillic and psychrotolerant organisms (Maccario et al., 2019; Chapters 3 & 4). The underlying selection mechanisms that govern the apparent convergent assembly of microbial communities in saline snow is explored here.

To access possible functional mechanisms for dominance of *Colwellia* and *Glacieocola*, we used a comparative genomics approach that can identify genomic and functional differences among bacterial populations across different taxonomic levels. However, until recently environmental comparative genomics has been limited by the requirement to obtain pure cultures in order to obtain fully sequenced their genomes. The development of sequencing technologies and bioinformatics tools circumvents some of the problems inherent with culture dependent techniques by facilitating recovery of genome sequences from metagenomic data. In this case, we used these environmental sequences to assemble sequences (contigs) that could be part of individual genomes (metagenomics assemble genomes – MAGs). Therefore, this preliminary study implemented metagenomic approaches followed by comparative genome analyses in order to better understand which functional traits enable colonisation and dominance in saline snow environments. This saline snow layer combines high salinity and high photochemical activity during daylight months. While Colwellia and Glacieocola are not particularly known for their resistance to photochemical activity, they are generally halotolerant, so we expect that the strains here would have acquired the photochemical resistance. We sought to identify the extent to which genomic differences among closely related taxa could be related to their environmental niche adaptation by comparing our assembled individual MAG's to the other *Colwellia* and *Glaciecola* genomes.

5.1 Methods

5.1.1 Description of samples and genomes used in this study

Table 5.1: Statistics of MAG's resolved in this study

MAG	Completion (%)	Redundancy (%)	N50	Number of contigs	Size (Mb)	GC content (%)
Colwellia	98.6	1.4	4665	1002	3.5	38.8
Glaciecola	80.3	8.5	1945	1494	2.8	41

We acquired 49 metagenomic datasets from two studies undertaken in this thesis and one published study from our lab (Maccario *et al*., 2019). All samples used in this study represent either (brine-wetted) saline snow overlying first-year ice (FYI), FYI cores, brine or the underlying seawater from Greenland (**Table S5.1**). Sampling strategies, DNA extraction and sequencing are all comparable; details can be found in Chapter 2, Chapter 3 and Maccario *et al*. (2019).

5.1.2 Assembly and Binning

Dynamic trimming and primer removal from metagenomic libraries was conducted using Trimmomatic (Bolger, Lohse, & Usadel, 2014) based on their quality scores (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) and read quality profiles were checked in FastQC (Andrews et al., 2012). Co-assemblies were performed using saline snow metagenomes and Megahit (D. Li et al., 2016) using the preset flag “meta-sensitive”. Bowtie2 (Langmead & Salzberg, 2012) was implemented to map reads from all metagenomes to the contigs produced in the previous step by Megahit. The recruited reads were stored as BAM files using samtools (H. Li et al., 2009). Anvi’o was used to estimate coverage and detection statistics of each contig (>1000 bp) that were subsequently exported to enable binning with CONCOCT. Anvi’o (Eren et al., 2015) was used to create a contigs database and combine mapping profiles into a merged profile database inclusive of binning results. Anvi’o was utilised in conjunction with Prodigal (Hyatt et al., 2010) to detect protein coding genes in each contig, HMMER (Eddy, 2011) to identify genes matching bacteria single-copy core gene collections and centrifuge to infer taxonomy of genes. Initial binning results were investigated to identify putative *Glaciecola* and *Colwellia* MAG’s, which were manually refined using anvi’o’s interactive interface. Bins with a putative annotation of *Colwellia, Glaciecola* were used for downstream pangenomic analysis.

5.1.3 Cultivar Genomes

Cultivar genomes used in this study can be found in **Table S5.2.** Genome sequence files were downloaded from the NCBI database and imported into anvi’o. Protein coding and single copy genes were detected using the strategies described in section 5.1.2.

5.1.4 Pangenomic analysis

The genome bins and cultivar genomes were used to generate an anvi'o genome storage and the pangenomic analysis run using the anvio function “anvi-pan-genome”. The main purpose of this function is to identify gene clusters and the details of the program can be found in (Delmont & Eren, 2018). Core and singleton gene clusters were identified using the anvi'o interactive interface. Singleton gene clusters identified in MAGs were annotated using EGGNOGmapper (Huerta-Cepas et al., 2016).

5.2 Results

5.2.1 Metagenome Assembled Genomes

The shotgun (metagenomics) sequencing of DNA from 11 saline snow samples resulted in a total of 9,795,523 quality-filtered 2 × 250 paired-end metagenomic reads. By co-assembling the saline snow samples, we recovered 694,954 contigs that were longer than 1 kbp and organized them into 88 genomic bins at various levels of completion (**Table S5.3**). There were two bins putatively annotated as *Colwellia* and *Glaciecola* with completion, redundancy and size statistics sufficient (completeness greater than 80% and redundancy less than 10%) designated as MAG's. Manual refinement using the anvi'o platform resulted in two refined MAG's for downstream analysis (**Table 5.1**). Figure 1 reports the detection and mean coverage statistics for each MAG. The coverage of *Glaciecola* and *Colwellia* MAG were highest in the saline snow metagenomes (**Figure 5.1**). The largest categories of gene annotations for both MAG's are shown in **Table S5.4 and S5.5**.

5.2.2 The pangenome *Colwellia*

A total of 11,557 gene clusters representing 35,562 gene calls contributed to the *Colwellia* pangenomic analysis (Figure 2). The core genes, defined as gene clusters found in all genomes analysed, were represented by 1,444 gene clusters totalling 12.5% of all clusters (13,584 genes). The *Colwellia* MAG assembled from our saline snow metagenomes harboured 423 singleton gene clusters (456 gene calls). We define singletons as gene clusters associated with a single genome and not found in any available genome. Annotation of these gene calls demonstrated that the largest portion (n=69) were annotated with the

COG functional category S (function unknown; Table S5.6). However, a large proportion (n=37) were denoted with the COG functional category T, which is related to “signal Transduction” and includes universal stress response proteins and histidine kinases. The next largest categories were P (n=24) and L (n=23) and G (n=22) corresponding to “inorganic ion transport and metabolism”, “replication and repair” and “carbohydrate metabolism and transport”. Of the singleton gene clusters identified by the pangenomic analysis, the contig “salinesnow__000000311699” encoded a number of unique gene functions involved in carotenoid biosynthesis and bacteriorhodopsin (Figure 5.3).

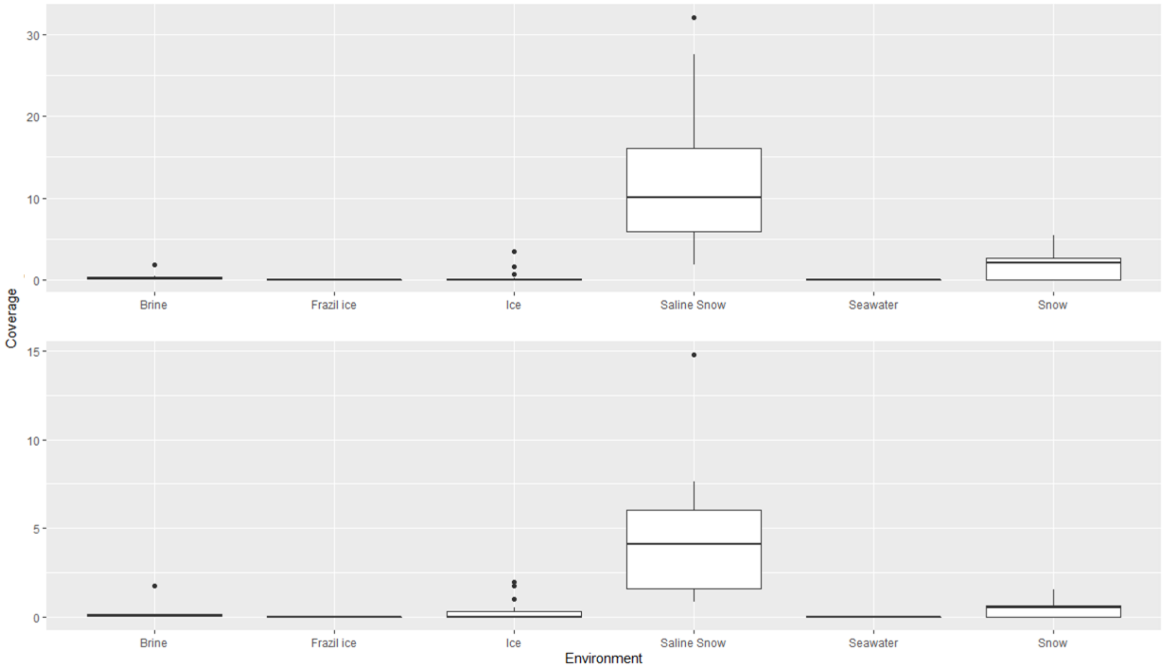


Figure 5.1: Coverage estimates for (A) *Glaciecola* sp. and (B) *Colwellia* sp. MAGs

5.2.3 The pangenome *Glaciecola*

A total of 10,567 gene clusters representing 15,517 gene calls contributed to the *Glaciecola* pangenomic analysis (Figure 5.2). There were only 71 core gene clusters (311 gene calls) found across all genomes representing just 0.672% of the total gene clusters. The number of core gene clusters was increased to 1285 (4396 gene calls) or 12.16% of the total gene clusters when *G. amylolitica* (GCF_003856375.1; denoted as *Glaciecola* THG in Figure 5.3) was excluded. The *Glaciecola* MAG assembled from our saline snow metagenomes harboured 491 singleton gene clusters (579 gene calls). Of the gene calls from the singleton gene cluster with an orthology assignment (n = 421), 380 could be annotated with either a

COG function or GO term. The largest portion were annotated as COG functional category S (function unknown) however, all of these were assigned a GO term including those related to gas vesicle formation (Table S6). The next largest COG category was T (signal transduction, n =53) followed by P (inorganic ion transport and metabolism, n = 37) and G (carbohydrate metabolism and transport, n = 31). Rhodopsin genes were identified in both the *Glaciecola* MAG and *G. nitratireducens* (FR1064).

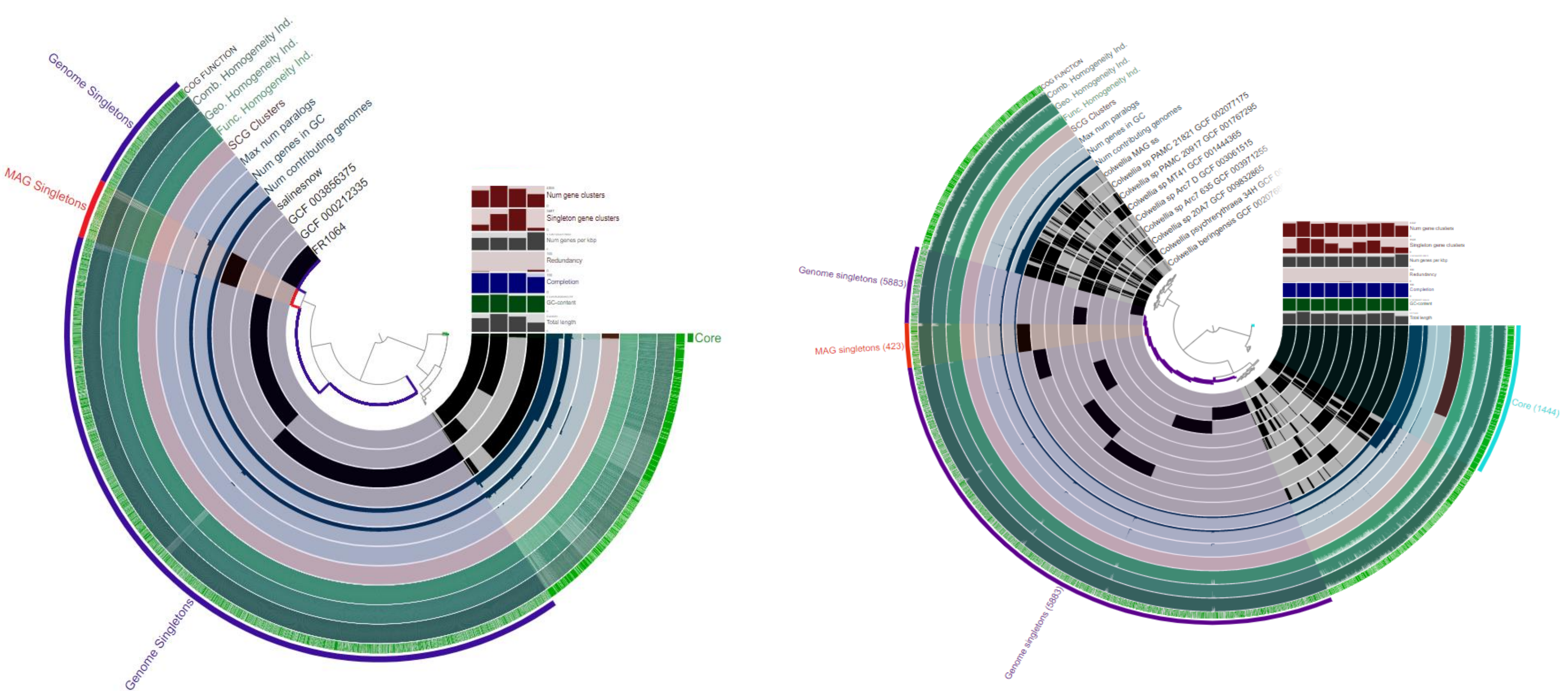


Figure 5.2: (A) Glaciecola and (B) *Colwellia* pangenome analysis of cultivar genomes and draft genomes recovered from saline snow metagenomes. Bars in the first 9 layers indicate the occurrence of a gene cluster in a given isolate genome. The next layers (blue) represent the number of genomes that contain the gene clusters, the number of genes in the gene cluster and the number of paralogs. The next layer, shows whether the gene clusters are single copy genes identified using the anvio function ‘anvi-run-hmms’. The next 3 layers indicate the geometric, functional and combined

5.2 Discussion

The metagenomics approach used here was able to extract to MAG's that provide indications of the microbial adaptation to the saline snow. However, the largest categories of genes in the *Colwellia* saline snow MAG were consistent with those considered globally most abundant in genomes such as transcriptional regulators (LysR and AcrB family), methyl-accepting chemotaxis proteins and outer membrane proteins (Aziz, Breitbart, & Edwards, 2010). However, a number of functional annotations with higher counts relative to all other genomes, included those in this study, may pertain to the unique conditions of saline snow. Snow is photochemical bioreactor in daylight and highly reactive to UV radiation that results in a hyper-oxidative environment (Domine et al., 2008). Specific genes involved in resistance to photochemical stress occurred at higher levels in the *Colwellia* MAG relative to all other genomes. These genes included those involved in DNA repair (*e.g.* exinuclease UvrABC ATPase, 5'3' exonuclease and DNA mismatch repair ATPase MutS) and enzymatic antioxidants (*e.g.* catalase). Additionally, the *Colwellia* and *Glaciecola* MAG harboured more genes annotated as, choline-glycine betaine transporter and choline dehydrogenase which are involved in the uptake of choline and catalyses the conversion of choline into the intermediate glycine betaine aldehyde leading to the formation of the osmoprotectant glycine betaine, respectively. Glycine betaine is a compatible solute with the ability to restore and maintain osmotic balance of living cells under stress and, hence, can be useful under saline conditions.

The multidrug efflux pump (subunit AcrB) was predicted to occur 20 and 18 times in the *Colwellia* and Glaciecola MAG, respectively and as few as twice in the other genomes. While multidrug efflux pumps are well known for their involvement in antibiotic resistance they also provide heavy metal resistance, regulate biofilm formation and quorum sensing. Although it is difficult to ascertain their exact function, free living microorganisms tend to have a larger repertoire of MDR pumps that are likely to contribute to adaptation to different environments by responding to assorted signals (Martinez et al., 2009). Further research would be required to elucidate their contribution to the colonisation of saline snow however, their increased occurrence in our MAG's relative to closely related species hint at a role.

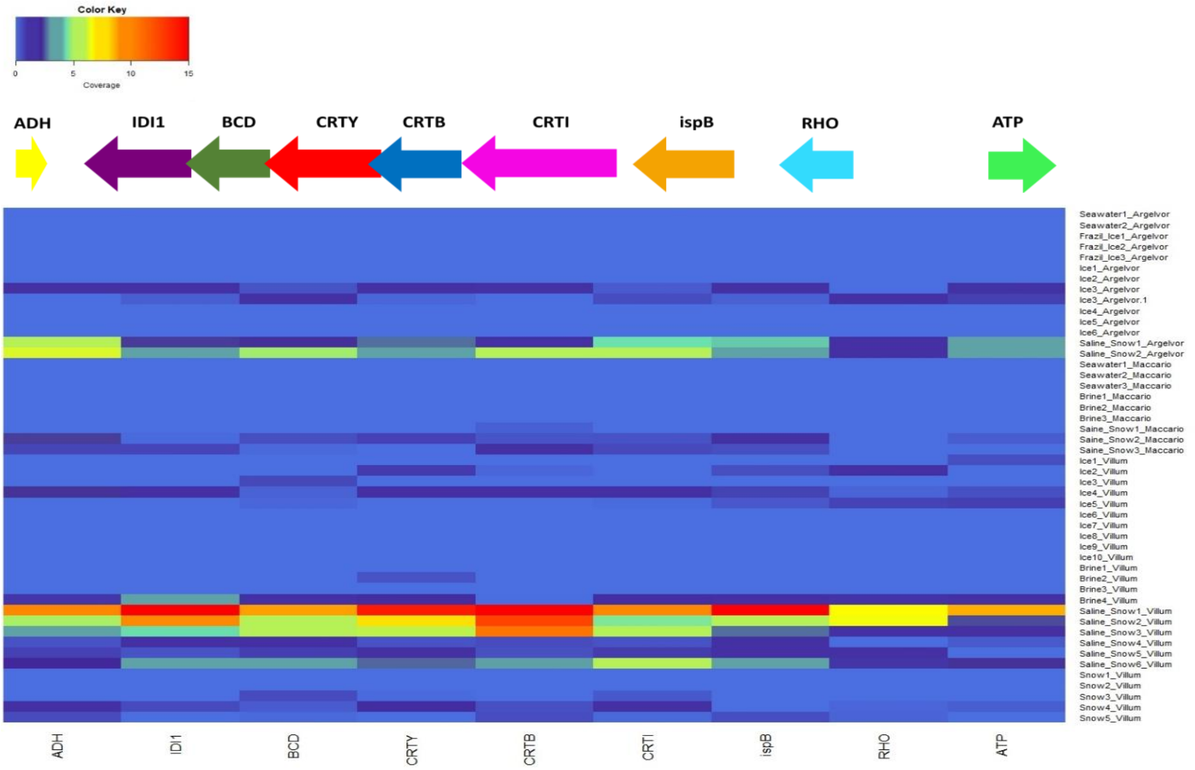


Figure 5.3: Organisation of contig “salinesnow__000000311699” encoding (from left to right) alcohol dehydrogenase, dimethylallyl diphosphate, beta-carotene 15,15'-dioxygenase, lycopene cyclase, phytoene synthase, phytoene, polyprenyl synthetase, bacteriorhodopsin and an ATPase. The heatmap displays the coverage values of this contig across all metagenomes

In heterotrophic bacteria, carotenoids are secondary metabolites associated with cell adaptability. Not only can carotenoids protect cells from UV damage and oxidative stress but they are implicated in the maintenance of membrane fluidity (Croce & Van Amerongen, 2014; Maoka, 2020; Seel et al., 2020). There is a high number of carotenoid-producing species identified in cold regions suggesting these pigments play an important role in the adaptation to this ecological niche (Vila, Hornero-Méndez, Azziz, Lareo, & Saravia, 2019). We identified a number of gene clusters unique to the *Colwellia* MAG corresponding to carotenoid biosynthesis. Indeed, the complete pathway for the production of 7,8-Dihydro-beta-carotene from the products of terpenoid backbone biosynthesis were encoded on a single contig from the *Colwellia* MAG. The closest homolog to one of the genes on this contig was predicted as a probable beta-carotene 15,15'-dioxygenase which converts betacarotene to yield two molecules of all-trans retinal. Interestingly, the gene encoding bacteriorhodopsin was found downstream of the carotenoid biosynthesis genes and adjacent to an ATPase gene. Bacteriorhodopsin is a light-driven proton pump which relies on an initial conformational change of retinal and eventually drives proton transport (Pinhassi, DeLong, Béjà, González, & Pedrós-Alió, 2016). The proton gradient which is formed can subsequently be used to generate chemical energy by ATP synthase. Bacteriorhodpsin was also predicted to be encoded in the *Glaciecola* MAG on the same contig as geranylgeranyl pyrophosphate synthase, an enzyme involved in synthesising a precursor to carotenoid biosynthesis. PR-bearing bacteria have been previously identified in the sea ice, but they are generally associated with the ice-water interface (Koh et al., 2010; Vader, Laughinghouse, Griffiths, Jakobsen, & Gabrielsen, 2018). However, we identified a greater coverage of these genes in the saline snow horizons relative to any of the underlying environments.

Within the singleton gene clusters, both MAG's encoded a number of genes related to gas vesicle formation. Gas vesicles are gas permeable membrane bound organelles and form the major component of gas vacuoles. Not only do these vesicles provide buoyancy and in turn selection of favourable environmental conditions but they have a light shielding capacity and can provide a larger surface to volume ratio for nutrient acquisition in order to adapt to high UV and low temperatures, respectively (Odaneth, 2018). Gas vacuoles have been associated with bacteriorhodopsins in some *Halobacterium* and may facilitate movement towards higher light intensities to drive the light driven proton pump (Oren, 2012).

5.4 Perspectives

Specific genes and singleton gene clusters were identified in MAG's assembled from saline snow samples which potentially pertain to adaptations unique to the environment including the use of light-driven bacteriorhodopsins. However, many of the functions associated in the singleton gene clusters were also found across the reference genomes suggesting that many functions are conserved. Nevertheless, the discrepancy in sequence homology, which defines the gene clusters, could confer variations in functioning that are beneficial under the unique conditions of saline snow. For instance, ABC transporters are one of the largest and most widespread groups of bacterial proteins but they are involved in a diverse range of functions corresponding to sequence variation. As such, further analysis will involve multiple sequence alignments of reference and metagenome recovered sequences in order to infer the functional manifestation of sequence differences. Additionally, we endeavour to resolve draft genomes for members of *Colwellia*, *Glaciecola* and others from the underlying ice, brine and seawater metagenomes. Research within this thesis suggests there may be a species or even strain specific enrichment in saline snow and as such these MAGs will serve as a comparison to their counterparts. Moreover, coverage values of recovered MAG's suggest differences in species or strains across sites and indicate the need to assemble by location in order to resolve difference in genome structure. This line of analysis will provide a better understanding of their abundance in saline snow and in turn the ecological drivers of this habitat.

5.5 References

Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., & Wingett, S. (2012, January). *FastQC*. Babraham, UK.

Aziz, R. K., Breitbart, M., & Edwards, R. A. (2010). Transposases are the most abundant, most ubiquitous genes in nature. *Nucleic Acids Research*, *38*, 4207–4217.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Croce, R., & Van Amerongen, H. (2014). Natural strategies for photosynthetic light harvesting. *Nature Chemical Biology*, Vol. 10, pp. 492–501. https://doi.org/10.1038/nchembio.1555

Delmont, T. O., & Eren, E. M. (2018). Linking pangenomes and metagenomes: The Prochlorococcus metapangenome. *PeerJ*, *2018*(1). https://doi.org/10.7717/peerj.4320

Domine, F., Albert, M., Huthwelker, T., Jacobi, H.-W., Kokhanovsky, A. A., Lehning, M., … Simpson, W. R. (2008). Snow physics as relevant to snow photochemistry. *Atmospheric Chemistry and Physics*, *8*(2), 171–208. https://doi.org/10.5194/acp-8-171-2008

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Computational Biology*, *7*(10), e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Eren, A. M., Esen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: An advanced analysis and visualization platformfor 'omics data. *PeerJ*, *2015*(10), e1319. https://doi.org/10.7717/peerj.1319

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M. C., … Bork, P. (2016). eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Research*, *44*(D1), D286–D293. https://doi.org/10.1093/nar/gkv1248

Hyatt, D., Chen, G. L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, *11*, 119. https://doi.org/10.1186/1471-2105-11-119

Koh, E. Y., Atamna-Ismaeel, N., Martin, A., Cowie, R. O. M., Beja, O., Davy, S. K., … Ryan, K. G. (2010). Proteorhodopsin-bearing bacteria in Antarctic sea ice. *Applied and Environmental Microbiology*, *76*(17), 5918–5925. https://doi.org/10.1128/AEM.00562-10

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*(4), 357–359. https://doi.org/10.1038/nmeth.1923

Larose, C., Berger, S., Ferrari, C., Navarro, E., Dommergue, A., Schneider, D., & Vogel, T. M. (2010). Microbial sequences retrieved from environmental samples from seasonal Arctic snow and meltwater from Svalbard, Norway. *Extremophiles*, *14*(2), 205–212. https://doi.org/10.1007/s00792-009-0299-2

Larose, C., Dommergue, A., & Vogel, T. M. (2013, February 5). The dynamic Arctic snow pack: An unexplored environment for microbial diversity and activity. *Biology*, Vol. 2, pp. 317–330. https://doi.org/10.3390/biology2010317

Larose, C., Prestat, E., Cecillon, S., Berger, S., Malandain, C., Lyon, D., … Vogel, T. M. (2013). Interactions between snow chemistry, mercury inputs and microbial population dynamics in an arctic snowpack. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0079972

Li, D., Luo, R., Liu, C. M., Leung, C. M., Ting, H. F., Sadakane, K., … Lam, T. W. (2016, June 1). MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods*, Vol. 102, pp. 3–11. https://doi.org/10.1016/j.ymeth.2016.02.020

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Liu, J., Zhang, Y., Cheng, X., & Hu, Y. (2019). Retrieval of Snow Depth over Arctic Sea Ice Using a Deep Neural Network. *Remote Sensing*, *11*(23), 2864. https://doi.org/10.3390/rs11232864

Maccario, L., Carpenter, S. D., Deming, J. W., Vogel, T. M., & Larose, C. (2019). Sources and selection of snow-specific microbial communities in a Greenlandic sea ice snow cover. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-38744-y

Maoka, T. (2020, January 1). Carotenoids as natural functional pigments. *Journal of Natural Medicines*, Vol. 74, pp. 1–16. https://doi.org/10.1007/s11418-019-01364-x

Martinez, J. L., Sánchez, M. B., Martínez-Solano, L., Hernandez, A., Garmendia, L., Fajardo, A., & Alvarez-Ortega, C. (2009, March 1). Functional role of bacterial multidrug efflux pumps in microbial natural ecosystems. *FEMS Microbiology Reviews*, Vol. 33, pp. 430–449. https://doi.org/10.1111/j.1574-6976.2008.00157.x

Odaneth, A. (2018). Understanding Gas Vesicles and Its Scope in Biotechnological Applications. *Advances in Biotechnology & Microbiology*, *11*(2), 1–2. https://doi.org/10.19080/aibm.2018.11.555806

Oren, A. (2012, December 27). The function of gas vesicles in halophilic archaea and bacteria: Theories and experimental evidence. *Life*, Vol. 3, pp. 1–20. https://doi.org/10.3390/life3010001

Pinhassi, J., DeLong, E. F., Béjà, O., González, J. M., & Pedrós-Alió, C. (2016). Marine Bacterial and Archaeal Ion-Pumping Rhodopsins: Genetic Diversity, Physiology, and Ecology. *Microbiology and Molecular Biology Reviews*, *80*(4), 929–954. https://doi.org/10.1128/mmbr.00003-16

Seel, W., Baust, D., Sons, D., Albers, M., Etzbach, L., Fuss, J., & Lipski, A. (2020). Carotenoids are used as regulators for membrane fluidity by Staphylococcus xylosus. *Scientific Reports*, *10*(1), 1–12. https://doi.org/10.1038/s41598-019-57006-5

Tedesco, L., & Vichi, M. (2014). Sea Ice Biogeochemistry: A Guide for Modellers. *PLoS ONE*, *9*(2), e89217. https://doi.org/10.1371/journal.pone.0089217

Vader, A., Laughinghouse, H. D., Griffiths, C., Jakobsen, K. S., & Gabrielsen, T. M. (2018). Proton-pumping rhodopsins are abundantly expressed by microbial eukaryotes in a high-Arctic fjord. *Environmental Microbiology*, *20*(2), 890–902. https://doi.org/10.1111/1462-2920.14035

Vila, E., Hornero-Méndez, D., Azziz, G., Lareo, C., & Saravia, V. (2019). Carotenoids from heterotrophic bacteria isolated from Fildes Peninsula, King George Island, Antarctica. *Biotechnology Reports*, *21*, e00306. https://doi.org/10.1016/j.btre.2019.e00306

6.0 Conclusion and perspectives

The snowpack-sea ice-ocean continuum represents an interesting ecosystem for studying colonisation and adaptation owing to its spatial and temporal heterogeneity and steep environmental gradients. Microbial assembly in sea ice are affected by the ability to transition from an open seawater environment (colonisation potential) and selection by new physiochemical parameters. Sea ice and snowpacks are characterised as dynamic environments that transition from full darkness to high UV irradiation, salinity values of up to x PPB or as low as X ppB and transient nutrient and contaminant inputs. In turn, we find a stratification of environmental gradients from the surface to the underlying ocean as well as a partitioning of physiochemical parameters such as those seen between the liquid brine fraction and the surrounding ice matrix or between MYI and FYI. These phases constitute different biological niches which may drive niche selection toward a specific community. Modern sequencing technologies and bioinformatic tools permit the exploration of the functional potential of microbial communities and assembly of genomes without the need to culture. While some studies have investigated the stratification of taxonomy within sea ice profiles, this has yet to be linked to functional potential and lifestyle insights from metagenomic assembled genomes (discussed in chapter 1) have yet to be accessed.

We investigated both the taxonomic and functional distribution of sea ice and snow microbial communities using microcosm and field based experiments in combination with molecular approaches (i.e qPCR, high throughput sequencing technologies). Our main objective was to contribute to answering questions such as “How do microorganisms colonise the sea ice?”, “Do microorganisms physically and metabolically partition in seawater, sea ice, snow and frost flowers?”, “What drives the physical partitioning of microbial communities in this continuum?” and “How does sea ice affect the microbial assembly of connected environments (i.e. frost flowers and saline snow)?”

Our work highlighted the significant contribution of niche based selection, even in the early stages of sea ice formation (chapter 2). Not only was there a taxonomic and functional partitioning of sea ice, frost flowers and the seeding seawater, but sea ice formed from seawater not usually subject to annual formation selected for archetypal sea ice bacteria. While previous studies have hypothesized a link between bacterial enrichment in sea ice and algal attachment, the low abundance of diatoms yet high bacterial loads in our artificial sea ice demonstrate an alternative mechanism of bacterial enrichment (chapter 2). This observation was reflected in natural sea ice profiles in which heterotrophic flagellate populations dominated and eukaryotic reads constituted a small proportion of the metagenomes (chapter 3 & 4). Newly formed frazil ice does not differ from the seawater in terms of bacterial community structure (chapter 3) and so processes that result in unique assembly are likely driven by post-consolidation selection mechanisms.

The factors that drive functional and taxonomic selection in the sea-ice, seawater and snow continuum have been linked to a number of genetic traits but only a few studies have integrated metagenomic approaches. Functional signatures linked to motility, chemotaxis, response to UV and genetic transfer are all found in greater abundance in the saline snow and some ice horizons relative to the seawater (chapter 3). To date, there are no representative metagenomes of winter/autumn ice. This approach could help us understand the major structuring factors and how they differ across seasons given the taxonomic changes already observed.

MYI is being rapidly replaced by FYI as the dominant sea ice type in the Arctic. As a consequence, the spatial and temporal extent of frost flowers and brine fed snow is predicted to increase. This thesis provides a novel insight into the microbial communities of these increasingly important ecosystems. In both microcosm (frost flowers) and field experiments (saline snow), a taxa specific enrichment was observed. Interestingly, independent of time and space, the basal snow horizons overlying FYI selected for an overwhelming dominance of *Colwellia* and *Glaciecola* (up to 99% relative abundance). This observation was not mirrored in the snow overlying MYI and LFI where brine channels were not evident. While the influence of brine on taxonomic structure of snow was striking, the underlying mechanisms could not be clearly linked to functional signatures in the metagenomes. As genomes can provide insight into the metabolic properties, functionality, biogeochemical cycling and interactions of microorganisms (discussed in chapter 1), we reconstructed *Colwellia* and *Glaciecola* genomes from the available saline snow metagenomes. The identification of a rhodopsin gene linked to an ATP pump in the *Colwellia* MAG that were not found in other *Colwellia* genomes, suggests possible niche selection. In turn, *Colwellia* inhabiting saline snow may gain a growth advantage from light using a proteorhodopsin proton pump. To accurately understand the contribution of proteorhopsin pumps to the acclimation of saline snow, further research would benefit from transcriptomic and proteomic studies which could identify whether the genes are actively expressed and translated *in situ*.

One of the central themes in this thesis is the adaptive potential of microorganisms across the sea ice profile. While we have linked functions to prevailing environmental conditions, we are interested in exploring the contribution of acquired adaptive traits relative to inherent traits to colonization of saline snow. We are continuing to recover MAGs from the overlying snow and underlying sea ice and ocean to contrast those recovered from the saline snow. This will enable direct comparisons to the organisms that seed the environment and putatively identify functions which have been acquired. One of the other remaining questions is whether the saline snow community is composed of organisms from the brine or whether nutrient supplementation stimulates the growth of snow microorganisms. Genome reconstruction will help address this issue but microcosm experiments will be key to providing a definitive answer.

Despite increasing the sequencing data available for sea ice mircoorganisms, the field of sea ice and snow microbiology could benefit from the application of modern research tools. While the use of genomics, transcriptomics and proteomics have been used in culture to infer the genetic mechanisms underpinning success of specific sea ice microorganisms, a lack of community data is evident. Further metagenomic studies of sea ice are required to elucidate the genetic potential of microorganisms, especially across seasons. Coupling studies with metatranscriptomics and metaproteomics are required to demonstrate whether the genes are expressed and translated. Ultimately, multi-disciplinary community studies present the logical next step in understanding sea ice, snow and frost flower ecosystems in the face of climate change.