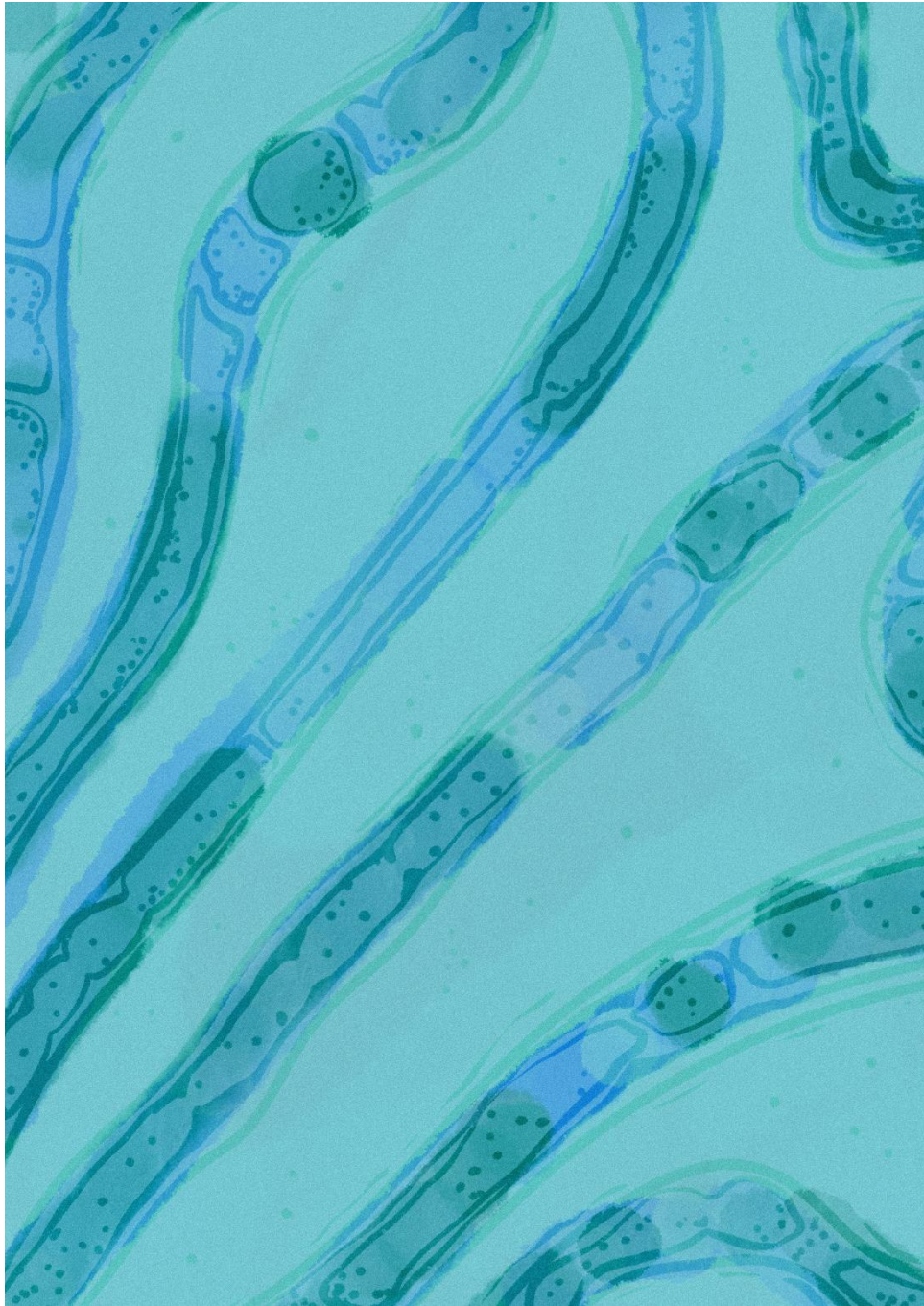




From Biogeography to Genes

Exploring the Diversity and Evolution of Antarctic Cyanobacteria

Benoit Durieu

Cover illustration: *Microbial Beauty* © Kawtar Waddi. This drawing is a free artistic interpretation inspired by our discussions and microbial observations.

From Biogeography to Genes

Exploring the Diversity and
Evolution of Antarctic Cyanobacteria

Benoit Durieu

Promotor

Dr. Annick Wilmotte

Co-promotor

Prof. Dr. Denis Baurain

Thesis submitted in partial fulfilment of the requirements for the degree
of Doctor of Philosophy (Ph.D.) in Sciences: *Biology of Organisms and
Ecology* at the University of Liège

Academic year
2025-2026

University of Liège
Faculty of Sciences
Department of Life Sciences
InBioS – Eukaryotic Phylogenomics
InBioS – Cyanobacterial Molecular Diversity and Ecology



Titre du doctorat en français : De la biogéographie aux gènes : étude de la diversité et de l'évolution des cyanobactéries antarctiques

Copyright © 2026, Benoit Durieu

All rights reserved. No part of this manuscript protected by its copyright notice may be reproduced or utilized in any form or any means, electronic or mechanical, including photocopying, recording or by any information storage or retrieval system without written permission of the author and promoters.

I take full responsibility for the content of this thesis.



Funding: this research was financially supported by a FRIA fellowship from the F.R.S.-FNRS, the Belgian Science Policy Office (BELSPO) Climate Change and Antarctic Microbial Biodiversity (CCAMBIO) project, and University of Liège. BCCM/ULC cyanobacterial collection is thanked for the use of ULC 065, ULC 081, ULC 082, ULC 084, ULC 065 and ULC 007 cyanobacterial strains. PhytoSYSTEMS' *durandal* computing cluster and CÉCI Consortium are also thanked.

Use of AI tools: during the redaction of this Ph.D. thesis, ChatGPT was used as a support tool for grammatical correction and reformulation.

Ph.D. Thesis, Faculty of Sciences, University of Liège, Liège, Belgium

Jury members

Prof. Dr. Monique Carnol (Chair, University of Liège, BE)

Prof. Dr. Denis Baurain (Co-promotor, University of Liège, BE)

Dr. Anne Jungblut (Natural History Museum, London, UK)

Assoc. Prof. Dr. H. Dail Laughinghouse (University of Florida, USA)

Prof. Dr. Nicolas Magain (Secretary, University of Liège, BE)

Assoc. Prof. Dr. Eveline Pinseel (Ghent University, BE)

Dr. Annick Wilmotte (Promotor, University of Liège, BE)

Acknowledgements

Finally, here we are... (sigh of gratitude)

This thesis would never have seen the light without all of you. Annick, thank you for supporting me from the beginning, thank you for your patience. The adventure is now coming to an end, and I wouldn't have had the resources to complete it without you. Thank you for giving me freedom, for sharing your precious knowledge, for offering me outreach activities that I enjoyed, and for giving me resources when it was over. Denis, you gave me new hope at a time when I thought I would never be able to finish this thesis. I will be grateful forever. You are a model mentor, the way you manage the team is remarkable, as is your empathy and availability despite being in high demand from all sides. Yannick, thank you for guiding and helping me from the beginning, without you I simply wouldn't have started! Your curiosity and critical thinking have greatly inspired me, as did your resilience. I also discovered a friend, and a kindred spirit in the fight! Valentina, I could never thank you enough for the help you gave me. If I finally finished, it wouldn't have been possible without you. In addition to being a friend, you are also an inspiration. Your own evolution has always amazed me. You will always be a model. Igor, I can't thank you enough for the help you've given me, both mentally and scientifically. You've clearly been the guide in shaping my research and scientific ideas. I'm so lucky to have a friend like you, with so much empathy and a profound mind. Luc, you arrived in the game like a blessed manager, your pragmatism and ability to stop when it is the good time has been decisive to put a final point to this thesis. For that, and all the other exciting side projects you involved me (even at the end, when I thought I couldn't do all at the same time :p), I thank you so much. Anne-Catherine, you have been a model of intelligence and resilience, your advice in friendship and at work has been and always will be very precious, thank you. Béa, thanks so much for your wise advices and friendship. Mick, I could never thank you enough for the mental and technical help in getting this work to completion. But more importantly, thank you for your friendship, your support, your light. What a team! As you already rightly said, with Loïc and Manu: *"from start to finish together, brothers!"*. A warm thanks to the Bionaute Lab (extended), what a team (again)! Louise, the rising star of the lab, Marie, the budding chief engineer, Valerian, il capo and the king, Edi, the GOAT, Guillaume, my favorite MacGyver, Daniella, the wisdom and quietude, Eline and Harris, we shared so many beautiful moments, Kiril, the brain and the athlete, Alain, the "Q" of the B22, Jerry, Mathieu, strength to you! The friends and colleagues from BCCM/ULC, Kim, Mariano, Haifaa, Manu, Sandeep, your help was crucial. Thanks also for your friendship, wonderful people! The friends from the CIP, Thomas, Alessia, Benoit, Loïc, Loïc and all the amazing folks with whom I began this adventure, thank you so much. Thanks to

the tomato team and friends, Yaprak, Cigdem, Gwendoline, Justine, Emeline. Thank you to my thesis committee, Dr. Georges Feller, Prof. Bernard Joris and Prof. Marc Hanikenne. I thank the members of my jury, Prof. Monique Carnol, Dr. Anne Jungblut, Prof. H. Dail Laughinghouse, Prof. Nicolas Magain and Prof. Eveline Pinseel. My dear (dear) friends who have supported me from the beginning, I could never thank you enough. Alice, Amandine, Antoine, Arnaud, Cécile, Clément, Constance, Élise, Elisa, Fiona, Greg, Jenn, Jens, Jules, Julie, Julien, Justin, Laura, Lennart, Méla, Nadjim, Quentin, Rémi, Romain, Rose, Sandra, Séverine, Sophie, Sylvain, Vincent, Zao, what a beautiful (mafia) family. The ones who joined during the road, Benoît and Benoît, Damien, Juliette, Kawtar, Luco, Nico, Souk, Sophie. I am so happy, and lucky to have you all in my life. Infinite thanks to my loving family, my beloved sister, Pauline, my beloved mom, Sophie, my beloved daddy, Denis, what genes we share! Thank you for making me who I am. To all my family here and in the sky, to all the friends in the heart. Thank you, my love Elisa, we fought together to achieve this!

I am so lucky you are here. This thesis will definitely have been a collective adventure...

Contents

Abstract	1
Graphical abstract.....	3
Résumé	4
Chapter 1 - General Introduction	7
1.1. Antarctica: a remote, at-risk continent with extreme environmental conditions	7
1.2. General presentation of Cyanobacteria	10
1.3. Cyanobacteria in Antarctica	12
1.3.1. Cyanobacterial diversity and adaptation strategies in Antarctica	12
1.3.2. Biogeography of Antarctic cyanobacteria	17
1.4. Genomics of Antarctic cyanobacteria: tools to study biogeography and evolution	20
1.4.1. Amplicon sequencing of cyanobacterial communities	20
1.4.2. Genomics of Antarctic cyanobacteria	21
1.5. From biogeography to genes: main objectives of the thesis and research questions	27
1.5.1. Main objectives	27
1.5.2. Research questions	28
Research chapters - Part one	29
Chapter 2 - (Sub-)Antarctic endemic cyanobacteria from benthic mats are rare and have restricted geographic distributions.....	30
2.1. Abstract	31
2.2. Introduction	31
2.3. Material and Methods	33
2.3.1. Sampling and measurements.....	33
2.3.2. DNA extraction, PCR and sequencing	34
2.3.3. Bioinformatic analysis to obtain quality filtered cyanobacterial OTUs	34
2.3.4. Taxonomic affiliation and geographic distribution of OTUs	35
2.3.5. Analysis of cyanobacterial community structure	36
2.4. Results	37
2.4.1. Cyanobacteria differ between Antarctica and the sub-Antarctic islands, but not between ACBRs	37

2.4.2. The distribution patterns and taxonomic composition differ between the abundant and rare fractions	39
2.5. Discussion	41
2.5.1. Cyanobacterial communities are different between Antarctica and the sub-Antarctic islands	41
2.5.2. Cyanobacterial communities are not clearly differentiated among ACBRs	42
2.5.3. The need for protecting the sub-Antarctic islands	44
2.6. Conclusions.....	45
2.7. References.....	46
2.8. Supplementary information.....	55

Chapter 3 - From single-family dominance to balanced assemblages: substrate-driven soil cyanobacterial communities in the Sør Rondane Mountains, East Antarctica

76

3.1. Abstract	77
3.2. Introduction	77
3.3. Materials & Methods	79
3.3.1. Sampling and site description	79
3.3.2. DNA extraction and sequencing	80
3.3.3. Data processing	81
3.3.4. Environmental data collection.....	82
3.3.5. Statistical analysis	82
3.4. Results	84
3.4.1. Community structure and diversity in relation to substrate and sampling sites	84
3.4.2. Community structure in relation to environmental factors, site and substrate	87
3.5. Discussion	88
3.5.1. Chroococcidiopsaceae dominates marble samples	88
3.5.2. Cyanothecaceae dominates gneiss samples.....	90
3.5.3. Gomontiellaceae dominates moraine samples of the Dry Valley	90
3.5.4. High morphological diversity and taxonomic evenness characterize granitic samples.....	91
3.6. Conclusion	93
3.7. Acknowledgments	94
3.8. Author contributions.....	94
3.9. References.....	95

3.10. Supplementary information.....	104
Research chapters - Part two	122
Chapter 4 - Cold adaptation and stress response of <i>Phormidesmis priestleyi</i> ULC007, a cyanobacterium isolated from Lake Bruehwiler (Larsemann Hills, Antarctica)	123
4.1. Introduction	124
4.2. Genome Announcement	124
4.3. Methods	125
4.3.1. Obtention of good quality assemblies and statistics	126
4.3.2. Cold adaptation and stress response genes	126
4.3.3. Clustering and statistical analysis	128
4.4. Results	128
4.4.1. Cold adaptation and stress response of <i>Phormidesmis priestleyi</i> ULC007 131	
4.5. Discussion	134
4.6. References.....	136
4.7. Supplementary Materials	139
4.8. Genome announcement	141
Chapter 5 - Out of Antarctica: new insights into Antarctic Subcluster 5.2 picocyanobacteria based on high-quality genomes	144
5.1. Abstract	145
5.2. Impact statement	145
5.3. Data summary.....	145
5.4. Introduction	146
5.5. Methods	147
5.5.1. Sampling	147
5.5.2. DNA extraction and de novo sequencing	148
5.5.3. Genome assembly and quality assessment.....	148
5.5.4. Multilocus maximum likelihood phylogeny (phylogenomic analysis)	148
5.5.5. 16S rRNA phylogeny	149
5.5.6. Genome annotations and pangenomic analysis	150
5.5.7. In-depth analysis of specific Antarctic gene clusters and functional enrichment analysis	150
5.6. Results	151
5.6.1. General characteristics of the assemblies.....	151

5.6.2. Antarctic genomes are paraphyletic in the phylogenomic analysis of subcluster 5.2	151
5.6.3. Densification of group D with environmental 16S rRNA sequences confirms the presence of a non-Antarctic lineage between the two Antarctic ones.....	154
5.6.4. Group D Antarctic genomes are also grouped together in the pangenomic analysis and contain unique GCs.....	155
5.6.5. Only two Antarctic gene clusters (GCs) display functional enrichment in the Antarctic genomes	155
5.6.6. In-depth analysis of higBA gene clusters reveals potential pseudogenes in the most-closely related non-Antarctic genomes within group D	156
5.6.7. Functional annotations enriched in Antarctic genomes of the group D ...	157
5.7. Discussion	158
5.7.1. Group D: an Antarctic lineage as an origin for dispersal in non-Antarctic ecological niches?.....	158
5.7.2. The specificities of genomes of the group D: no clear “Antarctic signal” and a pseudogene that could be a signature of ancestry	159
5.8. Conclusion	161
5.9. Author statements	162
5.9.1. Author contributions.....	162
5.9.2. Funding information	162
5.9.3. Acknowledgements	162
5.10. References.....	163
5.11. Supplementary information.....	171
Chapter 6 – General Discussion.....	181
6.1. Biogeographic patterns of Antarctic cyanobacterial communities	182
6.1.1. From continental to local scales	182
6.1.2. Endemism comparison between aquatic and terrestrial communities ...	185
6.2. Characterization of the specificities of Antarctic genomes.....	188
6.2.1. <i>A priori</i> (GOIs) and <i>a posteriori</i> (discovery-based) approaches	188
6.2.2. Antarctic-specific genes: comparison between picocyanobacteria and <i>Phormidesmis</i> groups	190
6.2.3. Determination of the origin of the Antarctic-specific genes.....	194
6.3. Methodological considerations and limitations.....	196
6.3.1. Sampling issues	196
6.3.2. Lack of (meta)data	197
6.3.3. Phylogenetic sampling depth and density, and pangenomic issues	198
6.3.4. Geographic location, habitat types and strain isolation	199

6.3.5. Sources of bias and signal robustness in amplicon sequencing	200
6.4. Final considerations and perspectives	204
6.4.1. Towards a better understanding of Antarctic uniqueness: complementary sampling and analyses	204
6.4.2. How could genomics help strengthen the concept of endemism?	207
6.4.3. Assessing and communicating microbial endemism: strategic tools for nature conservation.....	209
References.....	211
Publications.....	236

Abstract

Antarctica has undergone a long history of geographic and climatic isolation since its separation from Gondwana. Present-day Antarctic biodiversity is likely a mosaic of organisms derived from ancient lineages as well as taxa arrived on the continent at different time periods. Antarctic non-marine ecosystems encompass diverse lacustrine, hydroterrestrial, and terrestrial habitats that differ markedly in water availability and environmental conditions. Organisms inhabiting these ecosystems face multiple and often interacting stressors. The combination of long-term isolation, environmental heterogeneity, and extreme conditions has strongly shaped the diversity, distribution, and evolution of Antarctic biota.

In this context, Cyanobacteria are primary colonizers that facilitate ecosystem development and can constitute the dominant taxa in some Antarctic microbial ecosystems. Despite numerous studies on Antarctic cyanobacteria, current knowledge of their diversity and geographic distribution remains fragmented, limiting our ability to assess continent-wide biogeographic patterns and determine whether frameworks such as the Antarctic Conservation Biogeographic Regions (ACBRs) adequately reflect cyanobacterial distributions. Available studies reveal contrasting patterns, ranging from high levels of endemism to the distribution of cold-adapted ecotypes across both polar regions. Moreover, microbial biogeographic studies traditionally rely on conserved marker genes, whose limited genetic resolution may mask fine-scale geographic differentiation. Genome-scale approaches can complement these methods by revealing taxon-specific genes, including genes potentially associated with ecological niches, thereby providing finer-scale information to characterize microbial endemism.

This Ph.D. thesis aimed to improve our understanding of the biogeography of Antarctic cyanobacteria and to identify their genomic specificities. In the **first part**, we analyzed cyanobacterial 16S rRNA amplicon sequences from 84 lakes distributed across eight ACBRs and two sub-Antarctic islands (**Chapter 2**), and from 100 terrestrial samples collected at a finer spatial scale in the Sør Rondane Mountains (**Chapter 3**). In **Chapter 2**, at broad geographic scale, we discovered that **lacustrine** community structure was more strongly associated with geographic distance than with environmental parameters, although their relative contributions became more similar when considering the Antarctic continent alone. Lacustrine cyanobacterial community structures differed significantly between the Antarctic continent and the sub-Antarctic islands, but not among ACBRs. Both regions contained high proportions of taxa potentially endemic to Antarctica, particularly among low-abundance taxa, which also showed more restricted distributions within Antarctica, compared to high-abundance taxa. In **terrestrial**

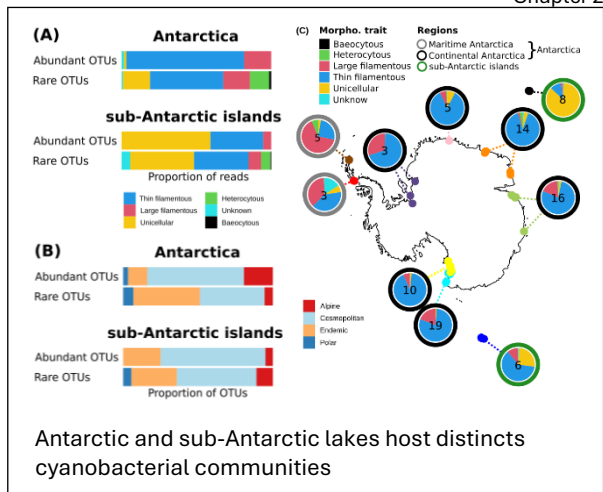
habitats, higher-resolution sampling revealed strong differences among substrates, with granitic soils supporting the highest cyanobacterial diversity, whereas other substrates showed lower diversity and strong single-family dominance (**Chapter 3**). At this local scale, environmental parameters and geographic distance had comparable effects on community structure. A preliminary comparison of both datasets in **Chapter 6** suggest that environmental and spatial factors explain a greater proportion of community variation in terrestrial than in lacustrine habitats, potentially reflecting differences in habitat heterogeneity. In contrast, the proportion of endemic taxa does not appear to differ substantially between the two habitat types, although these preliminary patterns require further investigation.

In the **second part** of this Ph.D. thesis, we used genomic approaches to identify genetic specificities associated with the Antarctic environment and provide complementary information for characterizing endemism. Two case studies focused on broadly distributed taxa: filamentous *Phormidesmis* and unicellular *Synechococcus*-like picocyanobacteria. In **Chapter 4**, an *a priori* approach based on genes of interest (GOIs) revealed higher copy numbers of genes involved in carbohydrate degradation, exopolysaccharide biosynthesis, DNA replication, and stress responses in the Arctic and Antarctic *Phormidesmis*, suggesting potential adaptations to polar conditions, although these results should be interpreted cautiously due to sampling and phylogenetic limitations. In **Chapter 5**, a discovery-based approach identified 16 gene clusters (GCs) specific to Antarctic picocyanobacteria, including two associated with functions enriched in Antarctic genomes. Pseudogenization of one of these GCs in the closest non-Antarctic relatives suggests relaxed selection following a potential ecological niche shift and supports a possible Antarctic origin of their ancestors. Finally, applying the same pangenomic approach to *Phormidesmis* in **Chapter 6** reveals more Antarctic-specific genes than in picocyanobacteria, although technical and biological differences between the two datasets limit robust comparisons.

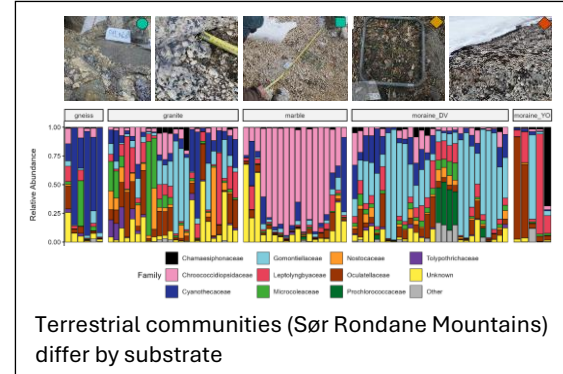
Altogether, the different studies and discussions presented in this thesis examine the concept of endemism from multiple perspectives. As endemism is an important criterion for biodiversity conservation, our findings provide additional scientific evidence to support conservation policies for Antarctic microbial ecosystems.

Graphical abstract

From biogeography...



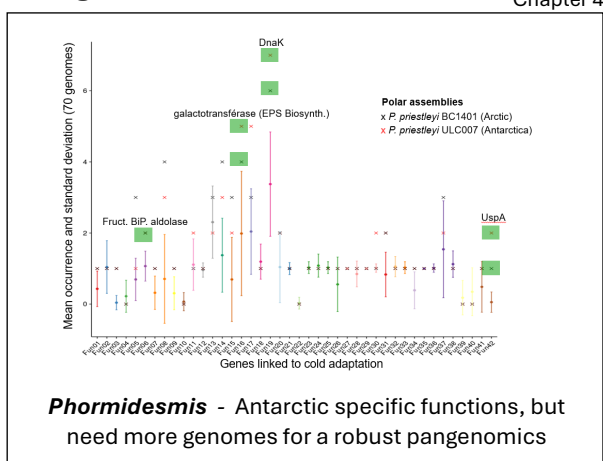
Chapter 3



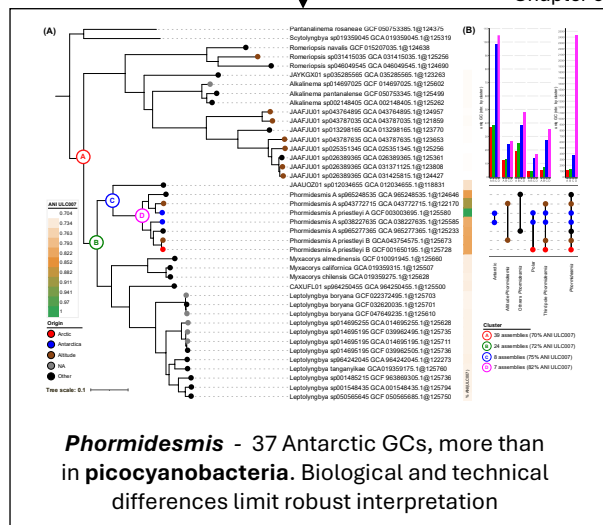
Chapter 6

- Low-abundance taxa show a higher proportion of endemism than high-abundance taxa
- More data are needed for a robust comparison between lacustrine and terrestrial communities

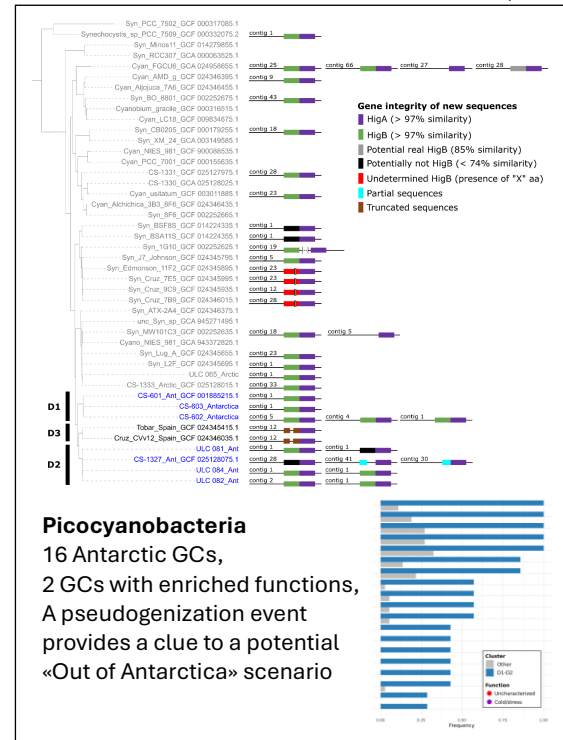
...to genes



Chapter 6



Chapter 5



A thought on the concept of endemism and its implication for ecosystem conservation

Résumé

L'Antarctique a connu une longue histoire d'isolement géographique et climatique depuis sa séparation du Gondwana. La biodiversité antarctique actuelle résulte probablement d'une mosaïque d'organismes issus de lignées anciennes ainsi que de taxons arrivés sur le continent à différentes périodes. Les écosystèmes antarctiques non-marins comprennent divers habitats lacustres, hydro-terrestres et terrestres, qui diffèrent fortement en termes de disponibilité en eau et de conditions environnementales. Les organismes vivant dans ces écosystèmes sont confrontés à de multiples facteurs de stress, souvent en interaction. La combinaison d'un isolement à long terme, d'une forte hétérogénéité environnementale et de conditions extrêmes a profondément façonné la diversité, la distribution et l'évolution du biote antarctique.

En Antarctique, les cyanobactéries sont des colonisateurs primaires favorisant le développement des écosystèmes. Elles peuvent constituer les taxons dominants dans certains écosystèmes microbiens. Malgré les nombreuses études consacrées aux cyanobactéries antarctiques, les connaissances actuelles sur leur diversité et leur distribution géographique restent fragmentées, ce qui limite notre capacité à évaluer les patrons biogéographiques à l'échelle du continent et à déterminer si des cadres tels que les Régions biogéographiques de conservation de l'Antarctique (*Antarctic Conservation Biogeographic Regions*, ACBRs) reflètent adéquatement leur distribution. Les études disponibles révèlent des patrons contrastés, allant de niveaux élevés d'endémisme à une distribution géographique d'écotypes adaptés au froid dans les deux régions polaires. En outre, les études de biogéographie microbienne reposent traditionnellement sur des gènes marqueurs conservés, dont la résolution génétique limitée peut masquer une différenciation géographique à fine échelle. Les approches à l'échelle du génome peuvent compléter ces méthodes en révélant des gènes spécifiques à certains taxons, notamment des gènes potentiellement associés à des niches écologiques, et ainsi fournir des informations à plus haute résolution pour caractériser l'endémisme microbien.

Cette thèse de doctorat vise à améliorer notre compréhension de la biogéographie des cyanobactéries antarctiques et à identifier leurs spécificités génomiques. Dans la **première partie**, nous analysons des séquences d'amplicons du gène de l'ARNr 16S de cyanobactéries provenant de 84 lacs répartis dans huit ACBRs et deux îles subantarctiques (**Chapitre 2**), ainsi que de 100 échantillons terrestres collectés à une échelle spatiale plus fine dans les montagnes Sør Rondane (**Chapitre 3**). Dans le **Chapitre 2**, à large échelle géographique, nous montrons que la structure des communautés **lacustres** est plus fortement associée à la distance géographique qu'aux paramètres environnementaux, bien que leurs contributions relatives deviennent plus

similaires lorsque seul le continent antarctique est considéré. Les structures des communautés cyanobactériennes lacustres diffèrent significativement entre le continent antarctique et les îles subantarctiques, mais pas entre les ACBRs. Les deux régions présentent de fortes proportions de taxons potentiellement endémiques de l'Antarctique, particulièrement parmi les taxons de faible abondance, qui présentent également des distributions plus restreintes au sein de l'Antarctique que les taxons de forte abondance. Dans les habitats **terrestres**, l'échantillonnage à plus haute résolution révèle de fortes différences entre les substrats, les sols granitiques présentant la diversité cyanobactérienne la plus élevée, tandis que d'autres substrats montrent une diversité plus faible et une forte dominance d'une seule famille taxonomique (**Chapitre 3**). À cette échelle locale, les paramètres environnementaux et la distance géographique ont des effets comparables sur la structure des communautés. Une comparaison préliminaire des deux jeux de données dans le **Chapitre 6** suggère que les facteurs environnementaux et spatiaux expliquent une plus grande proportion de la variation des communautés dans les habitats terrestres que dans les habitats lacustres, ce qui pourrait refléter des différences d'hétérogénéité des habitats. En revanche, la proportion de taxons endémiques ne semble pas différer substantiellement entre les deux types d'habitats. Ces résultats préliminaires nécessitent des analyses complémentaires.

Dans la **seconde partie**, nous utilisons des approches génomiques afin d'identifier des spécificités génétiques associées à l'environnement antarctique et de fournir des informations complémentaires pour caractériser l'endémisme. Deux études de cas portent sur des taxons largement distribués : les *Phormidesmis* filamenteuses et les picocyanobactéries unicellulaires de type *Synechococcus*. Dans le **Chapitre 4**, une approche *a priori* fondée sur des gènes d'intérêt (GOIs) révèle un nombre plus élevé de copies de gènes impliqués dans la dégradation des glucides, la biosynthèse des exopolysaccharides, la réplication de l'ADN et les réponses au stress chez les *Phormidesmis* arctiques et antarctiques, suggérant de potentielles adaptations aux conditions polaires. Ces résultats doivent toutefois être interprétés avec prudence en raison de limites liées à l'échantillonnage et au contexte phylogénétique. Dans le **Chapitre 5**, une approche fondée sur la découverte (*discovery-based*) permet d'identifier 16 groupes de gènes (gene clusters, GCs) spécifiques aux picocyanobactéries antarctiques, dont deux associés à des fonctions enrichies dans les génomes antarctiques. La pseudogénisation de l'un de ces GCs chez les plus proches apparentés non-antarctiques suggère un relâchement de la pression de sélection à la suite d'un possible changement de niche écologique, et soutient l'hypothèse d'une origine antarctique de leurs ancêtres. Enfin, l'application de la même approche pangénomique à *Phormidesmis* dans le **Chapitre 6** révèle davantage de gènes spécifiques à l'Antarctique que chez

les picocyanobactéries, bien que les différences techniques et biologiques entre les deux jeux de données limitent la robustesse des comparaisons.

Les différentes études et discussions présentées dans cette thèse abordent le concept d'endémisme sous plusieurs perspectives. L'endémisme constituant un critère important pour la conservation de la biodiversité, nos résultats apportent des éléments scientifiques supplémentaires susceptibles de soutenir les politiques de conservation des écosystèmes microbiens antarctiques.

Chapter 1

General Introduction

1.1. Antarctica: a remote, at-risk continent with extreme environmental conditions

Antarctica was part of the supercontinent Gondwana and separated from South America and Australia during the final stages of its breakup, estimated to have occurred around 40 Ma (Convey et al., 2020). This physical isolation led to the formation of the oceanic Antarctic Circumpolar Current (ACC) 20-25 Ma, which further isolated the Antarctic continent. This isolation accelerated the formation of the continent-wide ice sheets, which probably started to form around 34 Ma (Zachos et al., 2001). Since then, the succession of glacial and interglacial periods has induced fluctuations in the extent of ice-free ground, currently estimated to represent between 0.18% (Burton-Johnson et al., 2016) and 0.44% (Brooks, Jabour, et al., 2019) of the total continental area. Despite this long history that has progressively led to harsh environmental conditions and isolation, life has persisted in Antarctica. There is now substantial evidence that Antarctic biodiversity is not solely the result of recent colonization processes following the Last Glacial Maximum (20 ka ago), but that a part of this biodiversity traces back to ancient biota that lived in Antarctica before its separation from Gondwana (e.g., Allegrucci et al., 2006; De Wever et al., 2009; Pisa et al., 2014; Stevens & D'Haese, 2014). This biota has undergone an evolutionary radiation driven by the continent's geographic isolation and its climatic evolution, generating complex patterns of endemism (Convey et al., 2020).

The present-day Antarctic terrestrial and non-marine aquatic environments include the Antarctic continent and Peninsula, the islands of the Scotia Arc, and the sub-Antarctic islands (Convey et al., 2014) (**Fig. 1.1**). Antarctic ecosystems range from polar deserts, freshwater to hypersaline lakes and their ice cover, to grasslands and eutrophic ponds (Thomas et al., 2008), and habitats can be classified into three categories: lacustrine, hydroterrestrial (wetlands) and terrestrial (Elster, 2002). Lacustrine habitats are characterized by constant liquid water availability, hydroterrestrial habitats undergo freezing periods during the austral winter, and terrestrial habitats either experience only short periods of liquid water availability (e.g., snow melt) or rely on atmospheric humidity (Elster, 2002) (**Fig. 1.2**).

Antarctica is the coldest, driest and windiest place on Earth (Scambos et al., 2018), and its extreme environmental conditions are stressors for organisms inhabiting it (Pearce, 2017). In aquatic and terrestrial environments, the extreme **photoperiod**, from complete darkness in

winter to continuous daylight in summer, strongly controls primary production (Singh & Elster, 2007), as **UV-B radiation** causes transient damage on cell integrity (Turnbull & Robinson, 2009; Robinson & Waterman, 2014). In aquatic environments, **snow** covering the ice layer of water bodies, seasonally or permanently, induces a strong attenuation of photosynthetic active radiation (PAR) (Thomas et al., 2008). On soil, the presence of snow and ice cover buffers temperature variations but reduces light and warming of the ground (Laybourn-Parry et al., 2002; Block et al., 2009). The patchiness of snow and ice presence on soil creates contrasted and disconnected habitats. Also, seasonal thawing leads to soil surface instability and limits long-term colonization in certain terrestrial habitats (Convey et al., 2014).

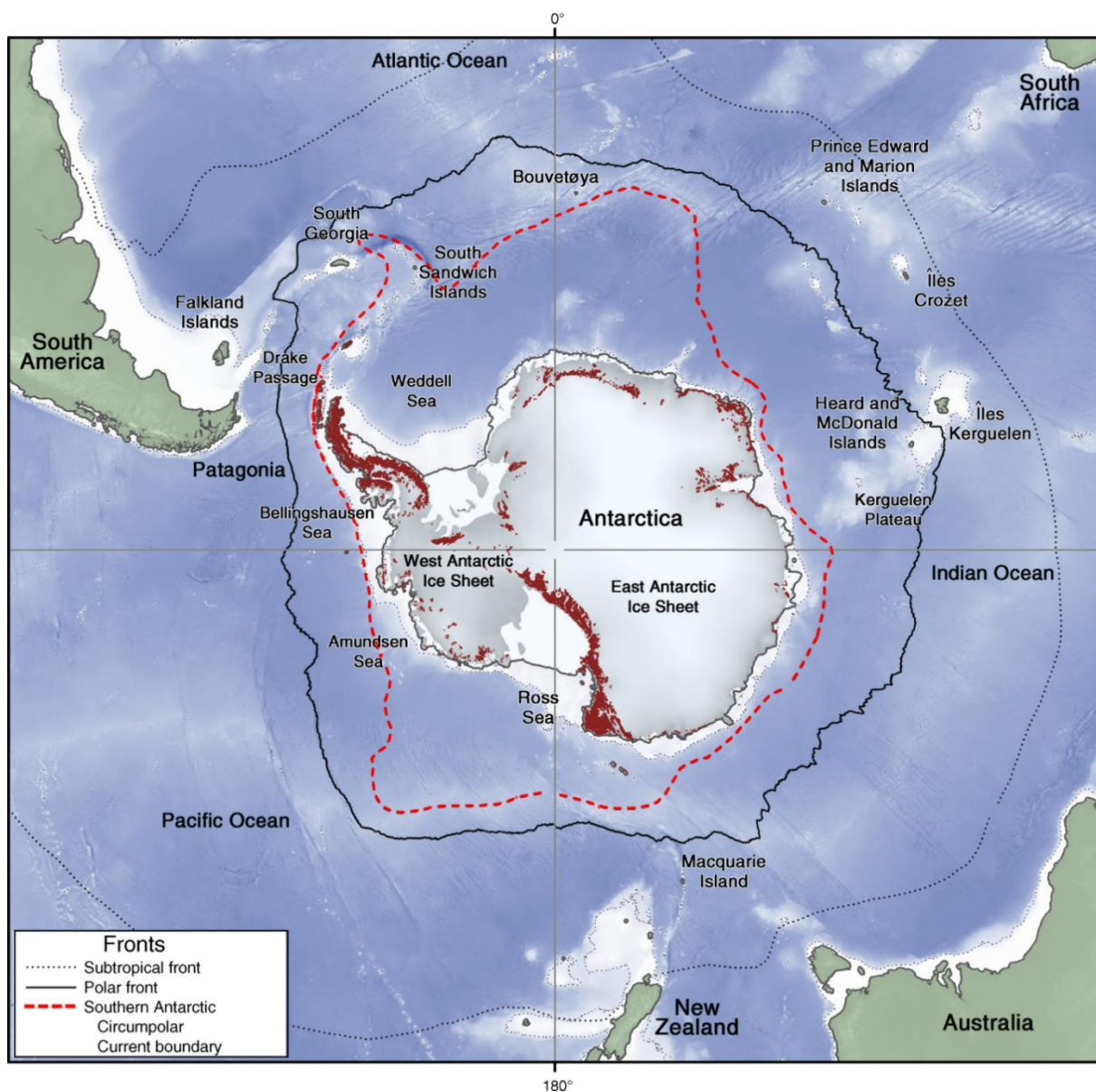


Fig. 1.1. Map of Antarctica, the Southern Ocean, and neighboring regions of the Southern Hemisphere, with the delineation of the Antarctic Circumpolar Current boundary (ACC; red) and ice-free areas (brown) added from [SCAR Antarctic Digital Database](#) using R packages “sf” and “ggplot2” (figure modified from Convey et al., 2014).

Cold temperatures, combined with large annual and daily **temperature variations** in small lakes, ponds, soil and rock surfaces, generate frequent **freeze-thaw cycles**. During freezing, water becomes biologically unavailable as it is converted into ice, resulting in short to long-term periods of **desiccation**. Water freezing, which induces salt exclusion during ice-crystal formation, also results in variations in salt concentrations in lakes or ponds. In meromictic water bodies (i.e., in which layers of water do not intermix), a strong **salinity** gradient can occur with depth (Bell & Laybourn-Parry, 1999). **(Hyper)saline** ponds and lakes as well as brine channels in the sea ice are other stressing habitats (Vincent, 1988). On soil, salt accumulation can also occur due to atmospheric deposit or desiccation (Convey et al., 2014). **Nutrients limitation** (i.e., very low concentrations of nitrogen, phosphorus and (dissolved) organic carbon) occurs both in aquatic and terrestrial habitats (Beyer et al., 200; Henshaw & Laybourn-Parry, 2002). Hence, in aquatic habitats, a temporal nutrient variability exists, with nutrients accumulation in winter and limitation of production in summer (Laybourn-Parry & Bayliss, 1996). In terrestrial habitats, strong **winds** induce both desiccation stress (by increasing evaporation and sublimation) and mechanical abrasion of microbial communities by wind-blown particles (Convey et al., 2014).

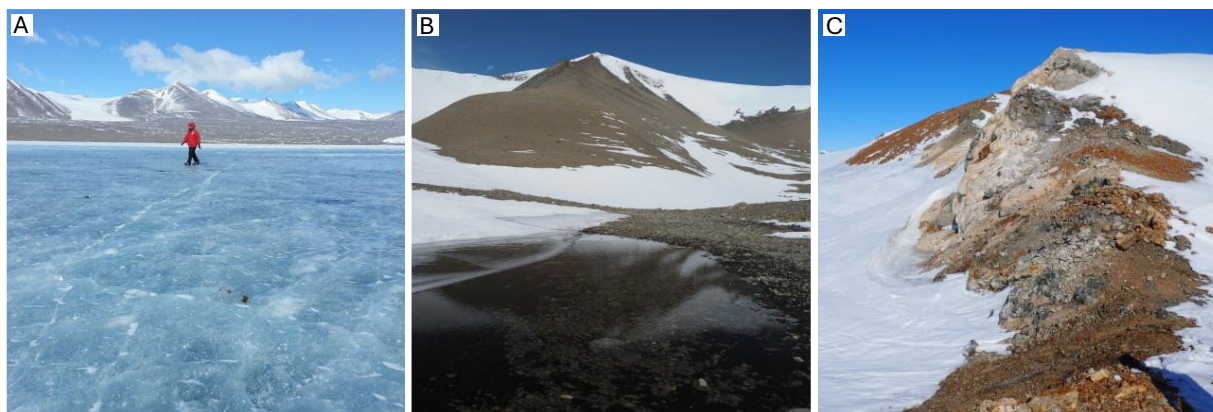


Fig. 1.2. Examples of Antarctic habitats following the classification of Elster et al. (2002). **A.** Lacustrine habitat: Lake Fryxell, a permanently ice-covered lake (Tallada et al., 2022). **B.** Hydroterrestrial (wetland) habitat: a lake in Yûboku Valley (Dronning Maud Land) that freezes completely during the austral winter. © Valentina Savaglia. **C.** Terrestrial habitat: the northeastern nunatak of Perlebandet (Dronning Maud Land). © Valentina Savaglia.

Impact of climate changes on Antarctic ecosystems is well documented (Turner et al., 2014; Lee et al., 2017; Ferreira et al., 2024). Likewise, anthropogenic footprint is recorded for the entire continent: buildings occupy more than 390,000 m², while the additional disturbance footprint just on ice-free land exceeds 5,200,000 m², affecting more than half of all large coastal ice-free areas (Brooks, Jabour, et al., 2019). Despite their exceptional scientific, environmental, and ecological value, ice-free areas are disproportionately affected by human activities. Disturbances were

found to alter the physical environment, reducing habitat suitability as well as directly damaging biota (Brooks, Tejedo, et al., 2019). In this context, nature conservation policies on the terrestrial Antarctic are of particular importance. Although Antarctica is designated as a “natural reserve, devoted to peace and science” under the Protocol on Environmental Protection to the Antarctic Treaty (Environmental Protocol, 1991), the activities concerned by this designation are not clearly defined. Consequently, activities with potentially significant impacts on ecosystems, including fishing, tourism, and station operations, may still be conducted (e.g. Tin et al., 2014; Hughes et al., 2016). Moreover, the Antarctic Specially Protected Areas (ASPAs) designated under the Protocol cover only 1.5% of the ice-free areas, encompassing 44% of its species (Wauchope et al., 2019; Wen et al., 2026). It is therefore essential to identify effective ways to strengthen and implement conservation policies across the continent.

1.2. General presentation of Cyanobacteria

Among the photosynthetic bacteria, Cyanobacteria are the only known phylum capable of oxygenic photosynthesis (Blankenship, 2002). Their popular name “blue-green algae” comes from the bluish color of some cyanobacterial taxa, which contain the pigment phycocyanin in a sufficient concentration (Whitton & Potts, 2012). They are often referred to as among the oldest recorded organisms on the planet, with earliest proposed microfossils dated from ~3500 Ma (Schopf et al., 2007) and oldest undisputed colonial microfossils from 1890-1840 Ma (Hofmann, 1976; Demoulin et al., 2019). Around 2500 Ma ago, they most likely played a key role in the oxidation of the atmosphere (i.e., the Great Oxidation Event), leading to the development of new habitats (Van Kranendonk et al., 2012). It has been proposed that multicellularity allowed cyanobacteria to explore these new ecological niches, subsequently facilitating their diversification (Schirmer et al., 2011) to reach maximum complexity ~2000 Ma ago (Butterfield, 2007).

Current morphological diversity of Cyanobacteria includes both unicellular (e.g., *Synechococcus* sp.) and multicellular forms, ranging from simple filaments with (e.g., *Phormidesmis* sp.) or without sheaths (e.g., *Pseudanabaena* sp.) to complex branching filaments (e.g., *Fischerella* sp.) (**Fig. 1.3A-D**). Some taxa harbor differentiated cells like heterocysts (for nitrogen fixation), akinetes (resting stage cells) or hormogonia (motile filaments) (**Fig. 1.3E-F**). The difference in cell division patterns (e.g., one plane, multiple planes, irregular, asymmetric) and ultrastructural patterns of thylakoids (e.g., parietal, radial, irregular thylakoids) are morphological criteria used

by taxonomists, in combination with phylogenetic analyses, for the separation of groups of taxa (Komárek et al., 2014) (**Fig. 1.3**).

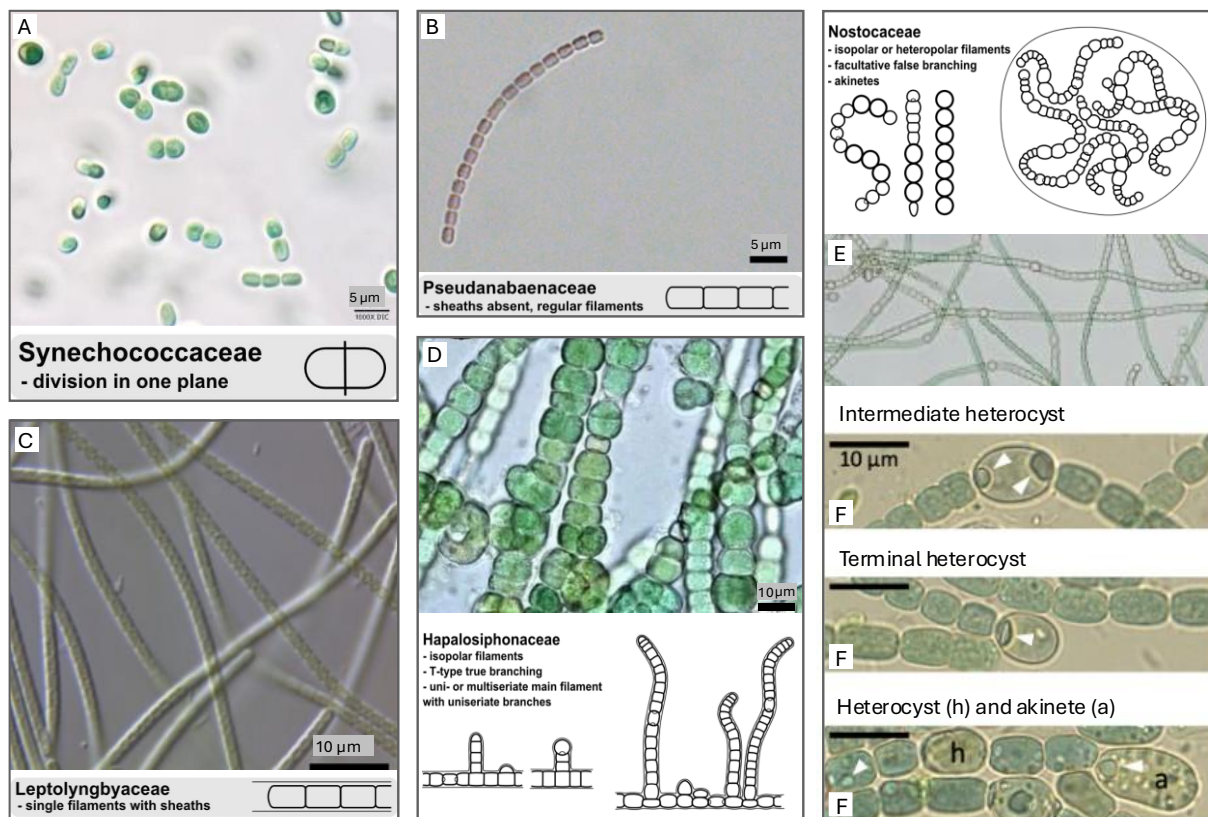


Fig. 1.3. Examples of the morphological diversity of Cyanobacteria, and the morphological characters used for taxonomical identification by Komárek et al. (2014). **A.** *Synechococcus* sp. UTEX 3154 (UTEX Culture collection of Algae), **B.** *Pseudanabaena epilithica* UHCC 1008^T (Christodoulou et al., 2023), **C.** *Leptodesmis lacustris* (Luz et al., 2023), **D.** *Fischerella* sp. NK 1-20 (Keshari et al., 2022), **E.** *Nostoc* sp. UTEX 2209 (UTEX Culture collection of Algae), **F.** *Nostoc azollae* (Sarasa-Buisán et al., 2025).

The cyanobacterial taxonomy is a work in progress (Kaštovský, 2024). Historically, their diversity was first studied using morphological and ecological approaches. But considering the plasticity of their morphology, partly dependent on the environmental or cultivation conditions (e.g., light, temperature), and their high level of cryptic diversity, these initial attempts led to taxonomic polyphyly and species miss-identification (Dvořák et al., 2014; Komárek et al., 2014). Nowadays, more robust polyphasic approaches, which combine molecular with morphological and/or ecological data, are used to delineate taxa (Komárek et al., 2014; Komárek, 2016). Even when morphological observation is difficult or not possible (e.g., complex environmental samples), the use of the 16S rRNA gene as a genetic marker for phylogeny is massively used to better understand the diversity of cyanobacteria. At and below the species level, given the limited taxonomic resolution of the 16S rRNA gene alone (Johnson et al., 2019), some taxa can only be determined using whole-genome data (Komárek, 2016). When morphological data are not available, the Code

of Nomenclature of Prokaryotes Described from Sequence Data (SeqCode; Hedlund et al., 2022), enables the taxonomic description of uncultured prokaryotes based on genome sequence (Sudianto et al., 2026). At higher taxonomic levels, the recent revision of the cyanobacterial taxonomy by Strunecký et al. (2023), based on phylogenomic approaches, resolved polyphyly at order and family levels. This work provided a robust framework to construct 16S rRNA curated databases, like Cyanoseq (Lefler et al., 2023) that greatly improved determination of taxonomic diversity in environmental samples. In **Fig. 1.4**, we redesigned the phylogenomic tree from Strunecký et al. (2023) to have a better overview of cyanobacterial classification.

Cyanobacteria are found in almost all environments receiving light, where they play major ecological roles. In the oceans, they are important contributors to primary production, with the cyanobacterium *Prochlorococcus* and *Synechococcus* producing an estimated 4 and 8 gigatons of fixed carbon per year, respectively, representing approximately 25% of ocean net primary productivity (Flombaum et al., 2013). By converting atmospheric nitrogen (N_2) into bioavailable ammonia (NH_3) or ammonium ion (NH_4^+), nitrogen-fixing cyanobacteria (e.g., *Trichodesmium*, *Crocospaera*, *Nordularia*, *Nostoc*) make a significant contribution to the nitrogen cycle in aquatic and terrestrial environments (Sánchez-Baracaldo et al., 2022). Moreover, some aquatic cyanobacterial taxa can access phosphorus pools in sediments and bottom waters that are generally unavailable to other phytoplankton, thereby contributing substantially to the phosphorus cycle (Cottingham et al., 2015). In polar environments, cyanobacteria are primary colonizers that facilitate ecosystem development and can constitute most of the microbial ecosystem biomass in the most extreme polar habitats (Vincent, 2002).

1.3. Cyanobacteria in Antarctica

1.3.1. Cyanobacterial diversity and adaptation strategies in Antarctica

In Antarctica, the diversity of cyanobacteria has long been studied based on morphological observations. More recently, polyphasic approaches combining molecular and morphological markers, together with the rise of Next-Generation Sequencing (NGS) technologies, have enabled the discovery and improved the characterization of Antarctic cyanobacterial diversity. Based on Velichko et al. (2021), we present the principal Antarctic cyanobacterial genera, and the principal sampling areas on the continent (**Fig. 1.5**). Overall, 38 genera have been identified, representing 16.8% of the total number of unambiguous genera listed by Strunecký et al. (2023) (**Fig. 1.5A**). These genera are distributed across 22 families (40% of the total) and 11 orders (55% of the total), thereby corresponding to a high proportion of the morphological diversity (**Fig. 1.5A, C**). Although

cyanobacteria have been sampled across the whole Antarctica, sampling resolution is not uniform: the McMurdo Dry Valley has been well studied, in contrast to the limited sampling of regions such as Dronning Maud Land or East Antarctica (**Fig. 1.5B**).

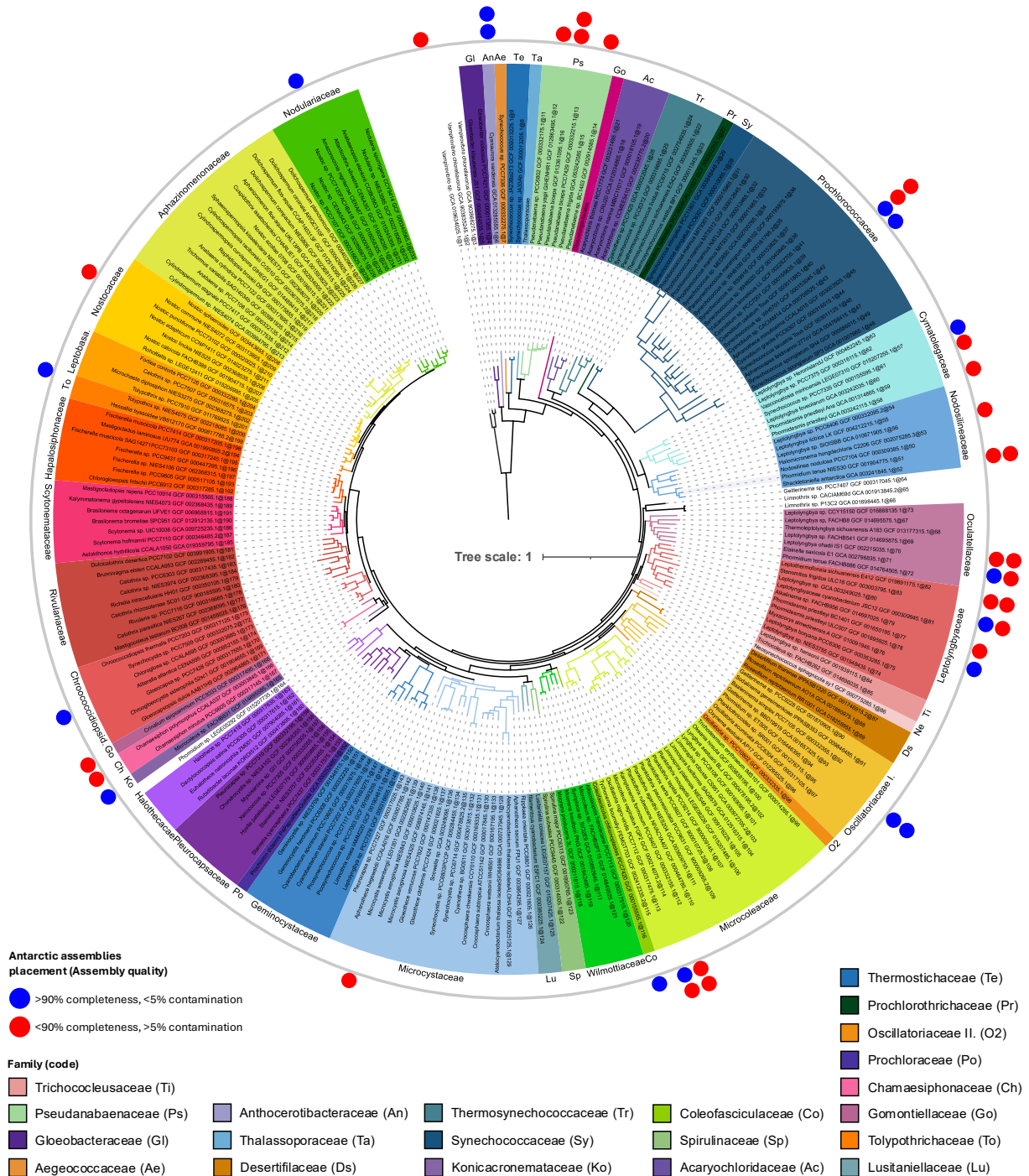


Fig. 1.4. Phylogenomic tree of Cyanobacteria representing taxonomic families and representative assemblies. The tree was constructed by (1) extracting the topology and the branch lengths from the tree of Strunecky et al. (2023) using the program Treemble (Allard & Kumar, 2026) and (2) formatting the Newick file in ITOL (Letunic & Bork, 2024). The blue and red dots indicate the placement of the 42 Antarctic cyanobacterial assemblies and metagenome-assembled genomes (MAGs) identified through a literature search complemented by a *Consensus* search (see **Section 1.4.2.** for details; **Table 1.2-1.3**).

Further sampling efforts in under-sampled areas would provide a more robust representation of Antarctic cyanobacterial diversity.

Along with heterotrophic microeukaryotes (e.g., fungi and non-photosynthetic Stramenopiles), autotrophic microeukaryotes (e.g., Chlorophyta), heterotrophic bacteria (e.g., Alphaproteobacteria) and other autotrophic bacteria (e.g., Chloroflexi), Cyanobacteria form the backbone of the food web in polar ecosystems (Ji et al., 2017; Tytgat et al., 2023). In these harsh environments, they play a central role in supporting the ecosystems through the provision of energy sources and fixed carbon and nitrogen (e.g. Singh & Elster, 2007; Komárek & Komárek, 2010; Christmas et al., 2015; Prieto-Barajas et al., 2018). In Antarctica, cyanobacteria occupy a wide variety of aquatic and terrestrial niches, including streams, rivers, ponds, lakes, sea water, snow, ice and soil, and they are also found in hypolithic and endolithic mineral-associated microbial assemblages (Quesada & Vincent, 1997; Vincent, 2007; Cameron et al., 2012; Pointing et al., 2014). The development of cyanobacteria in harsh environmental conditions is largely due to life strategies, including their ability to form biofilms (Chong et al., 2013; Chong et al., 2015).

Biofilms are thin lamellae (from several μm to a few mm thick) or small globules (up to several mm in diameter) formed by microbial associations in a self-produced exopolysaccharide matrix attached to a solid substrate (Rossi & De Philippis, 2015) (**Fig. 1.5 D1**). Biofilms are found in aquatic environments as well as on soil, ground, or rock surfaces. Most Antarctic biofilms are in the form of microbial mats (Chong et al., 2013; Chong et al., 2015). **Microbial mats** are massive (up to one meter thick), often stratified and sometimes mineralized, biofilms with a consistency ranging from gelatinous to hard and a shape from prostrate lamellae to pinnacles or cones (Stahl et al., 2013; Velichko et al., 2021) (**Fig. 1.5 D2**). Microbial mat ecotypes are ubiquitous and abundant on the Antarctic continent in lakes, small water reservoirs, streams, melt snow canals, or moist grounds around lakes and ponds (Ellis-Evans, 1996; Niederberger et al., 2012; Sumner et al., 2016; Koo et al., 2017).

In addition to cyanobacteria, microbial mats contain microeukaryotes as well as other phototrophic and heterotrophic bacteria. In lakes, the proportion of the different taxa in microbial mats vary with depth (Jungblut et al., 2016). By contrast, **planktonic** cyanobacteria are not abundant in lakes, except in the East Antarctica saline lake plankton, mainly represented by picocyanobacteria (Vincent & Quesada, 2012). Picocyanobacteria are unicells smaller than 2–3 μm in diameter that produce dense populations even in low nutrients environments (Vincent, 2007; Vincent & Quesada, 2012; Callieri et al., 2012) (**Fig. 1.5 D3**). In Antarctic terrestrial ice-free areas, cyanobacteria can form biological soil crusts, covering bare ground essentially free of liquid water.

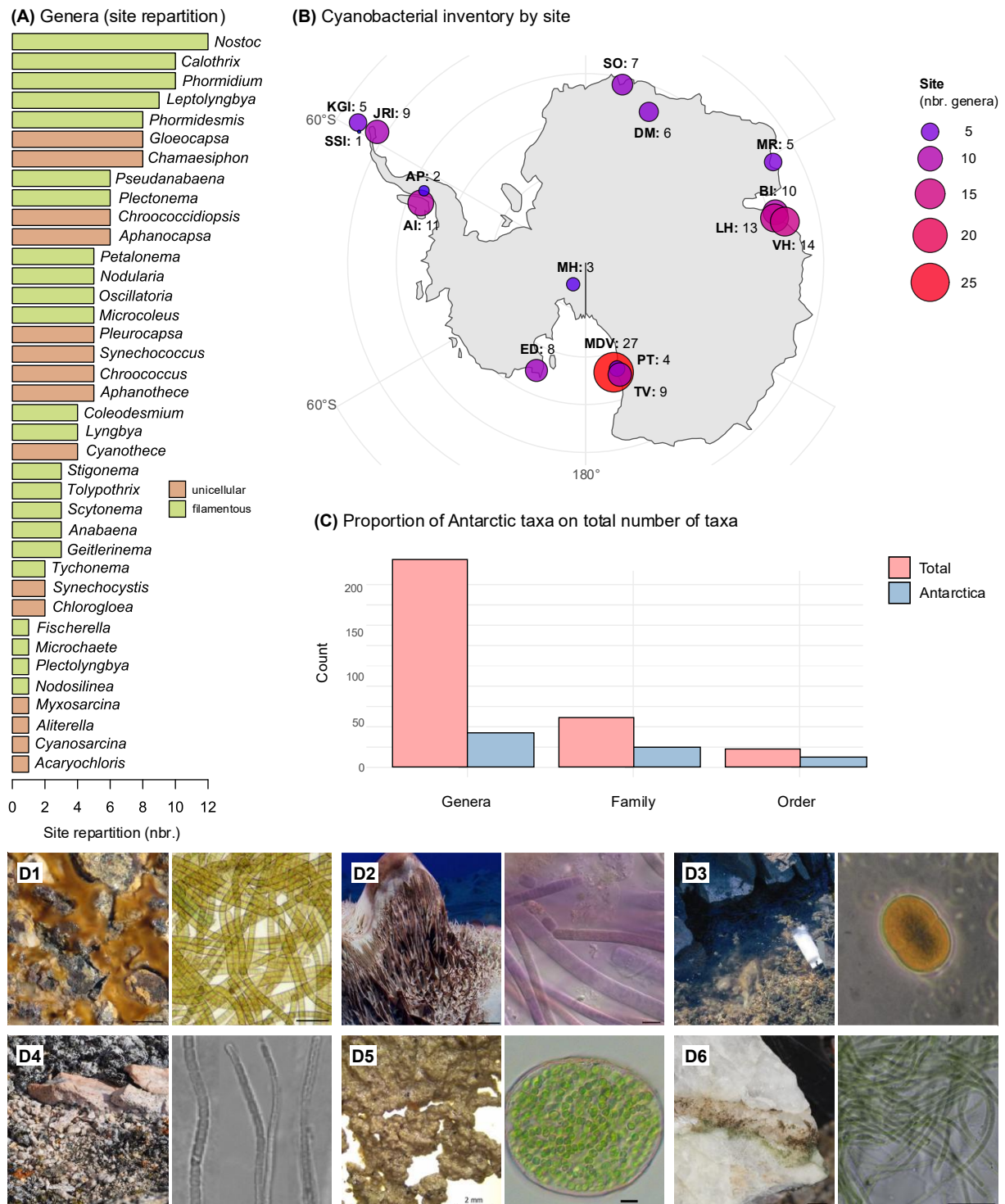


Fig. 1.5. Overview of cyanobacterial records in Antarctica. **A.** Antarctic genera and the number of sampling sites where they were recorded on the continent, as reported in Velichko et al. (2021). **B.** Geographic location and number of cyanobacterial samplings sites, as reported in Velichko et al. (2021). **C.** Number of Antarctic taxa (Velichko et al. 2021) compared with the total number of taxa following Strunecký et al. (2023). **D1.** Biofilm produced by *Oscillatoria* sp., Schirmacher Oasis. **D2.** Conical microbial mats with maroon top layer (© Chicago Tawani Foundation and NASA Exo-biology & Astrobiology Program Expedition-2019, Dale T. Andersen) produced by the cyanobacteria *Oscillatoria* sp. and *Phormidium* sp. (© Svetlana Smirnova) in the deep water of Lake Untersee. **D3.** Planktonic sample in Rock Pool1 (Vestfjella, Dronning Maud Land), containing *Cyanothece* sp. (from Arvola et al., 2024). **D4.** Biological soil crusts in Utsteinen Ridge (© Valentina Savaglia) dominated by *Stenomitos frigidus* (previously reported as '*Leptolyngbya*

frigida) (from Taton et al., 2006). **D5.** Microbial communities in cryoconites holes containing, e.g., *Nostoc* c.f. *microscopicum* (from Dory et al., 2025). **D6.** Endolithic (hypolithic) cyanobacteria, with prevalence of *Phormidesmis priestleyi*. Lake Untersee environs (from Velichko et al., 2021). **D1** and **D6:** © Russian Antarctic Expedition-2018 material.

Biological soil crusts (BSCs) are biofilms containing also other bacteria, green algae, lichens and mosses, where exopolysaccharides-producing filamentous cyanobacteria stabilize the assemblage and provide fixed nitrogen (Quesada & Vincent, 1997) (**Fig. 1.5 D4**). Mineral sediments, biological soil crusts and desiccated microbial mats can be eroded and transported by the wind, ending up on snow or ice. This (often dark) mix of mineral and biological materials causes the melting of the underlying snow or ice because of advection heat resulting from solar radiation, creating a local aquatic biotope called **cryoconite** where the biological material can grow even when it is entombed by an ice lid (Hodgson, 2012) (**Fig. 1.5 D5**). Finally, cyanobacteria can be found in lithic-associated habitats, forming biofilms underneath quartz rocks (**hypolithic**), tunnel into stone bodies (**euendolithic**), in stone fractures (**chasmoendolithic**) or in stone pores or cavities (**cryptoendolithic**) (Golubic & Schneider, 2003; Cockell & Stokes, 2004) (**Fig. 1.5 D6**).

In addition to their life strategies, a set of adaptation mechanisms is necessary to resist environmental stressors, both in terrestrial and aquatic habitats. Only a few psychrophilic Antarctic cyanobacteria, i.e., those with optimal growth at +10°C, have been characterized (Fritsen & Priscu, 1998), including *Microcoleus autumnalis*, which is abundant in cold ecosystems (Christmas et al., 2015). Most Antarctic cyanobacteria are considered psychrotolerant, as they can remain metabolically active near 0°C, although their optimal growth temperature is usually above 15°C (Tang & Vincent, 1999; Singh & Elster, 2007). Cold **temperatures** reduce membrane fluidity and enzymatic activity, induce abnormal protein folding or denaturation and may lead to intracellular ice crystallization (D'Amico et al., 2006). In response, cells can increase membrane fluidity by shortening fatty-acid chains and increasing their degree of unsaturation (Pittera et al., 2018). The production of extracellular polymeric substances (e.g., heteropolysaccharides) can limit the ice nucleation around cells and buffer the impact of **freeze-thaw** cycles (Vincent, 2007). The synthesis of ice-binding proteins (IBPs) and antifreeze proteins reduces the water freezing temperature and allows cyanobacteria to resist phase transitions (Raymond et al., 2007). In addition, the production of compatible solutes (e.g., sucrose, trehalose, glycine betaine) reduces the freezing point of intracellular fluids and limits cell **desiccation** by reducing the amount of water required to maintain osmotic equilibrium when salinity increases (Welsh, 2000). When sudden increase in **salt concentration** occurs, a rapid accumulation of salt helps maintain the osmotic equilibrium. To adapt to the large seasonal variations in **photoperiod**, cyanobacteria

require a set of pigment strategies not only for light harvesting when light is limited in winter but also for light protection in summer against high levels of PAR and **UV radiation** (Roos & Vincent, 1998; Singh & Elster, 2007). In low irradiance conditions, cyanobacteria use highly efficient phycobilin-based light-harvesting complexes that achieve photosynthetic quantum yields approaching the theoretical maximum (Hawes & Schwarz, 2001; Vincent, 2007). By contrast, high energy PAR and UV radiation can induce photo-inhibition, phycobiliprotein degradation, chlorophyll-bleaching, DNA damage, as well as the most deteriorative consequences: the production of reactive oxygen species (ROS), particularly singlet oxygen and superoxide anion-radical (Vincent, 2007; Velichko et al., 2021). ROS protective enzymes (e.g., superoxide dismutase) neutralize ROS molecules while carotenoids perform excited state quenching via heat dissipation (Vincent & Quesada, 2012). To protect their DNA from UV radiation, cyanobacteria produce photoprotective screening pigments and other quenching pigments (e.g., gloeocapsin, scytonemin, mycosporine-like), reduce the production of potent endogenous photosensitizers (Chl *a* and phycobiliproteins) (Proteau et al., 1993) and intensify their DNA repair processes, such as excision repair and photo-reactivation (Vincent, 2007; Zakhia et al., 2008; Castenholz & Garcia-Pichel, 2012). These processes are slower at lower temperatures, inducing an exacerbation of net damage due to high irradiance (Vincent, 2007). Differences in pigment contents were observed in microbial mat layers, with carotenoid accumulation in the surface water layer and high Chl *a* in the deeper horizons (Quesada & Vincent, 1997; Vincent, 2007).

1.3.2. Biogeography of Antarctic cyanobacteria

Biogeography is the study attempting to describe and explain spatial patterns of biodiversity and how these patterns change over time (Lomolino et al., 2006). Historically, biogeographic studies focused mainly on plants and animals. However, numerous studies now focus on microorganisms, including bacteria, archaea, fungi and microbial eukaryotes, highlighting the existence of microbial biogeographic patterns (e.g. Nemergut et al., 2011; Zheng et al., 2013; Tedersoo et al., 2014; Chow & Suttle, 2015; Ma et al., 2017; Verleyen et al., 2021). Microbial community composition across landscapes is not random. Microbial biogeographic patterns include, for example, evidence that different locations harbour distinct microbial communities (e.g., O'Mullan & Ward, 2005; Kociolek et al., 2017), distance-decay relationships, defined as a decline in community structure or taxonomic similarity with increasing geographic distance (e.g., Green et al., 2004; Milici et al., 2016), or even the existence of endemic taxa, defined as a taxon restricted to a particular location, region or habitat type (e.g., Whitaker et al., 2003; Oakley et al., 2010; Pessi et al., 2018). Four processes have been proposed to shape microbial biogeographic

patterns through a combination of evolutionary and ecological factors: selection, drift, dispersal, and mutation (Hanson et al., 2012). **Selection** acts at multiple biological levels, including genes, cells, individuals, and taxa, favouring entities with the highest fitness. **Drift** is the change in genotype or species frequencies resulting from random demographic fluctuations. **Dispersal** is the movement and establishment of a taxa, which are defined by their genotypes. **Mutation** generates non-directed changes in nucleotide sequences and, together with the other three processes, contributes to diversification among taxa. A key challenge is to identify the relative importance of these four processes in shaping the observed biogeographic patterns (Hanson et al., 2012), which can also be considered in terms of the relative contributions of contemporary selection and historical processes.

Among the biogeographic patterns observed in Antarctica, a major characteristic is the high level of endemic taxa, including plants, animals, and microorganisms (Allegrucci et al., 2006; Chong et al., 2015; Biersma et al., 2018; Pessi et al., 2018; Kociolek et al., 2017; Convey et al., 2020). Signals of regional endemism have led to the delineation of Antarctic Conservation Biogeographic Regions (ACBRs), which cover most of the ice-free areas of the Antarctic continent and Antarctic Peninsula (Terauds & Lee, 2016) (**Fig. 1.6A**). ACBRs are extensively used in a range of studies, including assessment of area protection levels (Chown et al., 2015; Shaw et al., 2014) and spatial analyses of biodiversity (e.g., Convey et al., 2014; Fraser et al., 2014; Pessi et al., 2018). However, for microorganisms, results are not consistent across taxonomic groups. For instance, while lake diatoms at the morphospecies level appear to cluster according to ACBRs (Verleyen et al., 2021), this pattern is not fully observed for soil bacteria (Varliero et al., 2024) (**Fig. 1.6B**). Moreover, the proportion of endemism may vary among taxonomic groups, including at phylum level, whether in bacteria or eukaryotes (Tytgat et al., 2023).

Regarding Antarctic cyanobacteria, several studies based on morphological and molecular traits have reported biogeographic patterns. In line with observations of other bacteria, a latitudinal gradient has been identified in terrestrial Antarctic cyanobacterial communities based on a meta-analysis of OTUs (Namsaraev et al., 2010). High similarity among cyanobacteria from the cold biosphere, including polar, high-altitude and cold deserts, has also been reported (Christmas et al., 2015; Jungblut et al., 2010; Kleinteich et al., 2017). This is further supported by numerous studies showing that Antarctic cyanobacterial communities consist of mixtures of cosmopolitan, cold-related, and endemic phylotypes (Taton et al., 2006; Jungblut et al., 2010; Namsaraev et al., 2010; Pessi et al., 2018). The reported proportion of endemism varies among studies, which may be explained by differences in sampling, including geographic resolution and spatial scale, as well as by the methods, including the taxonomic resolution of the markers used (Hanson et al., 2012;

Chase & Martiny, 2018). To go further, comparing biogeographic patterns inferred from conserved marker genes with those obtained from more variable markers, such as the 16S-23S ITS region, or (meta)genomic data, would allow to more robustly detect potential endemic signals (Chase & Martiny, 2018; J. Lumian, Sumner, et al., 2024).

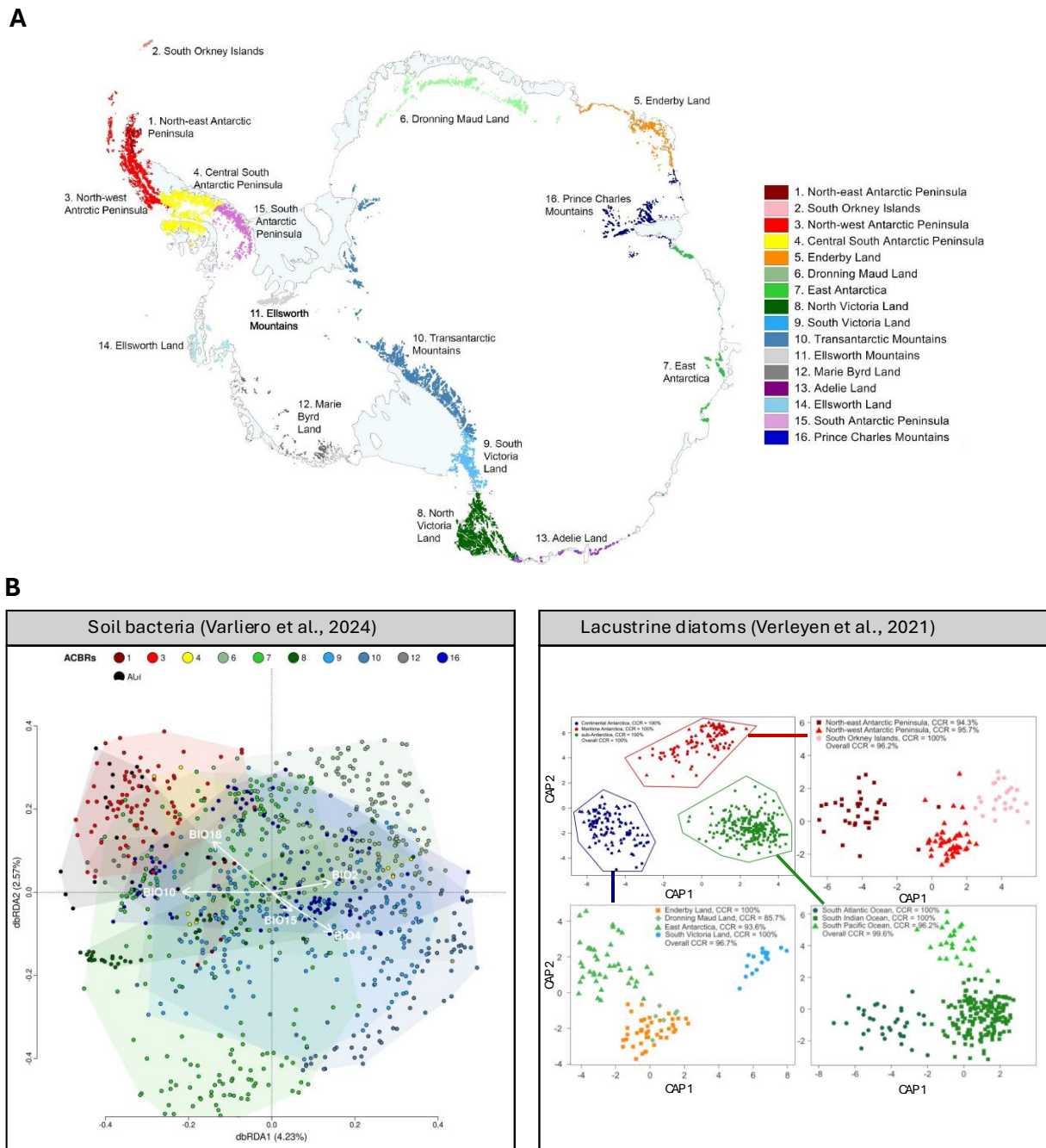


Fig. 1.6. A. Updated version of the Antarctic Conservation Biogeographic Regions (ACBRs v2), from Terauds & Lee (2016). **B.** Comparison of biogeographic patterns in soil bacteria and lacustrine diatoms reported from the literature: soil bacterial communities (Varliero et al., 2024) show no clustering according to ACBRs, whereas lacustrine diatoms communities (Verleyen et al., 2021) cluster according to ACBRs.

1.4. Genomics of Antarctic cyanobacteria: tools to study biogeography and evolution

In the research chapters of this thesis, two next-generation sequencing (NGS) approaches will be used. First, amplicon sequencing of 16S rRNA gene sequences will be used to characterize biogeographic patterns of cyanobacterial diversity and link them to environmental parameters (**Chapters 2 and 3**). Secondly, whole-genome sequencing (WGS) will be performed to obtain complete genome assemblies and identify Antarctic genomic specificities (**Chapters 4 and 5**).

1.4.1. Amplicon sequencing of cyanobacterial communities

Environmental amplicon sequencing, sometimes called targeted metagenomics, metagenetics or metabarcoding, involves extracting DNA from an environmental sample, amplifying fragments of a selected molecular marker gene by polymerase chain reaction (PCR), and subsequently sequencing the amplicons in parallel using a high-throughput sequencing (HTS) platform, such as 454 pyrosequencing (formerly Life Sciences; Margulies et al., 2005) or Illumina sequencing (formerly Solexa Inc.; Bentley et al., 2008). It produces sequence data at a scale of several orders of magnitude higher than Sanger sequencing (Sanger et al., 1977). This exceptional sequencing depth, coupled with increasingly affordable costs, has made it one of the preferred methods for obtaining molecular data in microbial ecology over the last two decades. For prokaryotes, the 16S rRNA gene is one of the most widely used markers in microbial evolution and ecology studies (Tringe & Hugenholtz, 2008). This is because it is found in all prokaryotes, is subject to similar evolutionary pressures across phyla, contains both conserved and variable regions, and show little evidence of horizontal gene transfer, making it a useful phylogenetic marker at different taxonomic scales (Woese, 1987). However, it has been shown to lack resolution below the genus level (Johnson et al., 2019). At finer taxonomic scales, other molecular markers can be used, such as the less conserved internal transcribed spacer (ITS) region, or even protein-coding genes to investigate specific taxa (Pessi, 2017). Although full-length amplicon sequencing has recently become more widely accessible (e.g., on PacBio Revio platform), most current amplicon sequencing studies focus on hypervariable regions or intergenic regions because of platform-specific read-length limitations (e.g., 2 x 300 bp on the Illumina MiSeq). For cyanobacteria, the V3-V4 variable region of the 16S rRNA gene (~422 bp), between two cyanobacteria-specific primers, is recommended (Nübel et al., 1997).

Bioinformatics processing of amplicon sequencing data typically consists of quality filtering, denoising, and the removal of chimeric sequences, followed by the clustering of amplicons in

operational taxonomic units (OTUs) using tools such as UPARSE (Edgar, 2013) or amplicon sequence variants (ASVs) using tools such as DADA2 (Callahan et al., 2016). Taxonomic assignment of amplicon reads can then be performed either through phylogenetic inference, using tools such as pplacer (Matsen et al., 2010), or through percentage-similarity matching against public databases such as NCBI GenBank (Sayers et al., 2025) or Cyanoseq (Lefler et al., 2023), using for example BLASTn (Altschul et al., 1990). Finally, amplicon sequencing pipelines provide a semi-quantitative overview of taxonomic diversity in each sequenced sample. This enables subsequent statistical analysis of alpha diversity, i.e., within-sample diversity, and beta diversity, i.e., between-sample diversity, to characterize and compare cyanobacterial communities. Several studies referenced above and focusing on cyanobacterial biogeography (**Section 1.3.2**) are based on 16S rRNA gene amplicon sequencing (e.g., Christmas et al., 2015; Kleinteich et al., 2017; Pessi et al., 2018). Moreover, amplicon sequencing is a powerful tool to study how microbial communities have changed over time (Jungblut & Hawes, 2017). Amplicon sequencing results can also be used to assess the conservation potential of specific areas (C. A. Guerra et al., 2022), or for long-term environmental monitoring to evaluate conservation management plans (Harmel et al., 2025).

To understand the functioning of microbial communities or to determine the genetic specificities of certain taxa, it is not sufficient to sequence marker genes, and other NGS techniques should be used. For example, metagenomics makes it possible to recover the total DNA from a given sample, reconstruct partial or complete genomes (the so-called Metagenome-Assembled Genomes or MAGs), and subsequently highlight their metabolic potential by analyzing MAG gene content. Hence, Antarctic cyanobacteria have been shown to contain a SigB-type sigma factor, which acts as a regulator of different stress responses (Varin et al., 2012). Metagenomic analyses have also revealed that, under cold conditions, the molecular chaperones DnaJ and DnaK are induced and help to maintain the native conformation of regulatory proteins such as DnaA, GyrA and RecA, which are involved in DNA topology (Varin et al., 2012). However, to determine the complete genomic potential of individual taxa and validate the presence or absence of specific genes, it is necessary to reconstruct MAGS as complete and as uncontaminated as possible (Cornet & Baurain, 2022).

1.4.2. Genomics of Antarctic cyanobacteria

Strunecký et al. (2023) provided a structured overview of the diversity of cyanobacterial genomes accessible for genomic research, allowing newly obtained culture-isolated assemblies and MAGs

to be placed within a robust phylogenomic context (**Fig. 1.4**). The availability of complete or near-complete genomic assemblies makes it possible not only to determine taxonomy, but also to characterize (1) the complete or near-complete metabolic potential of the corresponding organisms and (2), the genomic features that may provide insights into their evolutionary history (e.g., horizontal gene transfers). By placing analyses within an evolutionary context, genomics also allows highlighting genomic specificities linked, for example, to spatial and temporal parameters. In the context of understanding biogeographic patterns in Antarctica, such a genomic information complements estimates of Antarctic endemism based on marker genes, as it has the potential to identify genomic features that may reflect adaptations to Antarctic environments (Dsouza et al., 2015; Dragone et al., 2026).

How many Antarctic genomes are publicly available, and how are they being studied? Based on a literature survey, complemented by a *Consensus* search followed by manual refinement, we identified 15 papers that explicitly investigate cyanobacterial genomics and collectively report 42 Antarctic genomic assemblies available online in NCBI GenBank (**Table 1.2-1.3**), of which 86% are MAGs. Overall, 16 assemblies (38%) fall below the quality threshold of >90% completeness and <5% contamination (Bowers et al., 2017). The main sequencing technology used is Illumina short-read sequencing (13 papers, 86%), either alone or in combination with long-read sequencing technologies (3 papers, 20%). To visualize the genomic diversity represented in this dataset, we mapped the assemblies onto the phylogenomic tree of Strunecký et al. (2023) (**Fig. 1.4**).

Ten (66%) of these 15 papers focus on the comparative genomics of newly reported genomes, aiming to highlight genetic adaptations associated with survival under harsh Antarctic environmental conditions. Based on functional annotations, twelve studies (80%) list a set of targeted “genes of interest” (GOIs) and their functional categories, discussing their presence or absence to hypothesize potential adaptations to, for example, cold stress, desiccation and UV radiation. The recorded functional categories and examples of the genes studied are summarized in **Table 1.4**. Photosynthesis is the most studied functional category, with its different components analyzed in detail (cytochrome b6f, photosystem I and II, light-harvesting systems, the thylakoid membranes, the Calvin cycle, and CO₂/HO₃⁻ uptake systems). This category is followed by photoprotective compounds, mostly pigments (e.g., scytonemin, mycosporine-like amino acids (MAAs), carotenoids). Cold adaptation is also well studied, whereas other functional categories (e.g., EPS biosynthesis, circadian clock or desiccation) are less extensively studied and depend on the focus of each paper. Most often, these studies used generalist databases (e.g., COGs, KEGG) to infer annotations based on protein sequence similarity.

Table 1.2. (Part 1) Description of the 15 papers identified from the literature and the *Consensus* search.

Study	Format	Assembly name, NCBI accession, completeness (%), contamination (%)	Phyto.	Biogeo.	GOIs	Discovery	Sequencing
Lara et al., 2017	genome assembly	<i>Phormidesmis priestleyi</i> ULC007, GCF_001895925.1, 96.55, 0.16	no	no	yes	no	Illumina
Christmas et al., 2018	genome assembly	<i>Leptolyngbya</i> sp. BC1307, GCF_002286735.1, 92.65, 2.51	yes	no	yes	no	Illumina
Cornet et al., 2018	MAGs	<i>Cyanobium</i> sp. ULC082bin, GCA_003248995.1, 81.78, 3.93	yes	no	no	no	Illumina
		<i>Cyanobium</i> sp. ULC084bin3, GCA_003249115.1, 84.41, 5.14					
		<i>Leptolyngbya</i> sp. ULC077bin1, GCA_003249025.1, inconclusive					
		<i>Leptolyngbya foveolarum</i> ULC129bin1, GCA_003242035.1, inconclusive					
		<i>Phormidesmis priestleyi</i> ULC027bin1, GCA_003242115.1, inconclusive					
		<i>Shackletoniella antarctica</i> ULC041bin1, GCA_003241845.1, inconclusive					
Tang et al., 2019	genome assembly	<i>Leptolyngbya</i> sp. ULC073bin1, GCA_003249155.1, inconclusive	no	no	yes	no	PacBio, Illumina
		<i>Synechococcus</i> sp. SynAce01, GCF_001885215.1, 98.82, 0					
Grettenberger et al., 2020	MAGs	<i>Candidatus Cyanaurora vandensis</i> , GCF_013285555.1, 91.65, 1.9	yes	no	yes	no	Illumina
		<i>Candidatus Cyanaurora vandensis</i> , GCF_013285545.1, 92.06, 1.63					
Lumian et al., 2021	MAG	<i>Phormidium pseudopriestleyi</i> FRX01, GCF_017313335.1, 92.64, 1.57	no	no	yes	no	Illumina
Effendi et al., 2022	genome assembly	<i>Nostoc</i> SO-36, GCF_023734775.1, 84.75, 2.05	yes	no	yes	no	PacBio, Illumina
Cabello-Yeves et al., 2022	MAG	<i>Synechococcus</i> sp. Edmonson 11F2, GCF_024345895.1, 99.29, 0.12	no	yes	yes	yes	Illumina

Table 1.3. (Part 2) Description of the 15 papers identified from the literature and the *Consensus* search.

Study	Format	Assembly name, NCBI accession, completeness (%), contamination (%)	Phylo.	Biogeo.	GOIs	Discovery	Sequencing
Pessi et al., 2023	MAGs	Synechococcales cyanobacterium PMM_0082, GCA_949127865.1, 85.92, 0.00	no	yes	yes	no	Illumina
		Synechococcales cyanobacterium PMM_0072, GCA_949127675.1, 83.10, 1.41					
		Synechococcales cyanobacterium PMM_0058, GCA_949127795.1, 84.51, 1.41					
		Synechococcales cyanobacterium PMM_0002, GCA_949127925.1, 90.14, 1.41					
		Synechococcales cyanobacterium PMM_0083, GCA_949127665.1, 94.37, 1.41					
		Oscillatoriales cyanobacterium PMM_0100, GCA_949128045.1, 80.28, 2.82					
		Chroococcales cyanobacterium PMM_0086, GCA_949127655.1, 80.28, 2.82					
		Nostocales cyanobacterium PMM_0081, GCA_949127805.1, 91.55, 2.82					
		Oscillatoriales cyanobacterium PMM_0080, GCA_949127775.1, 94.37, 2.82					
		Oscillatoriales cyanobacterium PMM_0001, GCA_949127755.1, 94.37, 2.82					
Skoupý et al., 2022	genome assembly	Synechococcales cyanobacterium PMM_0084, GCA_949127885.1, 81.69, 4.23	yes	yes	no	no	Illumina
		Synechococcales cyanobacterium PMM_0085, GCA_949127705.1, 81.69, 4.23					
Dextro et al., 2023	genome assembly	Synechococcales cyanobacterium PMM_0099, GCA_949127875.1, 67.61, 5.63	yes	no	yes	no	PacBio
		<i>Candidatus Obscuribacteriales</i> bacterium PMM_0089, GCA_949127735.1, 77.46, 5.63					
Lumian et al., 2024a	MAGs	Synechococcales cyanobacterium PMM_0003, GCA_949128015.1, 78.87, 5.63	no	no	yes	no	Illumina
		Synechococcales cyanobacterium PMM_0073, GCA_949127695.1, 85.92, 5.63					
Lumian et al., 2024b	MAGs	Oscillatoriales cyanobacterium PMM_0059, GCA_949127745.1, 90.14, 5.63	no	yes	no	no	No sequencing
		Oscillatoriales cyanobacterium PMM_0004, GCA_949127945.1, 60.56, 7.04					
Roncero-Ramos et al., 2024	MAG	<i>Argonema antarcticum</i> A004/B2, GCF_0233333585.1, 99.52, 0.44 <i>Argonema gateatum</i> A003/A1, GCF_0233333595.1, 99.52, 0.64	no	no	yes	no	Illumina, Nanopore
Van Goethem et al., 2025	MAG	<i>Halotia brancorii</i> CENA392, GCF_029953635.1, 99.21, 1.15	yes	no	yes	yes	Illumina

Table 1.4. Synthesis of the literature survey on Antarctic cyanobacterial genomics: functional categories and examples of analyzed genes.

Functional category	Examples of genes involved	Papers
Photosynthesis	Cytochrome b6f (petA-D, petG, petM) ; Photosystem I (psaA, psaB, psI-K) ; Photosystem II (psbA, psbD, psbH-M, psbX-Y) ; Thylakoid membrane (sqdB, sqdX) ; Light harvesting complex (ApcA-F, CpcA-G, Peca-C) ; Calvin cycle (cbbL, cbbS, prk, RbcS, RbcL) ; CO2/HCO3-uptake systems (ndhB, ndhM, ndhAICE, bicA)	Chrismas et al. 2018, Dextro et al. 2023, Grettenberger et al. 2020, Lumian et al. 2021, Lumian et al. 2024a, Pessi et al. 2023, Tang et al. 2019, Van Goethem et al. 2025
Photoprotection	Scytonemin (trpA-G, tyrA, aroB, scyA-B, scyD-E) ; Mycosporine-like amino-acid (MAA) (mysA-B, mysE, mysC, NRPS-like) ; Carotenoid (crtB, crtE, crtH, crtO, crtP, cruA) ; Orange carotenoid protein (OCP)	Chrismas et al. 2018, Dextro et al. 2023, Grettenberger et al. 2020, Lumian et al. 2021, Pessi et al. 2023, Roncero-Ramos et al. 2024, Van Goethem et al. 2025
Cold adaptation	Cold shock protein (csp) ; Lipid desaturation (desA, desC, fabG) ; RNA helicases (dead, hrpB, helY) ; DNA repair system (holB, recA, recJ, uvrA) ; gyrA ; tig	Dextro et al. 2023, Effendi et al. 2022, Lara et al. 2017, Lumian et al. 2024a, Pessi et al. 2023, Van Goethem et al. 2025
EPS biosynthesis	wzx, wzy, wza, wzb, wzz, kpsM, kpsT, kpsU	Effendi et al. 2022, Pessi et al. 2023, Roncero-Ramos et al. 2024
Circadian clock	KaiA, KaiB, KaiC, LdpA, Pex, SasA	Lara et al. 2017, Lumian et al. 2024a, Pessi et al. 2023
Desiccation, osmotic stress and compatible solutes	Ion homeostasis (nhaS2, nhaS3, mrpA, mrpC, ktrA, ktrB, ktrE, aqpZ) ; osmoprotectant biosynthesis (sps, spp, treY, srlD, betaA, betB)	Roncero-Ramos et al. 2024, Tang et al. 2019, Van Goethem et al. 2025
Oxidative and other stress response	gshB, gpx, gor, dnaK, dnaJ, msrA	Lara et al. 2017, Roncero-Ramos et al. 2024, Van Goethem et al. 2025
Nitrogen and iron metabolism	Nitrogenases (nifH, nifD, nifK) ; Nitrate/nitrite reductases (narB, nirD, nasD) ; iron uptake (yusV, efeU, iscA, fbpA, feeF)	Pessi et al. 2023, Van Goethem et al. 2025
Secondary metabolites	NRPS, NRPS/PKS and PKS-like gene clusters	Dextro et al. 2023, Lara et al. 2017
Mobile Genetic Elements (MGEs)	Type I-C array CRISPR-Cas systems	Van Goethem et al. 2025
Sulfide tolerance	sqr, sat, cysNC, cysH, sir	Lumian et al. 2021

Outside the Antarctic context, specific curated databases also exist, such as Cyanorak for marine and brackish picocyanobacterial genomes (Garczarek et al., 2021), which allow functional annotations to be refined.

Finally, four (26%) of the 15 papers focus on biogeographic topics. One way to conduct a robust biogeographic analysis, especially when assessing genome “uniqueness” as a potential indicator of endemism, is to ensure that **(1)** the analysis is valid in terms of evolutionary context (Dewar et al., 2025) and **(2)** it identifies genes or gene clusters (GC) specifically associated with the geographic area studied, in a discovery-based approach (Delmont & Eren, 2018), rather than relying on a predefined list of genes of interest (GOIs).

Concerning the first point, 53% of these papers placed the genome(s) studied within a robust phylogenetic framework, in which the closest related taxa were compared, rather than taxa selected in a semi-random way. Placing the genomes of interest in a phylogenetic context provides a less biased comparative framework and allows hypotheses to be inferred about the evolutionary history of the genomes (Dewar et al., 2025).

The second point proposes an approach that makes it possible to identify unexpected or poorly characterized genomic features associated with geographic or environmental differentiation, as it is not constrained by the prior selection of candidate genes (Van Goethem et al., 2025; Dragone et al., 2026). Only 13% of the 15 papers listed have performed such a discovery-based approach (**Table 1.2-3**). **Pangenomic analysis** is a discovery-based approach that aims to represent all genomic variations found in a group of organisms (Matthews et al., 2024). For a given set of genomes, the “pangenome” represents the complete genetic repertoire of that genome set. It consists of two main components: (1) the ‘core’ genome, which contains gene clusters shared by all selected genomes, and (2) the ‘accessory’ genome, which comprises all non-core genes (Kaur et al., 2024; Matthews et al., 2024; Tettelin & Medini, 2020). The accessory genome can then be subdivided into strain-specific genes found in only one genome, also called singletons, and genes shared by a subset of genomes. If this subset consists of genomes sharing the same ecological niche (e.g., Antarctica environment), regardless of their evolutionary relationships, the genes they share can be defined as the ‘ecogenome’ (Cordero & Polz, 2014; Smillie et al., 2011). This distinction helps to avoid confusing taxonomic history with ecological adaptations when analyzing complex communities (Domingo-Sananes & McInerney, 2021). However, the pangenome is not a fixed biological entity, rather a statistical model whose outputs depend on dataset quality, curation, and taxonomic resolution (Guerra, 2026).

1.5. From biogeography to genes: main objectives of the thesis and research questions

1.5.1. Main objectives

Historically, this thesis brings together two research projects. The first part of the thesis is part of the project *Climate Change and Antarctic Microbial Biodiversity* (CCAMBIO), which aimed to investigate the diversity, biogeographic zoning, evolutionary history, and genomic composition of lacustrine microbial mat communities across the Antarctic Realm in order to assess their resilience and their local and regional responses to global change. The second part stems from the *ANTAREVOL* project, funded through a FRIA fellowship from the F.R.S.-FNRS. Its objective was to use a robust phylogenomic framework together with representative cyanobacterial strains to identify biological characteristics associated with Antarctic environmental conditions and to infer hypotheses regarding genome evolution.

By integrating these two projects, this thesis aims to improve our understanding of the diversity and evolution of Antarctic cyanobacteria and to further investigating microbial endemism through complementary methodological approaches at spanning multiple scales, from biogeographic analyses to pangenomics. The four research chapters are organized into two main parts:

- 1) The first part contains two research chapters that use amplicon sequencing to better understand the biogeographic patterns of (sub-)Antarctic cyanobacteria. The first research chapter analyses cyanobacterial communities from a set of lakes across eight Antarctic Conservation Biogeographic Regions (ACBRs) and two sub-Antarctic islands, using 16S rRNA sequences to assess the relevance of ACBRs for cyanobacteria and to determine their rate of endemism (**Chapter 2**). In the second research chapter, terrestrial biological soil crusts from the Sør Rondane Mountains (Dronning Maud Land) are analyzed using 16S rRNA sequences to describe the cyanobacterial diversity encountered and to identify the factors explaining observed differences in cyanobacterial community composition according to soil typology (**Chapter 3**).
- 2) The second part contains two research chapters that use whole-genome sequencing of cyanobacterial strains to identify Antarctic-specific genes in Antarctic genomes compared with non-Antarctic ones. First, by comparing genomes of polar *Phormidesmis* (Leptolyngbyaceae) with non-polar genomes, we highlight several genes and metabolic pathways that are more represented in these environments, based on an *a priori* approach using genes of interest (GOIs) (**Chapter 4**). Finally, in the last research chapter, we integrate the notion of Antarctic genomic specificities in a phylogenomic framework using

a discovery-based approach focused on genes clusters (GCs). To this end, the clade of picocyanobacteria, and more specifically the *Synechococcus*-like cyanobacteria of the subcluster 5.2, is studied using phylogenetic, phylogenomic and pangenomic approaches (**Chapter 5**).

Taken together, these results will be discussed in an integrated framework to provide a basis for a better characterization of microbial endemism in Antarctica (**Chapter 6**). Indeed, the concept of endemism is central to demonstrating the uniqueness of Antarctic ecosystems and for providing a scientific basis for their conservation.

1.5.2. Research questions

Each of the four chapters presented above addresses a specific research question:

- What are the biogeographic patterns of (sub-)Antarctic lacustrine cyanobacterial communities in terms of diversity and geographic distribution? (**Chapter 2**)
- What are the drivers shaping terrestrial cyanobacterial communities in the Sør Rondane Mountains? Does soil topology influence community structure? (**Chapter 3**)
- What are the genomic characteristics of the filamentous polar strain *Phormidesmis priestleyi* ULC007? (**Chapter 4**)
- What are the genomic characteristics of Antarctic picocyanobacteria of subcluster 5.2? (**Chapter 5**)

Research chapters - Part one

Antarctica hosts unique ecosystems that harbour many endemic organisms, and its preservation is one of the major priorities of the scientific community. Currently, the most widely used classification of Antarctic terrestrial biomes is the Antarctic Conservation Biogeographic Regions (ACBRs) (Terauds & Lee, 2016). However, these biogeographic regions, which were mostly defined using macroorganismic data, appear not to reflect the distribution of bacteria (**Fig. 1.6B**) (Pessi et al., 2018; Varliero et al., 2024). The lack of continent-scale biogeographic studies therefore represents a limitation for updating this framework for microorganisms.

Although cyanobacteria are among the most extensively studied microorganisms in Antarctica, current knowledge of their biogeography has mainly focused on specific bioregions of the continent (e.g., Taton et al., 2006; Pessi et al., 2018). A study based on a comprehensive continent-scale dataset to test the relevance of the ACBR framework for cyanobacteria has not yet been conducted. This is what we propose to address in **Chapter 2**. At the scale of Antarctic and sub-Antarctic lacustrine environments, we will compare cyanobacterial communities to assess the effects of key environmental variables (pH and salinity), as well as the effect of geographic distance on community structure. We will also compare Antarctic and sub-Antarctic endemism.

To refine the bioregional framework, analyses at a more local scale are also crucial for understanding how microorganisms are structured according to their environment. Thus, in **Chapter 3**, we will study terrestrial communities from the Sør Rondane Mountains (Dronning Maud Land) to highlight how cyanobacterial communities are structured across different soil types. Later, in the general discussion (**Chapter 6**), we will compare these two chapters in order to assess differences in community structure according to habitat type, i.e., lacustrine versus terrestrial environments, as well as patterns of endemism.

Chapter 2

(Sub-)Antarctic endemic cyanobacteria from benthic mats are rare and have restricted geographic distributions

Benoit Durieu^{1,2}, Valentina Savaglia^{1,3}, Yannick Lara⁴, Alexandre Lambion⁴, Igor S. Pessi⁵, Wim Wyverman³, Elie Verleyen³, Annick Wilmotte¹

¹ BCCM/ULC, InBioS–Molecular Diversity and Ecology of Cyanobacteria, University of Liège, 4000 Liège, Belgium

² InBioS–PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, 4000 Liège, Belgium

³ Laboratory of Protistology & Aquatic Ecology, Ghent University, Ghent, Belgium

⁴ Early Life Traces & Evolution-Astrobiology, University of Liège, 4000 Liège, Belgium

⁵ Finnish Environment Institute (Syke), Helsinki, Finland

Manuscript published in **Ecography**, 2025, Volume 2025, Issue 1, e07489.

Author contribution of BD:

BD led the data curation, performed DNA extractions, PCR amplifications, library preparations, bioinformatic analyses, statistical analyses and led the writing of the manuscript.

2.1. Abstract

The Antarctic terrestrial macrobiota are highly endemic and biogeographically structured, but whether this also holds true for microbial groups remains poorly understood. We studied the biogeographic patterns of Antarctic cyanobacteria from benthic microbial mats sampled in 84 lakes from two sub-Antarctic islands, as well as from eight Antarctic Conservation Biogeographic Regions (ACBRs) which were previously defined based mainly on macroscopic taxa. Analysis of 16S rRNA gene sequences revealed that Antarctic and sub-Antarctic lakes host significantly different cyanobacterial communities, yet that the bioregionalization pattern did not correspond to the division into ACBRs. Both Antarctic and sub-Antarctic lakes contain a high number of potentially endemic taxa (41% of the total diversity), of which 33.3% attain a relative abundance of <1%. Our findings highlight the uniqueness of Antarctic microbiota and the need for increased protection of inland waters in both Antarctica and the sub-Antarctic islands.

2.2. Introduction

Antarctic terrestrial macrobiota shows high levels of endemism (De Smet & Gibson, 2007; Pugh & Convey, 2008; McGaughan et al., 2010; Collins et al., 2023), which has been attributed to their long-term survival in ice-free refugia during glacial maxima and the long-term geographic isolation of the continent (Convey & Stevens 2007; Fraser et al., 2014). Continental-scale analyses of various taxonomic groups (Terauds et al., 2012; Dartnall 2017) revealed that the Antarctic terrestrial macrobiota are also biogeographically structured, which has prompted the delineation of 16 Antarctic Conservation and Biogeographic Regions (ACBRs) within the Continental and Maritime Antarctica (Terauds & Lee 2016). The ACBR framework is now considered an integral part of Antarctic science, policy and management, and is frequently used as the null hypothesis in Antarctic biodiversity studies (e.g. Chown et al., 2015; Convey and Peck 2019; Hughes et al., 2019; Colesie et al., 2023).

For the microorganisms studied so-far in Antarctic lakes and soils, results are not consistent between different taxonomic groups. Indeed, lake diatoms at the morphospecies level adhere to this ACBR scheme (Verleyen et al., 2021), while this biogeographic subdivision of the Antarctic does not fully reflect soil bacterial distribution (Varliero et al., 2024). The incidence of endemism might also differ between different taxonomic groups as shown in a recent large-scale comparison of microbiome composition of polar lakes (Tytgat et al., 2023). This bipolar comparison revealed that the number of sequences being unique to the Antarctic differed between bacteria and microeukaryotes, but also between different phyla in both domains of life.

This potential difference in the incidence of endemism and degree of biogeographic structuring between taxa might be partly related to taxon-specific differences in dispersal capacities and e.g. the ability to produce resting stages, go into dormancy and/or to survive wind-mediated dispersal (Tytgat et al., 2023). However, while this bipolar comparison of microbiome composition underscored the unique characteristics of Antarctic microbial communities, the geographic distribution of the sequence data in Tytgat et al. (2023) was not compared with that of related sequences stored in public databases and obtained from (sub-)tropical and temperate regions. This evidently prevents to assess whether the lineages are truly endemic to the Antarctic or, alternatively, have a wider geographic distribution. These comparisons exist only for a few microbial groups, including coccal green algae (De Wever et al., 2009) a terrestrial diatom species complex (Pinseel et al., 2020), and these indeed revealed the presence of distinct lineages or species that are restricted to the Antarctic. Albeit in these comparisons, the geographic coverage and extent of the reference database used evidently influences the assessment of whether taxa appear endemic or not, these DNA-based findings were congruent with a recent study in diatoms which revealed that 44% of the morphospecies are restricted to the Antarctic (Verleyen et al., 2021). For other microbial groups, such region-wide taxonomic inventories using a sufficient amount of samples are, however, still lacking, which prevents to assess their degree of endemism and whether their biogeographies adhere to the ACBR scheme.

A conspicuous feature of many Antarctic lakes is that primary production is largely limited to benthic microbial mats (Vincent, 2000a), which are dominated by filamentous cyanobacteria such as *Leptolyngbya* spp., *Oscillatoria* spp., *Phormidium* spp. and *Phormidesmis* spp. (Taton et al., 2003; Komárek et al., 2008; Vincent & Quesada 2012; Pessi et al., 2018; 2023). Some cyanobacterial taxa/OTU appear to be endemic to Antarctica (Jungblut et al., 2010, Komárek et al., 2012, Kleinteich et al., 2017) and novel taxa have yet to be discovered, as recently revealed by Pessi et al. (2023) using a metagenomics approach. Other cyanobacterial lineages from Antarctic lakes have been observed in the Arctic or other cold habitats such as alpine lakes, whereas others have a cosmopolitan distribution (e.g. they were reported from polar, temperate and tropical regions; Kleinteich et al., 2017; Pessi et al., 2018; 2023).

Only a few studies have compared the obtained sequences with all data present in public databases (i.e. GenBank) to determine their degree of endemism (e.g. Pessi et al., 2018). However, information on the biodiversity and incidence of endemism for lacustrine microorganisms is crucial for developing urgently-needed and science-based conservation strategies for Antarctic habitats in ice-free regions, which is needed for inland lake (Hawes et al., 2023) and edaphic ecosystems (Lebre et al., 2023).

Here, we aimed to provide a more complete picture of the diversity and community structure of benthic lacustrine cyanobacteria based on the analysis of 84 lakes' samples collected across the Continental and Maritime Antarctic regions in 8 ACRs, and in two sub-Antarctic islands. We specifically tested whether the distribution of Cyanobacteria is congruent with the ACRs framework. We hypothesize (1) that cyanobacterial communities differ between Antarctic (i.e. Continental and Maritime Antarctica) and the sub-Antarctic lakes due to their contrasting environmental characteristics and geographic isolation yet that their biogeographic structuring is incongruent with the ACRs scheme, in contrast to other life forms that are probably more dispersal-limited (e.g. macrobiota and diatoms; Convey et al., 2014; Schulte et al., 2022); and (2) that the degree of endemism is higher in the Antarctic lakes compared to those in the sub-Antarctic because long-distance dispersal in the latter region is aided by the Southern Hemisphere Westerly winds as shown for moss, liverwort, lichen, and pteridophyte floras (Muñoz et al., 2004).

2.3. Material and Methods

2.3.1. Sampling and measurements

Eighty-nine microbial mats were collected from 84 lakes or ponds across 8 ACRs from Antarctica (n = 75), and the sub-Antarctic Macquarie (n = 6) and Marion (n = 8) Islands from the Southern Ocean (**Fig. 2.1A**). Antarctica is classically divided into two zones with a different climate and separated by the Gressitt Line, namely the Maritime Antarctic (Peninsula) and the Continental Antarctic (Chown & Convey, 2007). Our dataset covers 2 and 6 ACRs, in the Maritime and Continental Antarctic, respectively. A detailed description of all the sampled sites can be found in **Table S2.1**.

The mats were sampled during the period 1993-2013 during field expeditions of national and international research programs following a standardized protocol (e.g. Sabbe et al., 2004). Briefly, visible microbial mats were collected from the littoral zone (~ 20-50 cm depth) using a sterilized spatula. All samples were immediately stored at -20°C and kept frozen upon DNA extraction. Measurements of conductivity and pH were conducted on 79 and 71 samples, respectively, as described in Pessi et al. (2018).

2.3.2. DNA extraction, PCR and sequencing

Partial 16S rRNA sequences were obtained as described by Pessi et al. (2016). Briefly, DNA was extracted using the PowerSoil DNA Isolation kit (MOBIO, Carlsbad, CA, USA). The V3-V4 region was targeted by PCR using cyanobacteria-specific primers CYA359F and CYA781Ra/CYA781Rb (Nübel et al., 1997). PCR mixture consisted of 1x PCR buffer with 1.5 mM MgCl₂, 1 mg.ML⁻¹ BSA, 200 µM of each dNTP, 0.2 µM of each primer, 1 U SUPER TAQ plus DNA polymerase (HT Biotechnology, Cambridge, UK), and 4 ng.µL⁻¹ template DNA in a final volume of 50 µL. PCR amplification was carried out with an initial denaturation step at 94°C for 5 min, followed by 34 cycles of 94°C for 1 min, 57°C for 1 min and 68°C for 1 min, and a final elongation at 68°C for 5 min. For 5 randomly chosen samples, PCR duplicates were made to validate the robustness of the workflow. For each sample, three independent PCR reactions were pooled, purified (GeneJet PCR Purification Kit, ThermoScientific, Waltham, MA, USA) and quantified (Quant-iT PicoGreen dsDNA Assay Kit, Life Technologies, Carlsbad, CA, USA). Libraries were pooled in equimolar concentrations and sent to Genewiz (South Plainfield, NJ, USA) where ligation of sequencing adapters and sequencing was done using the Illumina MiSeq platform (Illumina, San Diego, CA, USA) with 2x300 bp paired-end libraries.

2.3.3. Bioinformatic analysis to obtain quality filtered cyanobacterial OTUs

A read quality control and chimeric sequences filtering protocol based on UPARSE (Edgar 2013) was adapted from Pessi et al. (2016). Briefly, the paired-end reads were merged and filtered (--maxdiffs 10, --maxdiffpct 8), and reads that did not contain primer and barcodes in the 3' and 5' ends were removed. Two and zero mismatches were allowed for primer and barcode sequences, respectively. Reads with a maximum expected error of more than 0.5 and smaller than 370 bp were removed. After singleton removal, the remaining sequences were denoised to remove chimeras and sequencing errors using unoise3 (Edgar, 2016). The resulting denoised zero-radius operational taxonomic units (ZOTUs) were then clustered into OTUs using a 99 % similarity threshold (Edgar 2013). The most abundant sequence in each OTU cluster was selected as the representative sequence. OTUs with <10 reads in the whole OTU table were finally removed.

After occurrence inspection of the PCR duplicate samples (**Fig. S2.1**), mock communities and PCR duplicate samples were removed from the OTU table and only one sample with the highest read number was kept among PCR duplicate of the same sample.

2.3.4. Taxonomic affiliation and geographic distribution of OTUs

OTUs were first classified with the RDP classifier (Wang et al., 2007) implemented in QIIME pipeline (Caporaso et al., 2010). The taxonomic classification of cyanobacterial OTUs was subsequently refined by comparison with sequences from the GenBank nucleotide (nt) collection (accessed on 9 April 2024). For this, closely related sequences ($\geq 94.5\%$ similarity, corresponding to the *genus* taxonomic rank according to Yarza et al., 2014) were retrieved using BLAST 2.8.1+ (Boratyn et al., 2013). The sequences with the highest similarity to each OTU were manually selected. To address the taxonomy of “uncultured bacterium” or “uncultured cyanobacterium”, as well as to remove potential chimeras, a phylogenetic tree was built using RaxML version 8 (Stamatakis 2014) including the cyanobacterial OTUs, their close relative sequences taken from NCBI (nt) database and the CyanoSeq database (Lefter et al., 2023), which consist of (near) full length 16S rRNA gene sequences with curated taxonomy (**Fig. S2.2**). Based on their taxonomy (Komárek et al., 2014) and phylogenetic position, a morphologic trait was subsequently assigned to each OTU: “unicellular”, “thin filamentous”, “large filamentous”, “heterocytous” and “unknown” (i.e. in the case of sequences with taxonomy above family rank, for which the morphological trait could not be assigned).

The geographic distribution of the OTUs was inferred based on the distribution of similar sequences ($\geq 99\%$ similarity) from GenBank. They were categorized as follows: “endemic” (i.e. OTUs sharing less than 99% of similarity with any other published sequences or sharing at least 99% of similarity with sequences only reported from the southern polar region); “polar” (i.e. OTUs sharing more than 99% of similarity with sequences from the Arctic and the Antarctic), “alpine” (i.e. OTUs sharing more than 99% of similarity with sequences from polar and high altitude regions) and “cosmopolitan” (i.e. OTUs sharing more than 99% of similarity with at least one sequence from non-polar and non-alpine regions) (**Table S2.2**). These definitions are dependent of the geographic coverage of the public databases and the genetic marker used, knowing that only the cosmopolitan distribution can certainly be proven.

In order to examine the most abundant and the rarest OTUs fraction of Antarctica (i.e. Continental and Maritime Antarctica) and sub-Antarctic islands, OTUs encompassing $>90\%$ of the total OTU counts were considered as the “abundant fraction” whilst those encompassing $<1\%$ of the total OTU counts were considered as the “rare fraction”.

2.3.5. Analysis of cyanobacterial community structure

Beta diversity analyses were performed using R version 4.1.0 (R Core Team, 2021). Normal distribution and homoscedasticity of pH and conductivity values were tested through the Shapiro-Wilks test and the Bartlett's test using the package *stats* (version 4.1.0), whilst skewness and kurtosis were investigated with the package *moments* (version 0.14.1; Komsta & Novomestky, 2015). For conductivity, the dataset was $\log(x+1)$ transformed in order to attenuate skewness. Finally, the OTU abundance table was Hellinger transformed prior to multivariate analysis to minimize the effect of differences in library size (Legendre et al., 2005; Peres-Neto et al., 2006).

To visualize community structure relationships, the Hellinger transformed abundance OTU table was converted into a dissimilarity matrix using Bray-Curtis similarities and a nonmetric multidimensional scaling (NMDS) ordination was done using the *metaMDS* function ('vegan' package, version 2.5-7; Oksanen et al., 2019). Then, the *ordisurf* function ('vegan' package) was used to fit the latitude on the ordination in order to determine its correlation (Wood, 2003; Marra & Wood, 2011). The *ordisurf* function uses a generalized additive model (GAM) to model the latitude on the metaMDS ordination, with the function *gam* ('mgcv' package, version 1.8-35; Wood et al., 2016). To test for significant differences between the communities, PERMANOVA was implemented using the *pairwise.adonis* function ('pairwiseAdonis' package, version 0.4; Martinez Arbizu, 2020).

The effect of spatial parameters on cyanobacterial community structure was analysed using principal coordinates of neighborhood matrix (PCNMs; Borcard & Legendre, 2002) as described by Sakaeva et al. (2016). Briefly, the latitude and longitude of each sample were first used to calculate a geographic distance matrix using the function *distGeo* of the command *distm* ('geosphere' package version 1.5-14; Hijmans 2014). This matrix was used to calculate PCNM eigenvectors using the function *pcnm* ('vegan' package) with the smallest PCNM (PCNM1) representing large spatial scales and the biggest ones (PCNM*) fine spatial scales.

To explain variations in community structure, we performed redundancy analysis (RDA). The variation partitioning (Borcard et al., 1992; Peres-Neto et al., 2006) was used to determine the relative importance of geographic distance (spatial gradient), conductivity and pH on cyanobacterial community structure as in Sokol et al. (2013). First, a distance-based redundancy analysis (dbRDA; McArdle & Anderson, 2001) was used to model variation in community structures as a function of either environmental (pH and conductivity) or spatial (PCNMs) variables. Forward stepwise model selection (Blanchet et al., 2008) based on adjusted R^2 (Beisner

et al., 2006; Peres-Neto et al., 2006; Nabout et al., 2009; Legendre & Legendre 2012) was used to select PNCMs that best explained variation in community structures. Then, the PNCMs were combined with the environmental variables into a single dbRDA model to calculate the total variation in community structures, using the function *varpart* ('vegan' package). The adjusted R^2 values give us the among site variation in community structures explained by both environmental and spatial variables [E+S], by pure environmental [E|S], spatially structured by the environment [EnS], pure spatial [S|E] or unexplained.

2.4. Results

Illumina sequencing resulted in 3,981,506 reads of which 90.5% was retained after quality filtering. After further processing (dereplication, removal of singletons, chimeric and non-cyanobacterial reads, and removal of OTUs with <10 reads per sample), 310 cyanobacterial OTUs (99% similarity) representing 45.3% of the original bacterial reads were kept. Of the 310 OTUs, 146 OTUs (47.1%) were detected in both Antarctica (which includes both Continental and Maritime Antarctica) and the sub-Antarctic islands, 133 OTUs (42.9%) were restricted to Antarctica and 31 OTUs (10%) were detected only in the sub-Antarctic islands (**Table S2.2**).

2.4.1. Cyanobacteria differ between Antarctica and the sub-Antarctic islands, but not between ACBRs

The cyanobacterial communities differed between Antarctica and the sub-Antarctic islands as shown by the NMDS analysis (**Fig. 2.1A**), of which the first axis is correlated with latitude (Pearson correlation with NMDS1: $R^2 = 0.78$, $p < 0.01$). These differences were confirmed by PERMANOVA, which showed that the sub-Antarctic communities differed significantly from those in Continental Antarctica (pairwise.adonis: $r^2 = 0.081$, adjusted p-value = 0.003) and in Maritime Antarctica (pairwise.adonis: $r^2 = 0.120$, adjusted p-value = 0.006) (**Table S2.3**). Continental and Maritime Antarctic communities were also significantly different (pairwise.adonis: $r^2 = 0.0342$, adjusted p-value = 0.018), yet the bioregionalization pattern within Antarctica did not mirror the ACBR framework (**Fig. 2.1B**). Variation partitioning revealed that 14.8% of cyanobacterial community structure was explained by spatial parameters and only 4.9% by environmental factors (pH and conductivity) (**Table S2.4**). The significant spatial parameters were PNCMs reflecting large scale geographical distances (i.e. between Antarctica and the sub-Antarctic

Islands). In addition, variation partitioning analysis showed that latitude was included in the PCNMs component (**Fig. 2.2**).

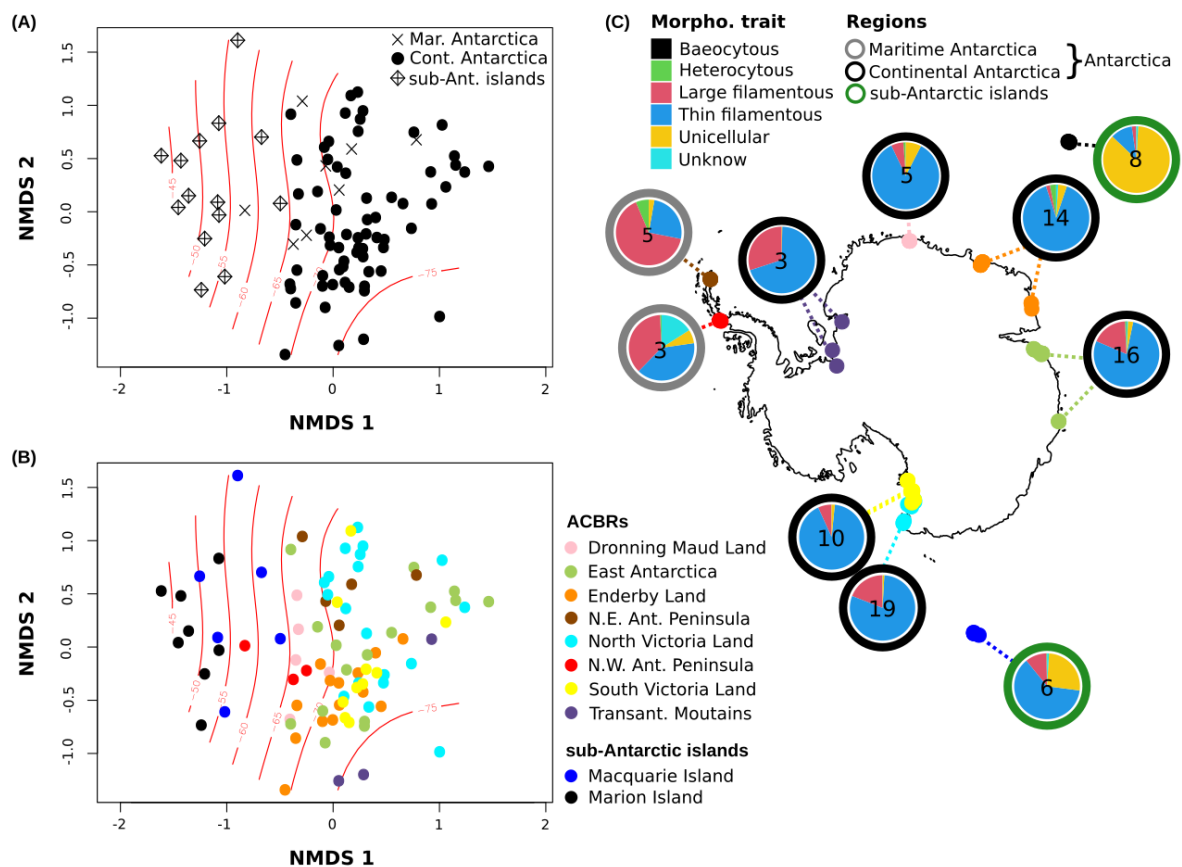


Fig. 2.1. A. metaMDS on the Bray-Curtis similarity matrix based on Hellinger transformed abundance data based on OTUs. Shapes represent the three main regions: Maritime Antarctica, Continental Antarctica and sub-Antarctic islands (see legend for details). Red curves represent the ordisurf function (R package ‘vegan’) fitting of generalized additive model (GAM) of latitude on the metaMDS ordination. **B.** metaMDS as in (A). Colors represent the different ACBRs or sub-Antarctic regions investigated (see legend). **C.** Map showing the sampling locations. The pie charts present the relative abundance of the different morphological types in each ACBRs. The numbers at the center of the pie charts represent the number of samples investigated in this study for each region.

The dominant genera also differed between Antarctica and the sub-Antarctic islands as revealed by the taxonomic assignment of OTUs morphotypes (**Fig. 2.1C, Table S2.2**). Sequences related to thin filamentous cyanobacteria of the genera *Leptolyngbya* (32% of the reads) and *Phormidesmis* (23.8%) were dominant in Continental Antarctica, while sequences related to large filamentous cyanobacteria of the genera *Phormidium* (6.1%) were more abundant in Maritime Antarctica. By contrast, sequences related to unicellular cyanobacteria of the genera *Cyanobium* (28.7%) and *Synechococcus* (16.4%) dominated the communities in the sub-Antarctic islands, and particularly Marion Island.

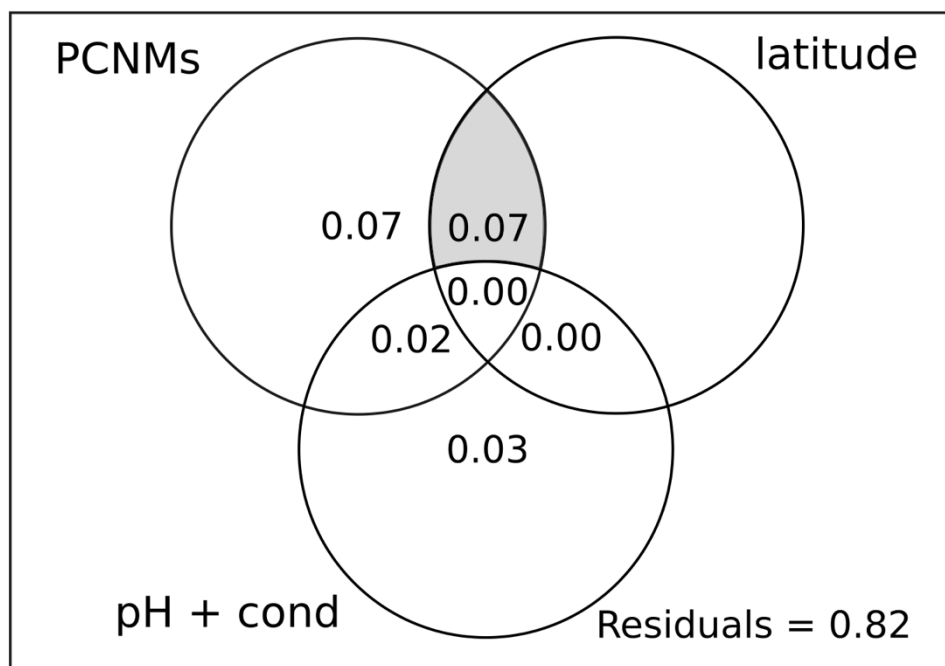


Fig. 2.2. Addition of latitude in the variation partitioning revealing its correlation with spatial parameter (in grey: 7% on the 14% of PCNMs). PCNMs (Principal Coordinate of Neighborhood Matrix) are eigenvectors ranged from large spatial scales (PCNM1) to fine spatial scale (PCNM*). Variation partitioning was calculated in R studio using function varpart ('vegan' package, version 2.5-7; Oksanen et al., 2019).

2.4.2. The distribution patterns and taxonomic composition differ between the abundant and rare fractions

The taxonomic composition of the rare fraction (<1% relative abundance) was markedly different from that of the abundant fraction (>90% relative abundance; **Fig. 2. 3A, Table S2.2**), as the rare fraction encompassed a higher number of heterocystous and unicellular morphotypes than the abundant fraction. The latter fraction in the Antarctic lakes was characterized by filamentous genera such as *Leptolyngbya* (34.3% of the number of reads), *Phormidesmis* (24.7%), *Stenomitos* (8.6%) and *Phormidium* (5.9%), whilst it was the unicellular *Cyanobium* (28.7%), the filamentous *Phormidesmis* (24.5%) and the unicellular *Synechococcus* (15.6%) for the sub-Antarctic lakes.

With no clear differences between ACBRs (**Fig. S2.3**), around 46% of the OTUs present in the Antarctic and sub-Antarctic lakes were cosmopolitan, whilst 41% were considered as endemic, as revealed by the manually curated blastn search performed in NCBI nucleotide (nt) database. This analysis allowed to retrieve the original information of each publicly available sequence with which our OTUs shared above 99% of similarity. Although the method is limited by the geographic coverage of the database and the genetic marker used, we considered an OTU as potentially

“endemic” if they shared less than 99% of similarity with any other sequences or if it shared more than 99% of similarity with sequences which were previously only detected in the Antarctic.

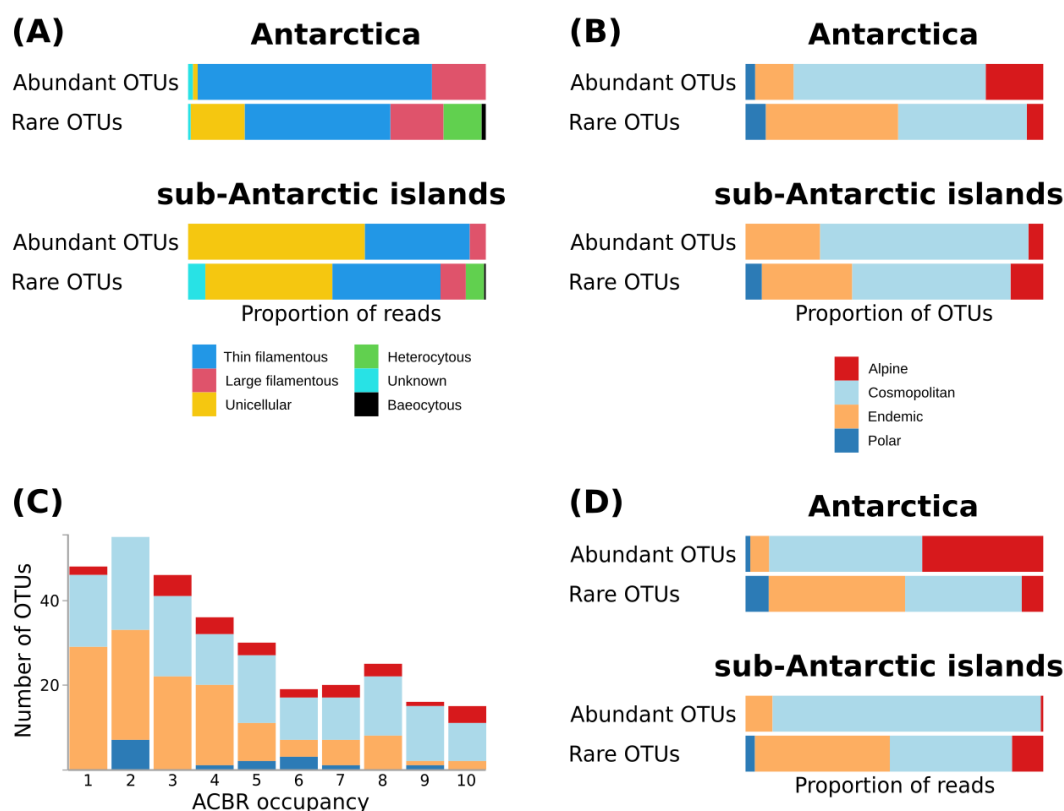


Fig. 2.3. Comparison of the 75 Antarctic and 14 sub-Antarctic cyanobacterial communities: **A.** The relative abundance of the different morphotypes of ‘abundant’ OTUs (i.e. OTUs encompassing >90% of the total OTU counts) and the ‘rare fraction’ (i.e. those encompassing <1% of the total OUT counts) in Antarctic (upper panel) and sub-Antarctic lakes (lower panel). Colors represent the OTU morphotypes. “Unknown” morphotype corresponds to OTUs whose taxonomic rank identification was too high (e.g. family) to discriminate between one or another morphological trait; **B.** The proportion of ‘endemic’, ‘alpine’, ‘polar’ and ‘cosmopolitan’ OTUs (expressed as number of OTUs) in the ‘abundant’ and ‘rare’ fraction of Antarctic and sub-Antarctic lakes; **C.** The number of ‘endemic’, ‘alpine’, ‘polar’ and ‘cosmopolitan’ OTUs occurring in 1, 2, ...10 ACBRs (i.e., the occupancy); **D.** The proportion of ‘endemic’, ‘alpine’, ‘polar’ and ‘cosmopolitan’ OTUs (expressed as number of reads) in the ‘abundant’ and ‘rare’ fraction. The color code in b-d represents the geographic distribution of the OTUs.

In this way, we observed that the geographic distribution of sequences belonging to the rare fraction was markedly different from that of members from the abundant fraction (**Fig. 2. 3B**, **Table S2.5**).

Endemic OTUs were more prevalent within the rare fraction whereas the abundant fraction consisted mostly of cosmopolitan OTUs. This difference was particularly prominent in the Antarctic lakes. The abundant fraction of the Antarctic lakes was represented by taxa with a cosmopolitan distribution (64.5% of the number of OTUs), followed by sequences with an alpine

(19.4%), endemic (12.9%) and polar (3.2%) distribution. The abundant fraction of the sub-Antarctic lakes was characterized by an even higher proportion of taxa with a cosmopolitan distribution (70%), followed by those with an endemic (25%) and alpine (5%) distribution. By contrast, the rare fraction in Antarctic lakes was characterized by OTUs with an endemic distribution (44.4% of the OTUs), followed by cosmopolitan (43.2%), polar (6.8%) and alpine (5.6%) distributions. The rare fraction in the sub-Antarctic islands encompassed a higher proportion of OTUs with a cosmopolitan distribution (53.2%), followed by those with an endemic (30.3%), alpine (11%) and polar (5.5%) distribution. Interestingly, when considering the relative abundances of the sequences instead of the number of OTUs (**Fig. 2. 3D, Table S2.5**), up to half of the abundant fraction in the Antarctic lakes is characterized by OTUs with a cold distribution (alpine: 40.7%; endemic: 6.3%; polar: 1.6%) whilst the vast majority of the sequences in the abundant fraction in the sub-Antarctic lakes consist of cosmopolitan taxa (90.1%). Finally, the analysis of ACBRs and sub-Antarctic islands occupancy showed that endemic OTUs are mostly restricted to a few ACBRs or islands, compared to cosmopolitan ones that are widespread across Antarctica (**Fig. 2. 3C**).

2.5. Discussion

2.5.1. Cyanobacterial communities are different between Antarctica and the sub-Antarctic islands

Our analyses revealed significant differences in cyanobacterial community composition and structure between Antarctic and sub-Antarctic lakes. These marked differences were also observed in other taxonomic groups such as terrestrial nematodes, mites and springtails (Chown & Convey, 2007), as well as in aquatic invertebrates (Dartnall, 2017) and other freshwater (Verleyen et al., 2021; Tytgat et al., 2023) and edaphic microorganisms (Lebre et al., 2023; Varliero et al., 2024). Our results also showed that the taxonomic composition, corresponding to specific morphological traits of the cyanobacteria in these mats, is different between Antarctica and the sub-Antarctic islands. Thin (1-3 μm) filamentous taxa (e.g. *Leptolyngbya*, *Phormidesmis*) followed by larger filamentous taxa (e.g. *Phormidium*, *Microcoleus*) dominated Antarctic lakes, whilst unicellular taxa (e.g. *Chamaesiphon*, *Cyanobium* and *Synechococcus*) dominated the sub-Antarctic lakes in Marion Island and were also particularly abundant, although not dominant, in the sub-Antarctic lakes in Macquarie Island. These findings agree with previous observations which reveal that cyanobacterial taxa dominate benthic mats, with filamentous taxa (e.g. *Leptolyngbya*, *Phormidium* and *Stenomitos*) particularly prominent in Antarctic lakes (Vincent

2000b; Taton et al., 2006; Fernandez-Carazo et al., 2011; Pessi et al., 2018; 2023; Sumner et al., 2016; Jungblut & Vincent 2017; Jackson et al., 2021). The high cyanobacterial abundance in benthic mats has been attributed to the general lack of large metazoan grazers (e.g. crustaceans) compared with lower latitude lakes (Zawierucha et al., 2019; Tytgat et al., 2023 and references therein). This lack of top-down control by grazers results in low levels of bioturbation of sediments and biofilms (Laybourn-parry et a. 1991; Jungblut et al., 2012) in Antarctic lakes promoting the formation of perennial microbial mats, which are physically structured by filamentous cyanobacteria and micro-algae (Tytgat et al., 2023). A higher proportion of unicellular OTUs in the sub-Antarctic islands compared to the Antarctic lakes was also encountered by Tytgat et al., (2023) and may be due to the higher planktonic primary production and/or the higher bioturbation and grazing by larger metazoan invertebrates in the sub-Antarctic islands compared to Antarctica. The higher planktonic primary production in the sub-Antarctic lakes is the result of the complete mixing of their water columns due to the lack of seasonal lake ice cover and high wind speeds, which keep non-motile planktonic taxa in suspension (Vincent & Laybourn-Parry, 2008). Moreover, in sub-Antarctic lakes, mosses and sometimes macrophytes cover the bottom of lakes and fill the niche occupied by the benthic cyanobacteria in Antarctic lakes (Smith & Ashton, 1971; Douce et al., 2023). Velázquez et al. (2017) also found differences in the cyanobacterial composition during ice-covered versus ice-free periods of lakes in Byers Peninsula, with unicellular taxa (e.g. *Synechococcus* spp.) occurring especially during the ice-free season.

2.5.2. Cyanobacterial communities are not clearly differentiated among ACBRs

No clear difference in cyanobacterial community structure was observed between the ACBRs. However, the communities were significantly different, albeit with a relatively high p-value, between the Maritime Antarctic and Continental Antarctica. However, this marginally significant difference might be related to the sampling effort, which was lower in Maritime Antarctica, hindering to reveal a more significant difference in community structure compared to Continental Antarctica. The absence of differences in cyanobacterial community structure among the ACBRs is in agreement with a biogeographic survey of soil bacteria which revealed clear differences between Antarctica and the sub-Antarctic islands, but not between the ACBRs and some difference but also overlap between Maritime and Continental Antarctica (Varliero et al., 2024). By contrast, the distribution of lake diatom species is congruent with the ACBR scheme (Verleyen et al., 2021). This difference can be in part attributed to differences in dispersal ability and life cycle characteristics between different freshwater taxa. Cyanobacteria are able to survive desiccation and other extreme conditions (Sakamoto et al., 2009) experienced during wind-

mediated dispersal, compared to, for example, lake diatoms (Souffreau et al., 2010). It has been shown that some Antarctic cyanobacteria have an ‘escape mechanism’ from the lakes, as they form so-called ‘lift-off’ mats (Dudeja et al., 2010), allowing benthic cyanobacteria to move from the bottom of the lake to the surface and afterwards disperse to other terrestrial and aquatic environments (Wharton et al., 1983; Tanabe & Kudoh, 2012). This might facilitate the colonization of different regions within the Antarctic continent (Herbold et al., 2014; Chown et al., 2015) and hence explain the lack of a clear structuring into the ACBRs. This efficient dispersal mechanism combined with a high resistance to the harsh conditions can also explain why quite a large number of the OTUs occur in all or nearly all ACBRs. Interestingly, the majority of the OTUs occurring in less than 5 ACBRs or sub-Antarctic islands are mostly endemic whilst those occurring between 5 and 10 ACBRs or sub-Antarctic islands were mainly cosmopolitan taxa. One scenario could be that these cosmopolitan taxa have colonized a lake system after its formation due to glacial retreat or isostatic uplift (Verleyen et al., 2012). These colonizing taxa might also be well-adapted for dispersal within Antarctica, leading to a quick spread between the different ACBRs and not leaving any footprint of biogeographical zonation at the 16S rRNA level considered in the present study. High incidence of endemism in Antarctic cyanobacteria, yet endemic taxa are rare and have restricted distributions

A total of 40.6% of all OTUs appear endemic to Antarctica or the sub-Antarctic islands. This level of endemism is comparable with lacustrine diatoms in both regions (Verleyen et al., 2021), with lichens in Antarctica (33–50%; Peat et al., 2007). It is however clearly lower than the incidence of endemism found in Antarctic springtails (90%; Baird et al., 2019), bdelloid rotifers (95% endemics in Antarctica; Iakovenko et al., 2015) and tardigrades (73%; Guidetti et al., 2019). The degree of endemism in cyanobacteria might, however, be an underestimation because 16S rRNA gene might not be variable enough and the used amplicon too short to delineate taxa at the species level (Cho & Tiedje, 2000), despite the fact that Taton et al. (2003) tested that the 16S rRNA V3-V4 region was representative of the complete sequence for cyanobacteria. On the other hand, this assessment of the endemism degree might also be an overestimation as the public database used to assess the geographic distribution of OTUs might not be complete. We are aware that our caring selection and characterization of OTUs (e.g. chimeric removing, removing OTUs ≤ 10 reads, blastn approach; see Material & Methods for details) might not fully solve these problems. However, despite these potential limits, our study showed that the endemic cyanobacteria only accounted for 9.4% of the total read counts. Moreover, while the most abundant OTUs were generally cosmopolitan, 68% of the endemic OTUs were rare in Antarctica and 56% in the sub-Antarctic islands. This has evidently important consequences for the conservation of these biota

in the face of climate change and increased human impacts (Chown et al., 2015) as rare taxa are evidently more difficult to protect than those with larger population sizes which are less prone to extinction (Hughes et al., 2015; Wauchope et al., 2019). However, protecting the rare biosphere is crucial, because those taxa may have important ecological roles, and serve as a large reservoir of genetic and functional diversity which may provide resilience to change (Lynch & Neufeld, 2015). What also became evident from this study is that the abundant and widespread OTUs in the sub-Antarctic islands had a cosmopolitan distribution, whereas in Antarctica, relatively more OTUs were restricted to cold environments (i.e. cold-ecotypes: endemic, polar or alpine distribution). Moreover, the three most abundant OTUs in Antarctica had an polar and alpine distribution. This is in line with previous observations about cold-habitat-specific cyanobacterial assemblages dominating Antarctica and often absent in other climatic zones (Jungblut et al., 2010; Kleinteich et al., 2017), which could be explained by a higher ecological success of taxa adapted to general cold and extreme conditions typical of Antarctic, Arctic and alpine environments (Hessen, 2007). Moreover, in both sub-Antarctic and Antarctic lakes, the abundant fraction was characterized by only a few but widespread and cosmopolitan OTUs, while the rare fraction encompassed mainly endemic OTUs with a limited range size. A possible explanation is that large population sizes may increase the probability of dispersal (Nemergut et al., 2011) and that some cyanobacterial taxa may be better adapted for dispersal via, e.g., aeolian transport (Herbold et al., 2014) than others. Alternatively, but data to fully test this is lacking, the rare and endemic OTUs might be the result of recent speciation events (Tytgat et al., 2023) and the time since speciation might have been too short for widespread dispersal within the Antarctic in contrast to the cosmopolitan taxa. In turn, this, as well as differences in dispersal capacities, could explain the higher proportion of cosmopolitan taxa being present in several ACBRs compared to the endemic OTUs being restricted to a few ACBRs. As already discussed, the cosmopolitan OTUs might have high dispersal capacities which might have resulted in their widespread distribution in the Antarctic and the lack of clear biogeographic structuring. In a climate change scenario, these widespread and cosmopolitan taxa might therefore be more likely to colonize new habitats.

2.5.3. The need for protecting the sub-Antarctic islands

Microbial communities in sub-Antarctic islands are poorly studied and have only recently been included in biogeographic studies of lakes (Tytgat et al., 2023; Verleyen et al., 2021) and soils (Lebre et al., 2023; Varliero et al., 2024). We encountered a relatively high degree of potential endemism in the sub-Antarctic lacustrine cyanobacterial communities, which has never been

previously reported. However, the incidence of endemism appeared to be lower in the sub-Antarctic islands than in the continent, possibly due to wind-driven dispersal between the islands and land masses at the temperate latitudes in the Southern Hemisphere was shown for cryptogamic floras (Muñoz et al., 2004). Tygtat et al. (2023) also revealed that a lower proportion of microorganisms was endemic in the sub-Antarctic lakes compared to the Antarctic systems. Similarly, the incidence of endemism in diatoms at the morphospecies level is also lower in sub-Antarctica compared with the Antarctic continent (Verleyen et al., 2021). However, given that the incidence of endemism is relatively high (34%) even compared with the estimations of cyanobacteria endemism to Antarctica (38%), our study further justifies including the sub-Antarctic islands surrounding the Antarctic continent to the ACBRs rationale, as already proposed by Lebre et al. (2023). The natural barriers occurring within the sub-Antarctic islands and with the rest of the world and the Antarctic are being progressively degraded through human activities (Perterra et al., 2017; Hughes et al., 2019). Therefore, conservation efforts should be enhanced to avoid human-induced introductions of non-native microbes in both Antarctica and the sub-Antarctic islands and homogenization of microbial floras between ice-free regions of the continent. Recommendations to minimize the human impact are given in the SCAR code of conduct for terrestrial research (Scientific Committee on Antarctic Research [SCAR], 2009; 2011).

2.6. Conclusions

In this study, we showed that the cyanobacterial community composition is different between Antarctic and sub-Antarctic lakes, but not among the ACBRs, which were generally delineated based on macroscopic organisms. We found a relatively high number of endemic taxa in both the Antarctic and the sub-Antarctic lakes, with a higher proportion in Antarctica. Also, a high amount of the encountered endemic taxa was rare, highlighting the urge for their protection. Finally, this study provides justifications to include the sub-Antarctic islands in the ACBRs rationale in order to improve the conservation measures of these regions.

2.7. References

- Baird, H.P., Janion-Scheepers, C., Stevens, M.I., Leihy, R.I. and Steven Loudon Chown (2019). The ecological biogeography of indigenous and introduced Antarctic springtails. 46(9), pp.1959–1973. doi: 10.1111/jbi.13639.
- Beisner, B.E., Peres-Neto, P.R., Lindström, E.S., Barnett, A. and Longhi, M.L. (2006). The role of environmental and spatial processes in structuring lake communities from bacteria to fish. *Ecology*, 87(12), pp.2985–2991. doi: 10.1890/0012-9658(2006)87[2985:troeas]2.0.co;2.
- Blanchet, F.G., Legendre, P. and Borcard, D. (2008), forward selection of explanatory variables. *Ecology*, 89: 2623–2632. doi: 10.1890/07-0986.1.
- Boratyn, G.M., Camacho, C., Cooper, P.S., Coulouris, G., Fong, A., Ma, N., Madden, T.L., Matten, W.T., McGinnis, S.D., Merezuk, Y., Raytselis, Y., Sayers, E.W., Tao, T., Ye, J. and Zaretskaya, I. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Research*, [online] 41(W1), pp.W29–W33. doi: 10.1093/nar/gkt282.
- Borcard, D, Legendre, P (2002), All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices, *Ecological Modelling*, 153(1–2): 51–68. doi: 10.1016/S0304-3800(01)00501-4.
- Borcard, D., Legendre, P. and Drapeau, P. (1992). Partialling out the Spatial Component of Ecological Variation. *Ecology*, 73(3), 1045–1055. doi: 10.2307/1940179.
- Caporaso, J., Kuczynski, J., Stombaugh, J. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7, 335–336 (2010). doi: 10.1038/nmeth.f.303.
- Cho, J.-C. and Tiedje, J.M. (2000). Biogeography and Degree of Endemicity of Fluorescent *Pseudomonas* Strains in Soil. *Applied and Environmental Microbiology*, 66(12), pp.5448–5456. doi: .1128/aem.66.12.5448-5456.2000.
- Chown, S.L. and Convey, P. (2007). Spatial and temporal variability across life's hierarchies in the terrestrial Antarctic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1488), pp.2307–2331. doi: 10.1098/rstb.2006.1949.
- Chown, S.L. & Convey, P. (2007). Spatial and temporal variability across life's hierarchies in the terrestrial Antarctic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1488), pp.2307–2331. doi: 10.1098/rstb.2006.1949.
- Chown, S.L., Clarke, A., Fraser, C.I., Cary, S.C., Moon, K.L. and McGeoch, M.A. (2015). The changing form of Antarctic biodiversity. *Nature*, 522(7557), pp.431–438. doi: 10.1038/nature14505.
- Colesie, C., Walshaw, C.V., Sancho, L.G., Davey, M.P. and Gray, A. (2023). Antarctica's vegetation in a changing climate. *Wiley Interdisciplinary Reviews: Climate Change*, 14(1), p.e810. doi: 10.1002/wcc.810.

Collins, G.E., Young, M.R., Convey, P., Chown, S.L., S. Craig Cary, Adams, B.J., Wall, D.H. and Hogg, I.D. (2023). Biogeography and Genetic Diversity of Terrestrial Mites in the Ross Sea Region, Antarctica. *Genes*, 14(3), pp.606–606. doi: 10.3390/genes14030606.

Convey, P. and Peck, L.S. (2019). Antarctic environmental change and biological responses. *Science Advances*, 5(11), p.eaaz0888. doi: 10.1126/sciadv.aaz0888.

Convey, P. and Stevens, M.I. (2007). ECOLOGY: Antarctic Biodiversity. *Science*, 317(5846), pp.1877–1878. doi: 10.1126/science.1147261.

Convey, P., Chown, S.L., Clarke, A., Barnes, D.K.A., Bokhorst, S., Cummings, V., Ducklow, H.W., Frati, F., Green, T.G.A., Gordon, S., Griffiths, H.J., Howard-Williams, C., Huiskes, A.H.L., Laybourn-Parry, J., Lyons, W.B., McMinn, A., Morley, S.A., Peck, L.S., Quesada, A. and Robinson, S.A. (2014). The spatial structure of Antarctic biodiversity. *Ecological Monographs*, 84(2), pp.203–244. doi: 10.1890/12-2216.1.

Dartnall H. (2017). The freshwater fauna of the South Polar Region: A 140-year review. *Papers and proceedings of the Royal Society of Tasmania*, 151, pp.19–57. doi: 10.26749/rstpp.151.19.

De Smet, W.H. and John A.E. Gibson (2007). *Rhinoglena kutikova* n.sp. (Rotifera: Monogononta: Epiphanidae) from the Bunger Hills, East Antarctica: a probable relict species that survived Quaternary glaciations on the continent. *Polar Biology*, 31(5), pp.595–603. doi: 10.1007/s00300-007-0393-9.

De Wever, A., Leliaert, F., Verleyen, E., Vanormelingen, P., Van der Gucht, K., Van Hodgson, D.A., Sabbe, K. and Wim Vyverman (2009). Hidden levels of phylodiversity in Antarctic green algae: further evidence for the existence of glacial refugia. *Proceedings of The Royal Society B: Biological Sciences*, 276(1673), pp.3591–3599. doi: 10.1098/rspb.2009.0994.

Douce, P., Mermillod-Blondin, F., Simon, L., Dolédec, S., Eymar-Dauphin, P., Renault, D., Sulmon, C., Vallier, F. and Bittebiere, A.-K. (2023). Biotic and abiotic drivers of aquatic plant communities in shallow pools and wallows on the sub-Antarctic Iles Kerguelen. *Polar Biology*, 46(4), pp.303–318. doi: 10.1007/s00300-023-03122-y.

Dudeja, S., Bhattacharjee, A.B. and Chela-Flores, J., 2010. Microbial mats in Antarctica as models for the search of life on the Jovian moon Europa. *Microbial Mats: Modern and ancient microorganisms in stratified systems*, pp.543-561. doi: 10.1007/978-90-481-3799-2_28.

Edgar, R.C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10(10), pp.996–998. doi: 10.1038/nmeth.2604.

Edgar, R.C. (2016). UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. [online] doi: 10.1101/081257.

Fernandez-Carazo, R., Hodgson, D.A., Convey, P. and Annick Wilmotte (2011). Low cyanobacterial diversity in biotopes of the Transantarctic Mountains and Shackleton Range (80-

82°S), Antarctica. *FEMS Microbiology Ecology*, [online] 77(3), pp.503–517. doi: 10.1111/j.1574-6941.2011.01132.x.

Fraser, C.I., Aleks Terauds, Smellie, J.L., Convey, P. and Chown, S.L. (2014). Geothermal activity helps life survive glacial cycles. *Proceedings of the National Academy of Sciences of the United States of America*, 111(15), pp.5634–5639. doi: 10.1073/pnas.1321437111.

Guidetti, R., Massa, E., Bertolani, R., Rebecchi, L. and Cesari, M. (2019). Increasing knowledge of Antarctic biodiversity: new endemic taxa of tardigrades (Eutardigrada; Ramazzottiidae) and their evolutionary relationships. *Systematics and Biodiversity*, 17(6), pp.573–593. doi: 10.1080/14772000.2019.1649737.

Hawes, I., Howard-Williams, C., Gilbert, N., Hughes, K. A., Convey, P., & Quesada, A. (2023). The need for increased protection of Antarctica's inland waters. *Antarctic Science*, 35(2), 64-88. doi: 10.1017/s0954102022000463.

Herbold, C. W., Lee, C. K., McDonald, I. R., and Cary, S. C. (2014). Evidence of global-scale aeolian dispersal and endemism in isolated geothermal microbial communities of Antarctica. *Nature communications*, 5(1), 3875. doi: 10.1038/ncomms4875

Hessen, D. O. (2007). Effects of UV radiation in arctic and alpine freshwater ecosystems. In *Arctic alpine ecosystems and people in a changing environment* (pp. 211-225). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-540-48514-8_13.

Hijmans, R.J. 2014. Geosphere: spherical trigonometry. R package version 1.3-11.<http://cran.r-project.org/web/packages/geosphere/index.html>.

Hughes, K.A., Cowan, D.A. and Wilmotte, A. (2015). Protection of Antarctic microbial communities – ‘out of sight, out of mind’. *Frontiers in Microbiology*, 6. doi: 10.3389/fmicb.2015.00151.

Hughes, K.A., Convey, P., Pertierra, L.R., Vega, G.C., Aragón, P. and Olalla-Tárraga, M.Á. (2019). Human-mediated dispersal of terrestrial species between Antarctic biogeographic regions: A preliminary risk assessment. *Journal of Environmental Management*, [online] 232, pp.73–89. doi: 10.1016/j.jenvman.2018.10.095.

Iakovenko, N., Smykla, J., Convey, P., Kašparová, E., Kozeretska, I.A., Trokhymets, V., Dykyy, I., Plewka, M., Devetter, M., Ďuriš, Z., and Janko, K. (2015). Antarctic bdelloid rotifers: diversity, endemism and evolution. *Hydrobiologia*, 761(1), pp.5–43. doi: 10.1007/s10750-015-2463-2.

Jackson, E. E., Hawes, I., and Jungblut, A. D. (2021). 16S rRNA gene and 18S rRNA gene diversity in microbial mat communities in meltwater ponds on the McMurdo Ice Shelf, Antarctica. *Polar Biology*, 44(4), 823-836. doi: 10.1007/s00300-021-02843-2

Jungblut, A. D., & Vincent, W. F. (2017). Cyanobacteria in polar and alpine ecosystems. *Psychrophiles: from biodiversity to biotechnology*, 181-206. doi: 10.1007/978-3-319-57057-0_9.

Jungblut, A.D., Lovejoy, C. and Vincent, W.F. (2010). Global distribution of cyanobacterial ecotypes in the cold biosphere. *The ISME Journal*, 4(2), pp.191–202. doi: 10.1038/ismej.2009.113.

Jungblut, A. D., Vincent, W. F., and Lovejoy, C. (2012). Eukaryotes in Arctic and Antarctic cyanobacterial mats. *FEMS Microbiology Ecology*, 82(2), 416-428. doi: 10.1111/j.1574-6941.2012.01418.x.

Kleinteich, J., Hildebrand, F., Bahram, M., Voigt, A.Y., Wood, S.A., Jungblut, A.D., Küpper, F.C., Quesada, A., Camacho, A., Pearce, D.A., Convey, P., Vincent, W.F., Zarfl, C., Bork, P. and Dietrich, D.R. (2017). Pole-to-Pole Connections: Similarities between Arctic and Antarctic Microbiomes and Their Vulnerability to Environmental Change. *Frontiers in Ecology and Evolution*, 5. doi: 10.3389/fevo.2017.00137.

Komárek, J., Nedbalová, L. and Hauer, T., 2012. Phylogenetic position and taxonomy of three heterocytous cyanobacteria dominating the littoral of deglaciated lakes, James Ross Island, Antarctica. *Polar Biology*, 35(5), pp.759-774. doi: 10.1007/s00300-011-1123-x.

Komárek J., Kaštovský J., Mareš J. & Johansen J. R. (2014) Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. – *Preslia* 86: 295–335

Komárek, J., Elster, J., and Komárek, O. (2008). Diversity of the cyanobacterial microflora of the northern part of James Ross Island, NW Weddell Sea, Antarctica. *Polar Biology*, 31, 853-865. doi: 10.1007/s00300-008-0424-1.

Komsta, L., and Novomestky, F. (2015). Moments, cumulants, skewness, kurtosis and related tests. *R package version*, 14(1).

Laybourn-Parry, J., Marchant, H. J., and Brown, P. (1991). The plankton of a large oligotrophic freshwater Antarctic lake. *Journal of Plankton Research*, 13(6), 1137-1149. doi: 10.1093/plankt/13.6.1137.

Lebre, P. H., Bosch, J., Coclet, C., Hallas, R., Hogg, I. D., Johnson, J., Moon, K. L., Ortiz, M., Rotimi, M. I., Stevens, M. I., Varliero, G., Convey, P., Vikram, S., Chown, S. L. and Cowan, D. A. (2023). Expanding Antarctic biogeography: microbial ecology of Antarctic island soils. *Ecography*, 2023(9), e06568. doi: 10.1111/ecog.06568.

Lefler, F.W., Berthold, D.E. and Laughinghouse IV, H.D. (2023). Cyanoseq: A database of cyanobacterial 16S rRNA gene sequences with curated taxonomy. *Journal of Phycology*, 59(3), pp.470-480. doi: 10.1111/jpy.13335.

Legendre, P. & Legendre, L. (2012). Multiscale analysis: spatial eigenfunctions. In *Developments in environmental modelling* (Vol. 24, pp. 859-906). Elsevier.. doi: 10.1016/B978-0-444-53868-0.50014-9.

Legendre, P., Borcard, D. and Peres-Neto, P.R. (2005). Analyzing beta diversity: partitioning the spatial variation of community composition data. *Ecological Monographs*, 75: 435-450. doi: 10.1890/05-0549

Lynch, M.D.J. and Neufeld, J.D. (2015). Ecology and exploration of the rare biosphere. *Nature Reviews Microbiology*, 13(4), pp.217–229. doi: 10.1038/nrmicro3400.

Marra, G. and Wood, S.N. (2011). Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis*, 55(7), pp.2372–2387. doi: 10.1016/j.csda.2011.02.004.

Martinez Arbizu, P. (2020). pairwiseAdonis: Pairwise multilevel comparison using adonis. R package version 0.4.

McArdle, B.H. and Anderson, M.J. (2001). Fitting multivariate models to community data: a comment on distance-based redundancy analysis. *Ecology*, 82(1), pp.290–297. doi: 10.1890/0012-9658(2001)082[0290:fmmtd]2.0.co;2.

McGaughan, A., Stevens, M.I. and Holland, B.R. (2010). Biogeography of circum-Antarctic springtails. *Molecular Phylogenetics and Evolution*, 57(1), pp.48–58. doi: 10.1016/j.ympev.2010.06.003.

Muñoz, J., Felicísimo, A.M., Cabezas, F., Burgaz, A.R. and Martínez, I. (2004). Wind as a long-distance dispersal vehicle in the Southern Hemisphere. *Science*, 304(5674), pp.1144-1147. doi: 10.1126/science.1095210

Nabout, J.C., Siqueira, T., Bini, L.M. and Nogueira, I. de S. (2009). No evidence for environmental and spatial processes in structuring phytoplankton communities. *Acta Oecologica*, 35(5), pp.720–726. doi: 10.1016/j.actao.2009.07.002.

Nemergut, D.R., Costello, E.K., Hamady, M., Lozupone, C., Jiang, L., Schmidt, S.K., Fierer, N., Townsend, A.R., Cleveland, C.C., Stanish, L. and Knight, R., 2011. Global patterns in the biogeography of bacterial taxa. *Environmental microbiology*, 13(1), pp.135-144. doi: 10.1111/j.1462-2920.2010.02315.x.

Nübel, U., Garcia-Pichel, F. and Muyzer, G. (1997). PCR primers to amplify 16S rRNA genes from cyanobacteria. *Applied and environmental microbiology*, 63(8), pp.3327–3332. doi: 10.1128/aem.63.8.3327-3332.1997.

Oksanen, J., Blanchet, F.G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Stevens, M.H. H., Szoecs, E., and Wagner, H. (2019) vegan: Community ecology package.

Peat, H.J., Clarke, A. and Convey, P. (2006). ORIGINAL ARTICLE: Diversity and biogeography of the Antarctic flora. *Journal of Biogeography*, 34(1), pp.132–146. doi: 10.1111/j.1365-2699.2006.01565.x.

- Peres-Neto, P. R., Legendre, P., Dray, S., and Borcard, D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology* 87, 2614–2625. doi: 10.1890/0012-9658(2006)87[2614:vposdm]2.0.co;2.
- Pertierra, L.R., Hughes, K.A., Vega, G.C. and Olalla-Tárraga, M.Á. (2017). High Resolution Spatial Mapping of Human Footprint across Antarctica and Its Implications for the Strategic Conservation of Avifauna. *PLOS ONE*, 12(1), p.e0168280. doi: 10.1371/journal.pone.0168280.
- Pessi, I.S., Maalouf, P.D.C., Laughinghouse, H.D., IV, Baurain, D. and Wilmotte, A. (2016), On the use of high-throughput sequencing for the study of cyanobacterial diversity in Antarctic aquatic mats. *J. Phycol.*, 52: 356-368. doi: 10.1111/jpy.12399.
- Pessi, I. S., Lara, Y., Durieu, B., Maalouf, P. D. C., Verleyen, E., and Wilmotte, A. (2018). Community structure and distribution of benthic cyanobacteria in Antarctic lacustrine microbial mats. *FEMS microbiology ecology*, 94(5), fiy042. doi: 10.1093/femsec/fiy042.
- Pessi, I.S., Popin, R. V., Durieu, B., Lara, Y., Savaglia, V., Roncero-Ramos, B., Hultman, J., Verleyen, E., Vyverman, W. and Wilmotte, A. (2023). Novel diversity of polar Cyanobacteria revealed by genome-resolved metagenomics. doi: 10.1099/mgen.0.001056.
- Pinseel, E., Janssens, S.B., Verleyen, E. Vanormelingen, P., Kohler, T.J., Biersma, E.M., Sabbe, K., Van de Vijver, B. and Vyverman, W. (2020). Global radiation in a rare biosphere soil diatom. *Nat Commun* 11, 2382. doi: 10.1038/s41467-020-16181-0.
- Pugh, P. and Convey, P. (2008). Surviving out in the cold: Antarctic endemic invertebrates and their refugia. *Journal of Biogeography*, 35(12), pp.2176–2186. doi: 10.1111/j.1365-2699.2008.01953.x.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: <https://www.R-project.org/>
- Sabbe, K., Hodgson, D.A., Elie Verleyen, Taton, A., Annick Wilmotte, Vanhoutte, K. and Wim Vyverman (2004). Salinity, depth and the structure and composition of microbial mats in continental Antarctic lakes. *Freshwater Biology*, 49(3), pp.296–319. doi: 10.1111/j.1365-2427.2004.01186.x.
- Sakaeva, A., Sokol, E.R., Kohler, T.J., Stanish, L.F., Spaulding, S.A., Howkins, A., Welch, K.A., Lyons, W.B., Barrett, J.E. and McKnight, D.M. (2016). Evidence for dispersal and habitat controls on pond diatom communities from the McMurdo Sound Region of Antarctica. *Polar Biology*, 39(12), pp.2441–2456. doi: 10.1007/s00300-016-1901-6.
- Sakamoto, T., Yoshida, T., Arima, H., Hatanaka, Y., Takani, Y. and Tamaru, Y. (2009). Accumulation of trehalose in response to desiccation and salt stress in the terrestrial cyanobacterium *Nostoc commune*. *Phycological Research*, 57(1), pp.66–73. doi: 10.1111/j.1440-1835.2008.00522.x.

Scientific Committee on Antarctic Research [SCAR]. (2009). "SCAR's environmental code of conduct for terrestrial scientific field research in Antarctica (IP04)," in Proceedings of the 32th Antarctic Treaty Consultative Meeting, Baltimore. Available at: http://www.ats.aq/documents/ATCM32/ip/ATCM32_ip004_e.doc

Scientific Committee on Antarctic Research [SCAR]. (2011). "SCAR's code of conduct for the exploration and research of subglacial aquatic environments (IP33)," in Proceedings of the 34th Antarctic Treaty Consultative Meeting, Buenos Aires. Available at: http://www.ats.aq/documents/ATCM34/ip/ATCM34_ip033_e.doc

Schulte, N.O., Khan, A.L., Smith, E.W., Zoumplis, A., Kaul, D., Allen, A.E., Adams, B.J. and McKnight, D.M. (2022). Blowin' in the wind: Dispersal, structure, and metacommunity dynamics of aeolian diatoms in the McMurdo Sound region, Antarctica. *Journal of Phycology*, 58(1), pp.36–54. doi: 10.1111/jpy.13223.

Smith, V. R., & Ashton, P. J. (1971). Bryophyte-cyanobacteria associations on sub-Antarctic Marion Island: are they important in nitrogen fixation?. Vol 4. Available at: <http://alp.lib.sun.ac.za/handle/123456789/7411>.

Sokol, E.R., Herbold, C.W., Lee, C.K., Cary, S.C. and Barrett, J.E. (2013). Local and regional influences over soil microbial metacommunities in the Transantarctic Mountains. *Ecosphere*, [online] 4(11), p.art136. doi: 10.1890/es13-00136.1.

Souffreau, C., Vanormelingen, P., Verleyen, E., Sabbe, K. and Vyverman, W. (2010). Tolerance of benthic diatoms from temperate aquatic and terrestrial habitats to experimental desiccation and temperature stress. *Phycologia*, 49(4), pp.309–324. doi: 10.2216/09-30.1.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), pp.1312–1313. doi: 10.1093/bioinformatics/btu033.

Sumner, D.Y., Jungblut, A.D., Hawes, I., Andersen, D.T., Mackey, T.J. and Wall, K. (2016). Growth of elaborate microbial pinnacles in Lake Vanda, Antarctica. 14(6), pp.556–574. doi: 10.1111/gbi.12188.

Tanabe, Y. and Kudoh, S. (2012). Possible ecological implications of floating microbial assemblages lifted from the lakebed on an Antarctic lake. *Ecological Research*, 27(2), pp.359–367. doi: 10.1007/s11284-011-0907-3.

Taton, A., Grubisic, S., Brambilla, E., De Wit, R. and Wilmotte, A. (2003). Cyanobacterial Diversity in Natural and Artificial Microbial Mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a Morphological and Molecular Approach. *Applied and Environmental Microbiology*, 69(9), pp.5157–5169. doi: 10.1128/aem.69.9.5157-5169.2003.

Taton, A., Grubisic, S., Balthasart, P., Hodgson, D.A., Laybourn-Parry, J. and Wilmotte, A. (2006). Biogeographical distribution and ecological ranges of benthic cyanobacteria in East Antarctic lakes. *FEMS Microbiology Ecology*, 57(2), pp.272–289. doi: 10.1111/j.1574-6941.2006.00110.x.

- Terauds, A. and Lee, J.R. (2016). Antarctic biogeography revisited: updating the Antarctic Conservation Biogeographic Regions. *Diversity and Distributions*, 22(8), pp.836–840. doi: 10.1111/ddi.12453.
- Terauds, A., Chown, S.L., Morgan, F., J. Peat, H., Watts, D.J., Keys, H., Convey, P. and Bergstrom, D.M. (2012). Conservation biogeography of the Antarctic. *Diversity and Distributions*, 18(7), pp.726–741. doi: 10.1111/j.1472-4642.2012.00925.x.
- Tytgat, B., Verleyen, E., Sweetlove, M., Den, V., Pinseel, E., Hodgson, D.A., Chown, S.L., Sabbe, K., Wilmotte, A., Willems, A. and Vyverman, W. (2023). Polar lake microbiomes have distinct evolutionary histories. *Science Advances*, 9(46). doi: 10.1126/sciadv.ade7130.
- Varliero, G., Lebre, P.H., Adams, B., Chown, S.L., Convey, P., Dennis, P.G., Fan, D., Ferrari, B., Frey, B., Hogg, I.D., Hopkins, D.W., Kong, W., Makhalanyane, T., Matcher, G., Newsham, K.K., Stevens, M.I., Weigh, K.V. and Cowan, D.A. (2024). Biogeographic survey of soil bacterial communities across Antarctica. *Microbiome*, 12(1). doi: 10.1186/s40168-023-01719-3.
- Velázquez, D., Jungblut, A.D., Rochera, C., Rico, E., Camacho, A. and Quesada, A. (2017). Trophic interactions in microbial mats on Byers Peninsula, maritime Antarctica. *Polar Biology*, 40(5), pp.1115–1126. doi: 10.1007/s00300-016-2039-2.
- Verleyen, E., Hodgson, D.A., Gibson, J., Imura, S., Kaup, E., Kudoh, S., De Wever, A., Hoshino, T., McMinn, A., Obbels, D. and Roberts, D. (2012). Chemical limnology in coastal East Antarctic lakes: monitoring future climate change in centres of endemism and biodiversity. *Antarctic Science*, 24(1), pp.23-33. doi: 10.1017/S0954102011000642.
- Verleyen, E., Van de Vijver, B., Tytgat, B., Pinseel, E., Hodgson, D.A., Kopalová, K., Chown, S.L., Van Ranst, E., Imura, S., Kudoh, S., Van Nieuwenhuyze, W., Sabbe, K. and Vyverman, W. (2021). Diatoms define a novel freshwater biogeography of the Antarctic. *Ecography*, 44(4), pp.548–560. doi: 10.1111/ecog.05374.
- Vincent, W. F. (2000a). Evolutionary origins of Antarctic microbiota: invasion, selection and endemism. *Antarctic Science*, 12(3), 374-385.
- Vincent, W.F. (2000b). Cyanobacterial dominance in the Polar Regions. *The Ecology of Cyanobacteria*, pp.321–340. doi: 10.1007/0-306-46855-7_12.
- Vincent, W. F., & Laybourn-Parry, J. (Eds.). (2008). *Polar lakes and rivers: limnology of Arctic and Antarctic aquatic ecosystems*. Oxford university press.
- Vincent, W.F. and Quesada, A. (2012). Cyanobacteria in High Latitude Lakes, Rivers and Seas. *Springer eBooks*, pp.371–385. doi: 10.1007/978-94-007-3855-3_13.
- Wang, Q., Garrity, G.M., Tiedje, J.M. and Cole, J.R. (2007). Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Applied and Environmental Microbiology*, [online] 73(16), pp.5261–5267. doi: 10.1128/aem.00062-07.

Wauchope, H.S., Shaw, J.D. and Terauds, A. (2019). A snapshot of biodiversity protection in Antarctica. *Nature Communications*, 10(1). doi: 10.1038/s41467-019-08915-6.

Wharton, R.A., Parker, B.C. and Simmons, G.M. (1983). Distribution, species composition and morphology of algal mats in Antarctic dry valley lakes. *Phycologia*, 22(4), pp.355–365. doi: 10.2216/i0031-8884-22-4-355.1.

Wood, S.N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1), pp.95–114. doi: 10.1111/1467-9868.00374.

Wood, S.N., N. Pya and B. Saefken (2016). Smoothing parameter and model selection for general smooth models. *Journal of the American Statistical Association* 111, 1548-1575. doi: 10.1080/01621459.2016.1180986

Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.-H., Whitman, W.B., Euzéby, J., Amann, R. and Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, 12(9), pp.635–645. doi: 10.1038/nrmicro3330.

Zawierucha, K., Marshall, C.J., Wharton, D. and Janko, K. (2019). A nematode in the mist: *Scottnema lindsayae* is the only soil metazoan in remote Antarctic deserts, at greater densities with altitude. *Polar Research*. [online] doi: 10.33265/polar.v38.3494.

2.8. Supplementary information

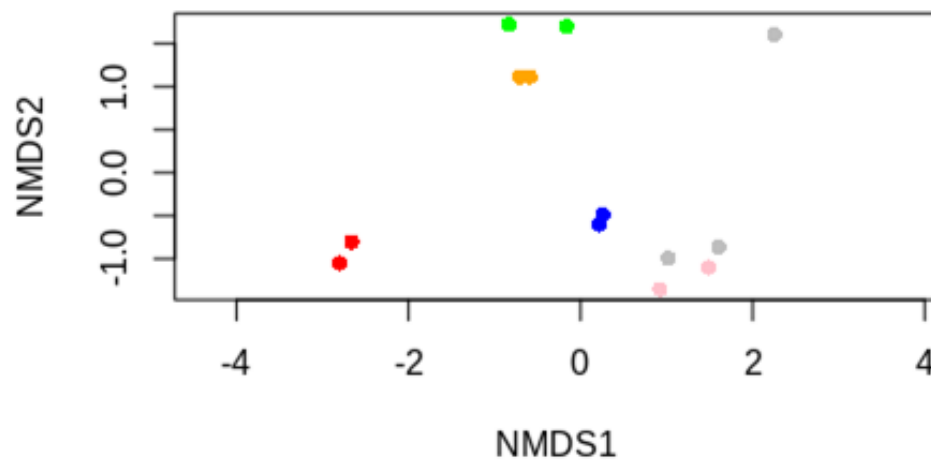


Fig. S2.1. metaMDS on the Bray-Curtis similarity matrix of the 5 replicates (green, red, blue, pink and orange) based on Hellinger transformed abundance data table. Three samples were added to test the robustness of the metaMDS analysis (in grey).



Fig. S2.2. (Part1) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

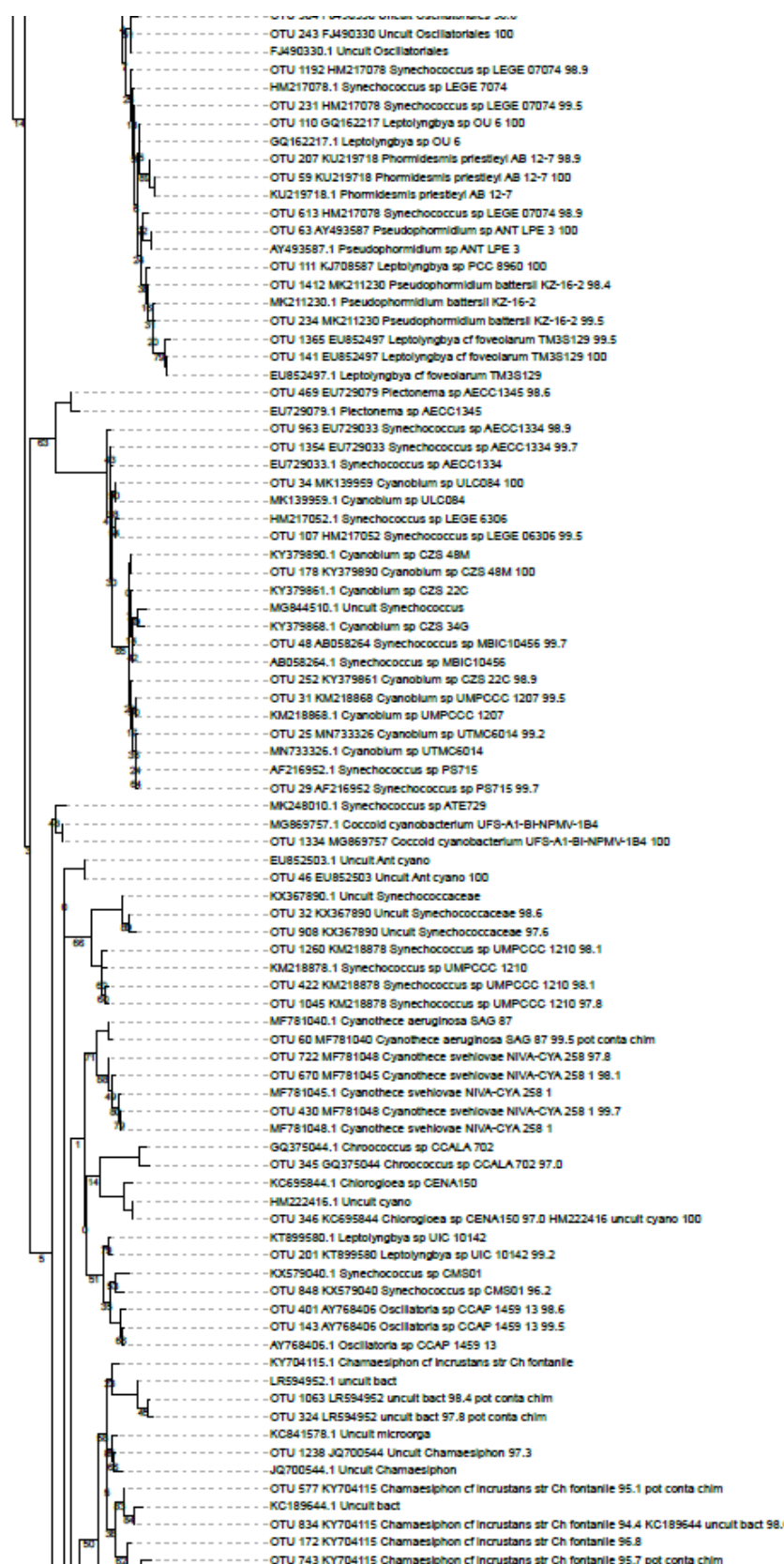


Fig. S2.2. (Part2) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

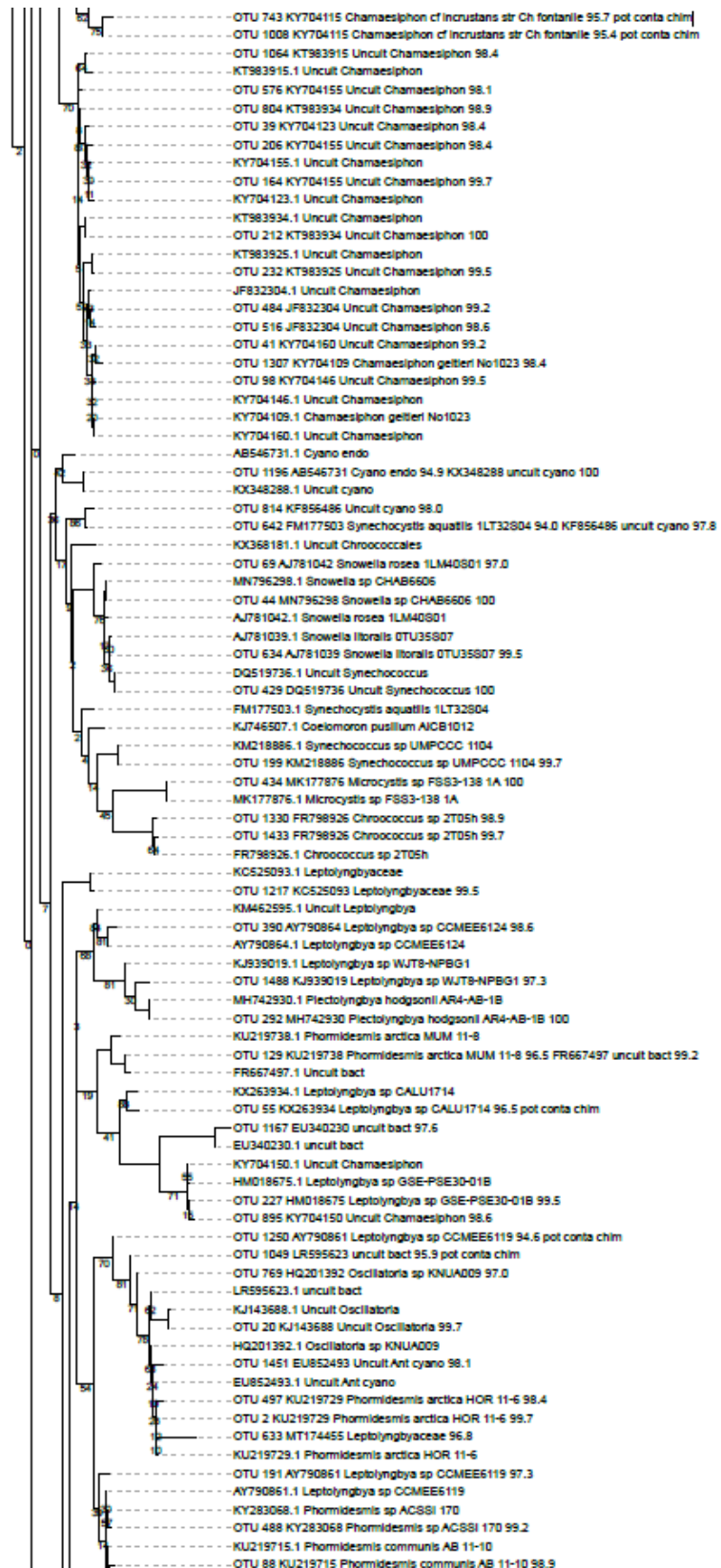


Fig. S2.2. (Part3) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

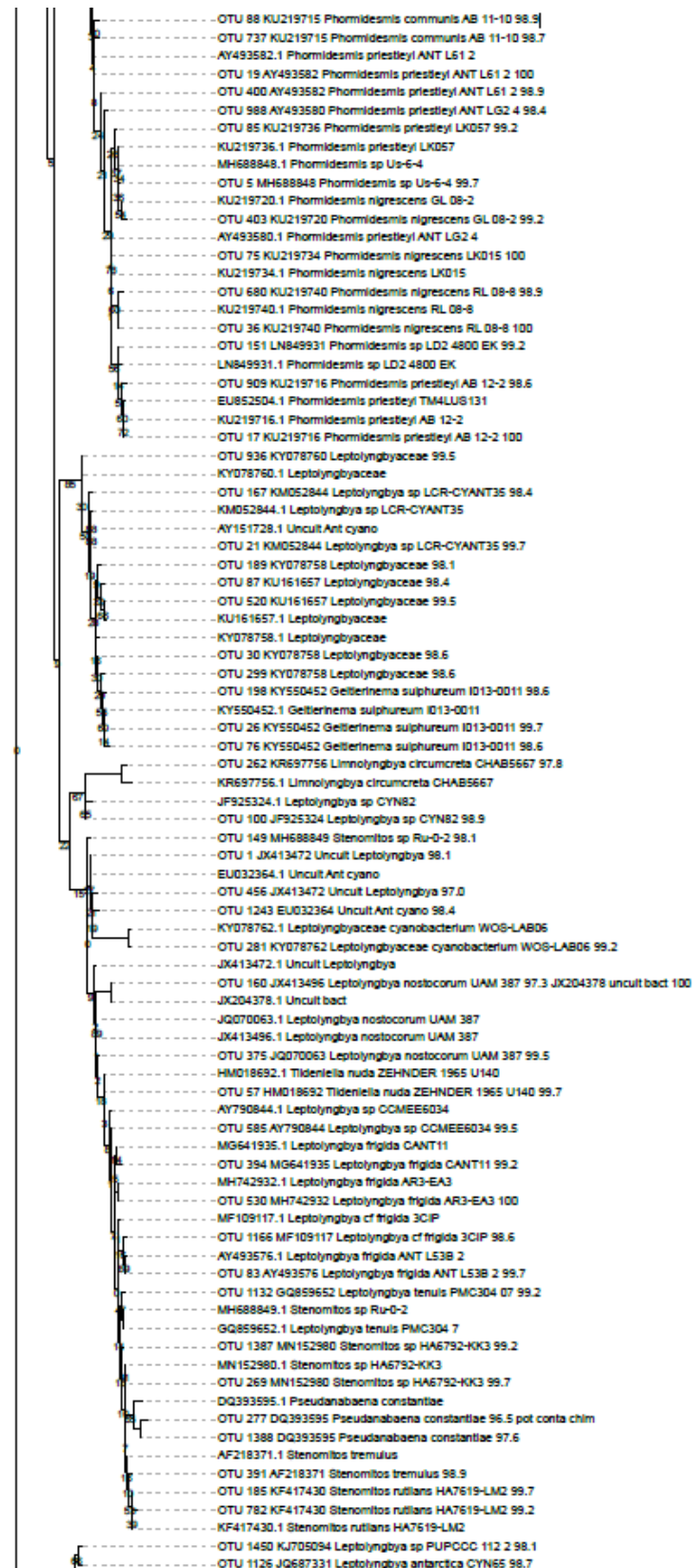


Fig. S2.2. (Part4) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

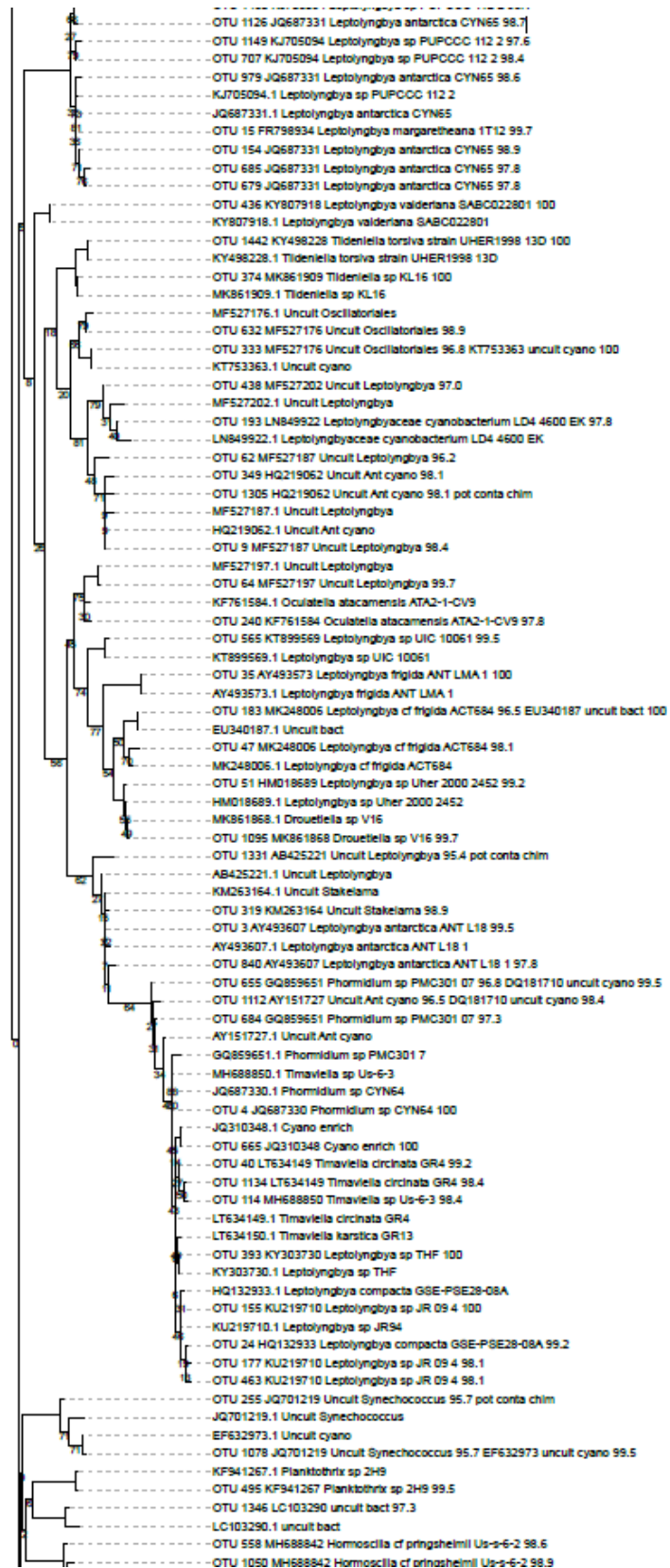


Fig. S2.2. (Part5) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

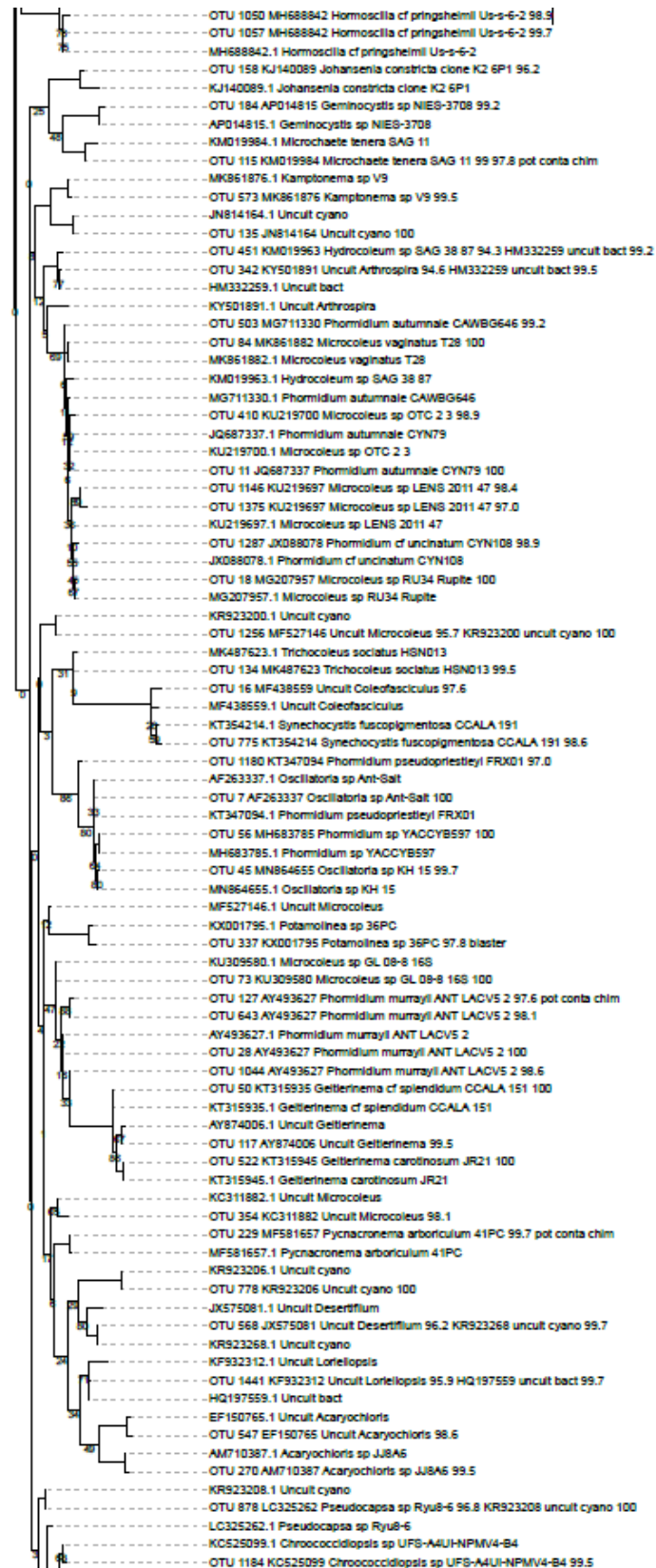


Fig. S2.2. (Part6) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

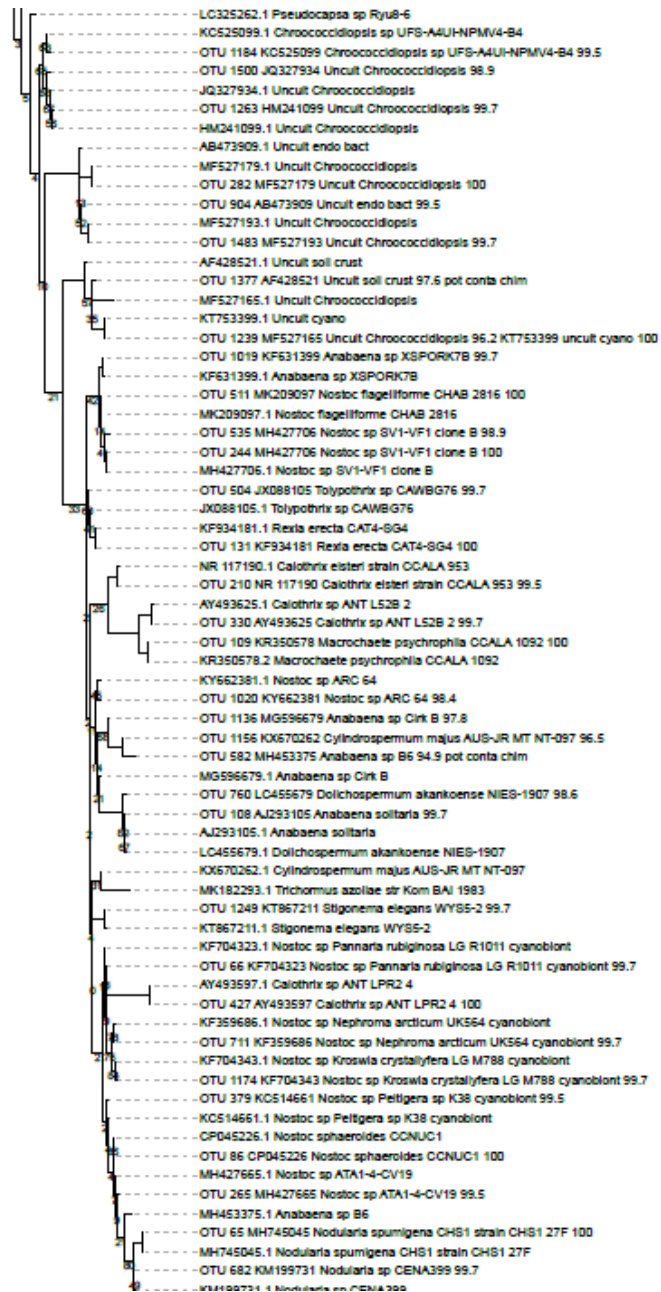


Fig. S2.2. (Part7) Phylogenetic analysis of the OTUs sequences, their close relative taken from the GenBank nucleotide (nt) collection (accessed on 9 April 2024) and the CyanoSeq database (Forrest et al. 2023). Maximum likelihood tree computed using GTR+GAMMA model and 100 bootstraps.

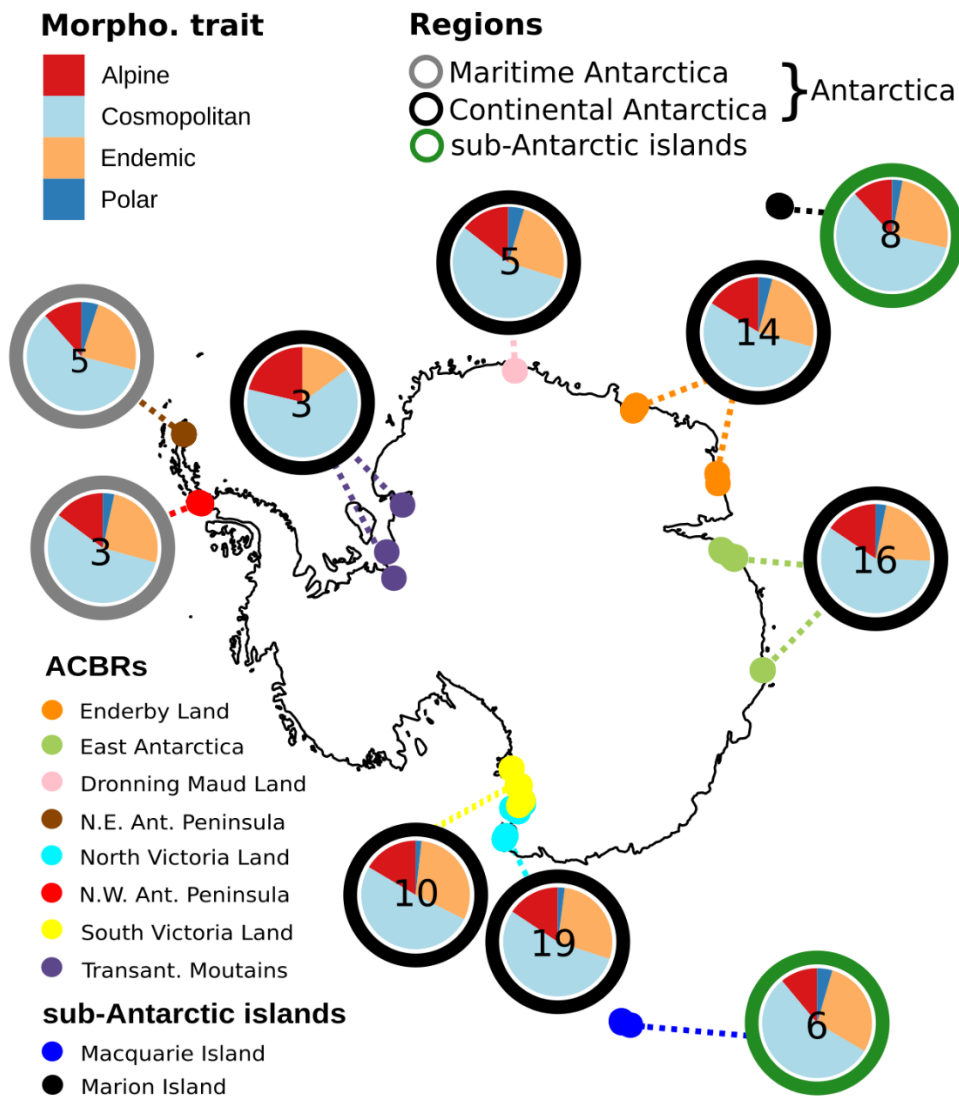


Fig. S2.3. Map showing the sampling locations. The pie charts present the geographic distribution of OTUs ('endemic', 'alpine', 'polar', 'cosmopolitan') in each ACBRs and in sub-Antarctic islands. The matrix of presence/absence of OTUs for each sample was used. The numbers at the middle of the pie charts represent the numbers of investigated samples for each region.

Table S2.1. Environmental data for the 89 samples.

Sample	Lake	Bioregion	ACBR	ACBR_code	Sample_year	North_deg	East_deg	Conductivity (mS cm ⁻¹) surface	pH surface
DRMAU_SC1_litt	S1 (=SC1)	continental Antarctica	Dronning Maud Land	DRMAU	2003	-70,774	11,813	0,006	8,51
DRMAU_SC2_litt	S2 (=SC2)	continental Antarctica	Dronning Maud Land	DRMAU	2003	-70,767	11,746	0,002	6,79
DRMAU_SC5_litt	S5 (=SC5)	continental Antarctica	Dronning Maud Land	DRMAU	2003	-70,766	11,765	0,05	7,28
DRMAU_SC6_litt	S6 (=SC6)	continental Antarctica	Dronning Maud Land	DRMAU	2003	-70,768	11,792	0,009	7,5
DRMAU_SC7_litt	S7 (=SC7)	continental Antarctica	Dronning Maud Land	DRMAU	2003	-70,771	11,803	0,154	7,1
EDRBY_EO7	EO7	continental Antarctica	Enderby Land	EDRBY	2003	-69,015	39,579	0,234	7,6
EDRBY_L_Lorna_H2	Lorna (mat sample)	continental Antarctica	Enderby Land	EDRBY	2001	-67,785	62,788	0,436	8,94
EDRBY_LA7_litt_0.1A	LA7	continental Antarctica	Enderby Land	EDRBY	2003	-69,193	39,635	5,046	8,09
EDRBY_Lake_Barkell	Lake Barkell	continental Antarctica	Enderby Land	EDRBY	2001	-67,466	60,967	0,61	8,49
EDRBY_SK1_G0_1_Mago_Ike_L	SK1 - Mago Ike lake	continental Antarctica	Enderby Land	EDRBY	2001	-69,472	39,631	0,484	8,95
EDRBY_SK5	SK5 - Naka-Tempyo Ike lake	continental Antarctica	Enderby Land	EDRBY	2007	-69,504	39,714	0,014	8,58
EDRBY_SK5_surface	SK5 - Naka-Tempyo Ike lake	continental Antarctica	Enderby Land	EDRBY	2007	-69,504	39,714	0,014	8,58
EDRBY_Ufs_Isl_ID2_6_Hemsworth_I	Ufs Island Lake (edge mat sample)	continental Antarctica	Enderby Land	EDRBY	2001	-67,481	61,126	NA	NA
EDRBY_WO1_litt	WO1	continental Antarctica	Enderby Land	EDRBY	2003	-69,504	39,714	0,249	8,1
EDRBY_WO4_litt_0.1m	WO4	continental Antarctica	Enderby Land	EDRBY	2003	-69,021	39,571	0,33	7,63
EDRBY_WO6_litt_0.1m	WO6	continental Antarctica	Enderby Land	EDRBY	2003	-69,023	39,576	0,668	7,64

EDRBY_WO6_litt_1.5m_east_side	WO6	continental Antarctica	Enderby Land	EDRBY	2003	-69,023	39,576	0,668	7,64
EDRBY_WO8_litt	WO8	continental Antarctica	Enderby Land	EDRBY	2003	-69,029	39,523	0,273	7,76
EDRBY_WO8_litt_epips_0.2m	WO8	continental Antarctica	Enderby Land	EDRBY	2003	-69,029	39,523	0,273	7,76
ESANT_Ambio_14	Fold Lake	continental Antarctica	East Antarctica	ESANT	NA	-69,350	76,333	0,885	7,22
ESANT_Ambio_19	Lake Burgess	continental Antarctica	East Antarctica	ESANT	NA	-69,415	76,128	0,186	6,53
ESANT_Ambio_24	Ace_L	continental Antarctica	East Antarctica	ESANT	NA	-68,473	78,185	13,73	NA
ESANT_Corner_L_litt	Corner Lake 1	continental Antarctica	East Antarctica	ESANT	1997	-68,600	78,533	0,0376	6,73
ESANT_Highway_VH26	Highway Lake	continental Antarctica	East Antarctica	ESANT	1996	-68,465	78,188	4,5	NA
ESANT_L_Alanna	Lake Alanna	continental Antarctica	East Antarctica	ESANT	1997	-69,467	75,917	2,54	6,85
ESANT_L_Jack	Lake Jack	continental Antarctica	East Antarctica	ESANT	1997	-69,417	76,100	0,1113333333	6,79
ESANT_M15_Bunger_Hills	M15	continental Antarctica	East Antarctica	ESANT	2002	-66,240	100,708	NA	NA
ESANT_M16_Bunger_Hills	M16	continental Antarctica	East Antarctica	ESANT	2002	-66,251	100,605	NA	NA
ESANT_M5_Bunger_Hills	M5	continental Antarctica	East Antarctica	ESANT	2002	-66,296	100,691	NA	NA
ESANT_Organic_L	Organic Lake	continental Antarctica	East Antarctica	ESANT	1995	-68,458	78,192	74,12	NA
ESANT_Pendant_L	Pendant Lake	continental Antarctica	East Antarctica	ESANT	1995	-68,462	78,242	13,17	NA
ESANT_Rauer2_RI2	Rauer 2	continental Antarctica	East Antarctica	ESANT	1993	-68,817	77,700	131,500	7,52
ESANT_Reid_litt	Lake Reid	continental Antarctica	East Antarctica	ESANT	1993	-69,383	76,383	7,38	7,06
ESANT_Waterfall_L	Waterfall Lake	continental Antarctica	East Antarctica	ESANT	2004	-68,549	78,336	NA	NA
ESANT_Watts_L	Watts Lake	continental Antarctica	East Antarctica	ESANT	1995	-68,600	78,183	2,55	NA

ESPEN_L17FR	L17	Antarctica maritime	North East Antarctic Peninsula	ESPEN	2013	-63,828	-57,322	0,088	9,1
ESPEN_L18FR	L18	Antarctica maritime	North East Antarctic Peninsula	ESPEN	2013	-63,826	-57,317	0,035	9,1
ESPEN_L19FR	L19	Antarctica maritime	North East Antarctic Peninsula	ESPEN	2013	-63,820	-57,316	0,148	9,4
ESPEN_L30FR	L30	Antarctica maritime	North East Antarctic Peninsula	ESPEN	2013	-63,819	-57,310	0,353	10,2
ESPEN_L32FR	L32	Antarctica maritime	North East Antarctic Peninsula	ESPEN	2013	-63,817	-57,332	3,37	9,9
INDIA_MI02	MI2	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,864	37,858	1,848	7,29
INDIA_MI16	MI16	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,888	37,811	0,43	6,46
INDIA_MI20	MI20	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,892	37,866	0,024	5,96
INDIA_MI3	MI3	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,864	37,856	0,15	7,85
INDIA_MI4	MI4	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,861	37,847	0,061	7,85
INDIA_MI44	MI44	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,943	37,590	0,292	6,32
INDIA_MI52	MI52	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,929	37,588	1,875	6,35
INDIA_MI60	MI60	sub-Antarctic islands	Subantarctic Indian ocean	INDIA	2007	-46,885	37,615	0,895	7,28
NVICT_06_12	Lake 15	continental Antarctica	North Victoria Land	NVICT	2005	-74,326	165,126	0,483	8,78
NVICT_06_51	Lake 7	continental Antarctica	North Victoria Land	NVICT	2005	-74,723	162,694	0,492	8,48
NVICT_147_06	Skua Lake	continental Antarctica	North Victoria Land	NVICT	2005	-74,700	164,083	7,52	9,24
NVICT_153_06	Carezza Lake	continental Antarctica	North Victoria Land	NVICT	2005	-74,710	164,058	0,245	7,33
NVICT_161_06	Ponds	continental Antarctica	North Victoria Land	NVICT	2005	-74,733	163,950	0,263	8,62

NVICT_161_XX	Enigma Lake	continental Antarctica	North Victoria Land	NVICT	2004- 2005	-74,725	164,033	0,216	9,12
NVICT_170_06	Lake 3	continental Antarctica	North Victoria Land	NVICT	2005	-74,361	165,131	1,016	9,79
NVICT_178_06	Lake 15 A	continental Antarctica	North Victoria Land	NVICT	2005	-74,313	165,069	0,049	7,8
NVICT_183_06	Lake	continental Antarctica	North Victoria Land	NVICT	2005	-74,233	164,750	0,154	8,53
NVICT_21_06	Lake Gondwana (alto)	continental Antarctica	North Victoria Land	NVICT	2005	-74,613	165,214	0,251	8,59
NVICT_211_XIX	Lake D (4 - vicino mare)	continental Antarctica	North Victoria Land	NVICT	2003	-74,882	163,717	0,533	8,2
NVICT_230_06	Lake North	continental Antarctica	North Victoria Land	NVICT	2005	-72,346	169,924	0,123	8,37
NVICT_240_06	Lake	continental Antarctica	North Victoria Land	NVICT	2005	-72,371	169,888	0,392	8,73
NVICT_251_XX	Lake	continental Antarctica	North Victoria Land	NVICT	2004	-72,600	169,350	0,013	7,4
NVICT_26_06	Lake Gondwana Base Lake	continental Antarctica	North Victoria Land	NVICT	2005	-74,633	166,227	1,124	9,52
NVICT_36_06	Laguna salmastra	continental Antarctica	North Victoria Land	NVICT	2005	-74,894	163,683	60	8,05
NVICT_37_XX	Lake	continental Antarctica	North Victoria Land	NVICT	2004	-74,776	163,654	0,105	7,58
NVICT_6_06	Lake 14	continental Antarctica	North Victoria Land	NVICT	2005	-74,329	165,133	1,803	9,71
NVICT_70_06	Lake 20 A	continental Antarctica	North Victoria Land	NVICT	2005	-74,968	162,516	0,909	8,1
PACIF_MAC10	MAC10	sub-Antarctic islands	Southwestern Pacific ocean	PACIF	2013	-54,684	158,869	0,169	8,59
PACIF_MAC2	MAC2	sub-Antarctic islands	Southwestern Pacific ocean	PACIF	2013	-54,497	158,901	0,393	8,08
PACIF_MAC4	MAC4	sub-Antarctic islands	Southwestern Pacific ocean	PACIF	2013	-54,521	158.8950	0,17	8,26
PACIF_MAC6	MAC6	sub-Antarctic islands	Southwestern Pacific ocean	PACIF	2013	-54,656	158,876	0,134	8,37
PACIF_Square_L	Square Lake	sub-Antarctic	Southwestern Pacific	PACIF	2013	-54,570	158,900	0,202	9,94

PACIF_Tulloch_Lake	Tulloch Lake	islands sub-Antarctic islands	ocean Southwestern Pacific ocean	PACIF	2013	-54,579	158,895	0,072	7,25
SVICT_06_265	Pond	continental Antarctica	South Victoria Land	SVICT	2005	-75,899	162,579	0,203	7,57
SVICT_227_XX	Lake E (5)	continental Antarctica	South Victoria Land	SVICT	2004	-76,162	161,997	0,252	7,96
SVICT_254_06	Lake G	continental Antarctica	South Victoria Land	SVICT	2005	-76,163	162,000	0,125	9,17
SVICT_278_06	Pond	continental Antarctica	South Victoria Land	SVICT	2005	-76,244	162,790	1,293	9
SVICT_282_06	Oasi Lake	continental Antarctica	South Victoria Land	SVICT	2005	-74,697	164,099	6,74	9,62
SVICT_69_06	Lake 20	continental Antarctica	South Victoria Land	SVICT	2005	-74,968	162,517	0,34	8,64
SVICT_AMB37	Hoare_L_SS	continental Antarctica	South Victoria Land	SVICT	NA	-77,633	162,850	0,567	NA
SVICT_FRY2	Lake Fryxell	continental Antarctica	South Victoria Land	SVICT	NA	-77,600	163,117	NA	NA
SVICT_L_Fryxell	Lake Fryxell	continental Antarctica	South Victoria Land	SVICT	1999	-77,600	163,117	NA	NA
SVICT_L_Hoare	Lake Hoare	continental Antarctica	South Victoria Land	SVICT	1995	-77,633	162,850	NA	NA
TRMNT_Forlidas_TM2	Forlidas Pond	continental Antarctica	Transantantarctic Mountains	TRMNT	1999	-82,450	-51,350	2,22	8,15
TRMNT_Lundstrom_TM4	Lundström Lake	continental Antarctica	Transantantarctic Mountains	TRMNT	1999	-80,450	-29,480	0,227	9,04
TRMNT_Mount_Lowry	Lake Moat	continental Antarctica	Transantantarctic Mountains	TRMNT	2009	-84,000	-63,000	NA	NA
WSPEN_Col_L	Col Lake 1	maritime Antarctica	North West Antarctic Peninsula	WSPEN	NA	-67,831	-67,232	0,1132	NA
WSPEN_PP_mat	Narrows Lake	maritime Antarctica	North West Antarctic Peninsula	WSPEN	NA	-67,601	-67,207	0,0927	NA
WSPEN_PP_Shallow	Narrows Lake	maritime Antarctica	North West Antarctic Peninsula	WSPEN	NA	-67,601	-67,207	0,0927	NA

Table S2.2 Metadata of the 310 OTUs.

OTUs	blast_ntmatch_OTUs_Closed_	Genus	Family	Order	traitMorphologic_	_distributionGeographic	Abund_total	distributionSampling_	Abund_Ant	Freq_Ant	Masse_Ant	Abund_Sub	Freq_Sub	Masse_Sub
OTU_2	OTU_2_KU219729_Phormidesmis_arctica_HOR_11-6_99,7	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	243523	Both	242304	14,7563	Abundant	1219	0,7563	Abundant
OTU_3	OTU_3_JN979956_Leptolyngbya_borchgrevinkii_KOVACIK-ANT_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	153111	Both	153078	9,3225	Abundant	33	0,0205	Rare
OTU_1	OTU_1_ON803467_Stenomitros_frigidus_BEA_1253B_99,2	Stenomitros	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	126469	Both	126439	7,7001	Abundant	30	0,0186	Rare
OTU_5	OTU_5_MJF810642_Leptolyngbya_sp_M1.3_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	117661	Both	117644	7,1645	Abundant	17	0,0105	Rare
OTU_4	OTU_4_MH688850_Timaviella_dunensis_Us-6-3_100	Timaviella	Oscillatoriaceae	Synechococcales	thin_filamentous	Cosmopolitan	84655	Both	84524	5,1475	Abundant	131	0,0813	Average
OTU_6	OTU_6_AY493588_Leptolyngbya_antarctica_ANT_LAC1_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	89558	Both	89524	5,4520	Abundant	34	0,0211	Rare
OTU_11	OTU_11_JQ687337_Phormidium_autumnale_CYN79_100	Phormidium	Oscillatoriales	Oscillatoriales	large_filamentous	Cosmopolitan	70378	Both	65121	3,9659	Abundant	5257	3,2615	Abundant
OTU_7	OTU_7_AF263337_Oscillatoria_sp_Ant-Salt_100	Oscillatoria	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	62786	Both	62782	3,8234	Abundant	4	0,0025	Rare
OTU_19	OTU_19_AY493582_Phormidesmis_priestleyi_ANT_L61_2_100	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	56910	Both	21155	1,2883	Abundant	35755	22,1830	Abundant
OTU_9	OTU_9_MF527187_Uncult_Leptolyngbya_98,4	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	44998	Both	42393	2,5817	Abundant	2605	1,6162	Abundant
OTU_18	OTU_18_MG207957_Microcoleus_sp_RU34_Rupite_100	Microcoleus	Microcoleaceae	Oscillatoriales	large_filamentous	Cosmopolitan	43265	Both	42805	2,6068	Abundant	460	0,2854	Average
OTU_16	OTU_16_MF438559_Uncult_Coleofasciculus_97,6	Coleofasciculus	Coleofasciculaceae	Oscillatoriales	large_filamentous	Endemic	38851	Both	38849	2,3659	Abundant	2	0,0012	Rare
OTU_17	OTU_17_KU219716_Phormidesmis_priestleyi_AB_12-2_100	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	38772	Both	36161	2,2022	Abundant	2611	1,6199	Abundant
OTU_21	OTU_21_KM052844_Leptolyngbya_sp_LCR-CYANT35_99,7	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	38356	Both	38351	2,3356	Abundant	5	0,0031	Rare
OTU_15	OTU_15_JQ687331_Leptolyngbya_antarctica_CYN65_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	37838	Both	37833	2,3040	Abundant	5	0,0031	Rare
OTU_25	OTU_25_CP003495_Cyanobium_gracile_PCC_6307_99,5	Cyanobium	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	18516	Both	3068	0,1868	Average	15448	9,5842	Abundant
OTU_31	OTU_31_KM218868_Cyanobium_sp_UMPPCC_1207_99,5	Cyanobium	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	21505	Both	1798	0,1095	Average	19707	12,2266	Abundant
OTU_13	OTU_13_MH118732_Phormidesmis_priestleyi_ULC062_100	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	31287	Both	31284	1,9052	Abundant	3	0,0019	Rare
OTU_30	OTU_30_MH118741_Leptolyngbya_sp_ULC077_99,1	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	12797	Antarctica	12797	0,7793	Abundant	0	0,0000	Not_present
OTU_41	OTU_41_KY704160_Uncult_Chamaesiphon_99,2	Chamaesiphon	Chamaesiphonaceae	Synechococcales	unicellular	Cosmopolitan	24117	Both	23990	1,4610	Abundant	127	0,0788	Average
OTU_26	OTU_26_KY550452_Geitlerinema_sulphureum_I013-0011_99,7	Geitlerinema	Coleofasciculaceae	Oscillatoriales	large_filamentous	Cosmopolitan	21786	Both	21780	1,3264	Abundant	6	0,0037	Rare
OTU_20	OTU_20_KU219729_Phormidesmis_arctica_HOR_11-6_96,5	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	24638	Antarctica	24638	1,5005	Abundant	0	0,0000	Not_present
OTU_10	OTU_10_JQ793008_Uncult_cyano_100	Uncult_cyano	unknow	unknow	unknow	Polar	24179	Both	23506	1,4315	Abundant	673	0,4175	Average
OTU_23	OTU_23_JX413429_Uncult_Pseudanabaena_98,1	Pseudanabaena	Pseudanabaenaceae	Synechococcales	thin_filamentous	Endemic	23498	Both	23482	1,4301	Abundant	16	0,0099	Rare
OTU_28	OTU_28_AY493627_Phormidium_murrayii_ANT_LACV5_2_100	Phormidium	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	22602	Both	22599	1,3763	Abundant	3	0,0019	Rare
OTU_29	OTU_29_AF216952_Synechococcus_sp_PS715_99,7	Synechococcus	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	17524	Both	54	0,0033	Rare	17470	10,8387	Abundant
OTU_24	OTU_24_OR129827_Timaviella_sp_CNP3-B1-c4_99,5	Timaviella	Oscillatoriaceae	Synechococcales	thin_filamentous	Cosmopolitan	19188	Antarctica	19188	1,1686	Abundant	0	0,0000	Not_present
OTU_40	OTU_40_LT634149_Timaviella_circinata_GR4_99,2	Timaviella	Oscillatoriaceae	Synechococcales	thin_filamentous	Cosmopolitan	14499	Antarctica	14499	0,8830	Abundant	0	0,0000	Not_present
OTU_50	OTU_50_OR582731_Anagnostidinema_sp_LHM-MM_100	Anagnostidinema	Coleofasciculaceae	Oscillatoriales	large_filamentous	Cosmopolitan	11606	Both	11592	0,7060	Abundant	14	0,0087	Rare
OTU_34	OTU_34_MK139959_Cyanobium_sp_ULC084_100	Cyanobium	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	11582	Both	2976	0,1812	Average	8606	5,3393	Abundant
OTU_35	OTU_35_AY493573_Leptolyngbya_frigida_ANT_LMA_1_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	11576	Antarctica	11576	0,7050	Abundant	0	0,0000	Not_present
OTU_32	OTU_32_NR_172618_Altericista_violacea_strain_CALU_1077_98,4	Altericista	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	10610	Both	313	0,0191	Rare	10297	6,3884	Abundant
OTU_39	OTU_39_KY704123_Uncult_Chamaesiphon_98,4	Chamaesiphon	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	10448	Both	6140	0,3739	Average	4308	2,6728	Abundant
OTU_49	OTU_49_LT600732_Pseudanabaena_sp_RL-08-7_100	Pseudanabaena	Pseudanabaenaceae	Synechococcales	thin_filamentous	Cosmopolitan	8463	Both	7076	0,4309	Abundant	1387	0,8605	Abundant
OTU_47	OTU_47_NR_186875_Drouetiella_epilithica_strain_ACKU670_98,9	Drouetiella	Oscillatoriaceae	Synechococcales	thin_filamentous	Endemic	7775	Both	7772	0,4733	Abundant	3	0,0019	Rare
OTU_48	OTU_48_AB058264_Synechococcus_sp_MBIC10456_99,7	Synechococcus	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	7693	Both	60	0,0037	Rare	7633	4,7356	Abundant
OTU_46	OTU_46_KP688591_Scytolyngbya_timolectis_XS01_100	Scytolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	7329	Both	7328	0,4463	Abundant	1	0,0006	Rare
OTU_85	OTU_85_KU219736_Phormidesmis_priestleyi_LK057_99,2	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	6904	Both	6871	0,4184	Abundant	33	0,0205	Rare
OTU_51	OTU_51_ON897678_Drouetiella_hepatica_KPABG_4161_99,4	Drouetiella	Oscillatoriaceae	Synechococcales	thin_filamentous	Cosmopolitan	6062	Both	6046	0,3682	Average	16	0,0099	Rare
OTU_45	OTU_45_MN864655_Oscillatoria_sp_KH_15_99,7	Oscillatoria	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	5947	Both	5942	0,3619	Average	5	0,0031	Rare
OTU_65	OTU_65_MH745045_Nodularia_spumigena_CHS1_100	Nodularia	Aphanizomenonaceae	Nostocales	heterocystous	Cosmopolitan	5497	Both	5495	0,3346	Average	2	0,0012	Rare
OTU_63	OTU_63_AY493587_Pseudophormidium_sp_ANT_LPE_3_100	Pseudophormidium	Microcoleaceae	Oscillatoriales	large_filamentous	Endemic	5294	Antarctica	5294	0,3224	Average	0	0,0000	Not_present
OTU_57	OTU_57_HM018692_Tildenella_nuda_ZEHNDER_1965_U140_99,7	Tildenella	Oscillatoriaceae	Synechococcales	thin_filamentous	Alpine	5230	Both	5227	0,3183	Average	3	0,0019	Rare
OTU_55	OTU_55_KX263934_Leptolyngbya_sp_CALU1714_96,5	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	4937	Both	13	0,0008	Rare	4924	3,0549	Abundant
OTU_52	OTU_52_AY493620_Phormidesmis_priestleyi_ANT_LPR2_5_100	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	4876	Both	4863	0,2962	Average	13	0,0081	Rare
OTU_84	OTU_84_MK861882_Microcoleus_vaginatus_T28_100	Microcoleus	Microcoleaceae	Oscillatoriales	large_filamentous	Cosmopolitan	4825	Both	4805	0,2926	Average	20	0,0124	Rare
OTU_36	OTU_36_KU219740_Phormidesmis_nigrescens_RL_08-8_100	Phormidesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	4433	Both	3716	0,2263	Average	717	0,4448	Average
OTU_66	OTU_66_KF704323_Nostoc_sp_Pannaria_rubiginosa_LG_99,7	Nostoc	Nostocaceae	Nostocales	heterocystous	Cosmopolitan	4539	Both	4537	0,2763	Average	2	0,0012	Rare

OTU_59	OTU_59_KU219718_Phormidiesmis_priestleyi_AB_12-7_100	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	3974	Both	3868	0.2356	Average	106	0.0658	Average
OTU_64	OTU_64_MF527197_Uncult_Leptolyngbya_99,7	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	4058	Antarctica	4058	0.2471	Average	0	0.0000	Not_present
OTU_684	OTU_684_GQ859651_Phormidium_sp_PMC301_07_97,3	Phormidium	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	3357	Antarctica	3357	0.2044	Average	0	0.0000	Not_present
OTU_56	OTU_56_MH683785_Phormidium_sp_YACCYB597_100	Phormidium	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	3704	Both	3702	0.2255	Average	2	0.0012	Rare
OTU_60	OTU_60_MF781040_Cyanothece_aeruginosa_SAG_87_99,5	Cyanothece	Cyanothecaceae	Oscillatoriales	unicellular	Cosmopolitan	3675	Both	3671	0.2236	Average	4	0.0025	Rare
OTU_75	OTU_75_KU219734_Phormidiesmis_nigrescens_LK015_100	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	3263	Both	3255	0.1982	Average	8	0.0050	Rare
OTU_81	OTU_81_AB075995_Pseudanabaena_sp_PCC_7403_96,2	Pseudanabaena	Pseudanabaenaceae	Synechococcales	thin_filamentous	Endemic	3161	Antarctica	3161	0.1925	Average	0	0.0000	Not_present
OTU_88	OTU_88_KU219715_Phormidiesmis_communis_AB_11-10_98,9	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	2822	Antarctica	2822	0.1719	Average	0	0.0000	Not_present
OTU_655	OTU_655_GQ859651_Phormidium_sp_PMC301_07_96,8	Phormidium	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	2522	Both	2521	0.1535	Average	1	0.0006	Rare
OTU_98	OTU_98_KY704146_Uncult_Chamaesiphon_99,5	Chamaesiphon	Chamaesiphonaceae	Synechococcales	unicellular	Cosmopolitan	2786	Antarctica	2786	0.1697	Average	0	0.0000	Not_present
OTU_83	OTU_83_AY493576_Leptolyngbya_frigida_ANT_L53B_2_99,7	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	2698	Both	2631	0.1602	Average	67	0.0416	Rare
OTU_131	OTU_131_KF934181_Rexia_erecta_CAT4-SG4_100	Rexia	Tolypothrichaceae	Nostocales	heterocystous	Cosmopolitan	2665	Both	2539	0.1546	Average	126	0.0782	Average
OTU_87	OTU_87_MT733179_Leptolyngbya_sp_CHAB4063_clone_1_98,6	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	2582	Antarctica	2582	0.1572	Average	0	0.0000	Not_present
OTU_73	OTU_73_KU309580_Microcoleus_sp_GL_08-8_16S_100	Microcoleus	Microcoleaceae	Oscillatoriales	large_filamentous	Cosmopolitan	2516	Both	1161	0.0707	Average	1355	0.8407	Abundant
OTU_86	OTU_86_CP045226_Nostoc_sphaeroides_CCNUC1_100	Nostoc	Nostocaceae	Nostocales	heterocystous	Cosmopolitan	2295	Both	2221	0.1353	Average	74	0.0459	Average
OTU_114	OTU_114_MH688850_Timaviella_sp_Us-6-3_98,4	Timaviella	Oculatellaceae	Synechococcales	thin_filamentous	Cosmopolitan	2323	Antarctica	2323	0.1415	Average	0	0.0000	Not_present
OTU_151	OTU_151_LN849931_Phormidiesmis_sp_LD2_4800_EK_99,2	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	2242	Both	2139	0.1303	Average	103	0.0639	Average
OTU_91	OTU_91_KF417657_Synechococcus_sp_SAG_2387_96,7	Synechococcus	Synechococcaceae	Synechococcales	unicellular	Endemic	2151	Both	2148	0.1308	Average	3	0.0019	Rare
OTU_135	OTU_135_JN814164_Uncult_cyano_100	Uncult_cyano	unknow	unknow	unknow	Endemic	2132	Antarctica	2132	0.1298	Average	0	0.0000	Not_present
OTU_69	OTU_69_AJ781042_Snowella_rosea_1LM40S01_97,0	Snowella	Coelospaeriaceae	Synechococcales	unicellular	Endemic	2104	Antarctica	2104	0.1281	Average	0	0.0000	Not_present
OTU_108	OTU_108_AJ293105_Anabaena_solitaria_99,7	Anabaena	Nostocaceae	Nostocales	heterocystous	Cosmopolitan	2069	Both	2066	0.1258	Average	3	0.0019	Rare
OTU_76	OTU_76_ON897672_Leptodesmis_sp_KPABG_610043_98,8	Leptodesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	1993	Both	1943	0.1183	Average	50	0.0310	Rare
OTU_154	OTU_154_JQ687331_Leptolyngbya_antarctica_CYN65_98,9	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	1908	Both	1907	0.1161	Average	1	0.0006	Rare
OTU_252	OTU_252_KY379861_Cyanobium_sp_CZS_22C_98,9	Cyanobium	Synechococcaceae	Synechococcales	unicellular	Endemic	1438	Both	50	0.0030	Rare	1388	0.8611	Abundant
OTU_110	OTU_110_GQ162217_Leptolyngbya_sp_OU_6_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	1719	Both	1717	0.1046	Average	2	0.0012	Rare
OTU_100	OTU_100_JF925324_Leptolyngbya_sp_CYN82_98,9	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	1589	Both	2	0.0001	Rare	1587	0.9846	Abundant
OTU_403	OTU_403_KU219720_Phormidiesmis_nigrescens_GL_08-2_99,2	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	1446	Both	749	0.0456	Average	697	0.4324	Average
OTU_92	OTU_92_NR_172596_Nodosilinea_signiensis_strain_USMFM_100	Nodosilinea	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	1368	Antarctica	1368	0.0833	Average	0	0.0000	Not_present
OTU_115	OTU_115_KM019984_Microchaete_tenera_SAG_11_99_97,8	Microchaete	Rivulariaceae	Nostocales	heterocystous	Endemic	1331	Antarctica	1331	0.0811	Average	0	0.0000	Not_present
OTU_129	OTU_129_KU219738_Phormidiesmis_arctica_MUM_11-8_96,5	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	1311	Both	1310	0.0798	Average	1	0.0006	Rare
OTU_117	OTU_117_OM732238_Anagnostidinema_sp_BACA0434_100	Anagnostidinema	Coleofasciculaceae	Oscillatoriales	large_filamentous	Cosmopolitan	1299	Sub_Antarctica	0	0.0000	Not_present	1299	0.8059	Abundant
OTU_111	OTU_111_MT946564_Pseudophormidium_battersii_NN-3-1-D_100	Pseudophormidium	Microcoleaceae	Oscillatoriales	large_filamentous	Cosmopolitan	1088	Both	1086	0.0661	Average	2	0.0012	Rare
OTU_400	OTU_400_AY493582_Phormidiesmis_priestleyi_ANT_L61_2_98,9	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	1160	Both	194	0.0118	Rare	966	0.5993	Abundant
OTU_175	OTU_175_CP003594_Synechococcus_sp_PCC_7502_99,8	Synechococcus	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	1140	Both	1135	0.0691	Average	5	0.0031	Rare
OTU_160	OTU_160_OQ079418_Leptolyngbya_sp_PS30-1.1_99,8	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	1130	Antarctica	1130	0.0688	Average	0	0.0000	Not_present
OTU_178	OTU_178_KY379890_Cyanobium_sp_CZS_48M_100	Cyanobium	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	1124	Both	19	0.0012	Rare	1105	0.6856	Abundant
OTU_127	OTU_127_HQ012544_Phormidium_arthurensis_LCR-OSC4b_99,7	Phormidium	Oscillatoriaceae	Oscillatoriales	large_filamentous	Alpine	1126	Both	1113	0.0678	Average	13	0.0081	Rare
OTU_107	OTU_107_HM217052_Synechococcus_sp_LEGE_06306_99,5	Synechococcus	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	944	Both	377	0.0230	Average	567	0.3518	Average
OTU_737	OTU_737_KU219715_Phormidiesmis_communis_AB_11-10_98,7	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	1021	Both	1004	0.0611	Average	17	0.0105	Rare
OTU_134	OTU_134_MK487623_Funiculus_sp_HSN013_99,5	Funiculus	Coleofasciculaceae	Oscillatoriales	large_filamentous	Cosmopolitan	1031	Both	956	0.0582	Average	75	0.0465	Average
OTU_222	OTU_222_AY493620_Phormidiesmis_priestleyi_ANT_LPR2_5_98,9	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	1022	Both	1020	0.0621	Average	2	0.0012	Rare
OTU_136	OTU_136_KC815479_Uncult_Leptolyngbya_98,9	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	944	Both	813	0.0495	Average	131	0.0813	Average
OTU_191	OTU_191_AY790861_Leptolyngbya_sp_CCME6119_97,3	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	912	Both	699	0.0426	Average	213	0.1321	Average
OTU_680	OTU_680_KU219740_Phormidiesmis_nigrescens_RL_08-8_98,9	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	903	Both	541	0.0329	Average	362	0.2246	Average
OTU_206	OTU_206_KY704155_Uncult_Chamaesiphon_98,4	Chamaesiphon	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	901	Both	77	0.0047	Rare	824	0.5112	Average
OTU_1044	OTU_1044_AY493627_Phormidium_murrayii_ANT_LACV5_2_98,6	Phormidium	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	897	Antarctica	897	0.0546	Average	0	0.0000	Not_present
OTU_149	OTU_149_MH688849_Stenomitos_sp_Ru-0-2_98,1	Stenomitos	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	861	Both	14	0.0009	Rare	847	0.5255	Average
OTU_497	OTU_497_KU219729_Phormidiesmis_arctica_HOR_11-6_98,4	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	837	Both	834	0.0508	Average	3	0.0019	Rare
OTU_174	OTU_174_JX413429_Uncult_Pseudanabaena_94,1	Pseudanabaena	Pseudanabaenaceae	Synechococcales	thin_filamentous	Alpine	798	Antarctica	798	0.0486	Average	0	0.0000	Not_present
OTU_184	OTU_184_AP014815_Geminocystis_sp_NIES-3708_99,2	Geminocystis	Chroococcaceae	Chroococcales	unicellular	Cosmopolitan	780	Both	1	0.0001	Rare	779	0.4833	Average
OTU_758	OTU_758_MH118732_Phormidiesmis_priestleyi_ULC062_98,6	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	766	Antarctica	766	0.0466	Average	0	0.0000	Not_present
OTU_189	OTU_189_MH118741_Leptolyngbya_sp_ULC077_98,3	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	746	Both	227	0.0138	Rare	519	0.3220	Average
OTU_167	OTU_167_KM052844_Leptolyngbya_sp_LCR-CYANT35_98,4	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Polar	743	Antarctica	743	0.0452	Average	0	0.0000	Not_present
OTU_231	OTU_231_HM217078_Synechococcus_sp_LEGE_07074_99,5	Synechococcus	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	719	Antarctica	719	0.0438	Average	0	0.0000	Not_present
OTU_155	OTU_155_KU219710_Leptolyngbya_sp_JR_09_4_100	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	714	Antarctica	714	0.0435	Average	0	0.0000	Not_present
OTU_185	OTU_185_KF417430_Stenomitos_rutilans_HA7619-LM2_99,7	Stenomitos	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	693	Both	2	0.0001	Rare	691	0.4287	Average
OTU_243	OTU_243_FJ490330_Uncult_Oscillatoriales_100	unknow	unknow	Oscillatoriales	large_filamentous	Polar	691	Both	684	0.0417	Average	7	0.0043	Rare
OTU_456	OTU_456_ON803467_Stenomitos_frigidus_BEA_1253B_98,1	Stenomitos	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	687	Antarctica	687	0.0418	Average	0	0.0000	Not_present
OTU_1049	OTU_1049_HQ201392_Oscillatoria_sp_KNUA009_95,9	Oscillatoria	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	628	Antarctica	628	0.0382	Average	0	0.0000	Not_present
OTU_207	OTU_207_KU219718_Phormidiesmis_priestleyi_AB_12-7_98,9	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	663	Both	662	0.0403	Average	1	0.0006	Rare
OTU_143	OTU_143_AY768406_Oscillatoria_sp_CCAP_1459_13_99,5	Oscillatoria	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	647	Antarctica	647	0.0394	Average	0	0.0000	Not_present
OTU_183	OTU_183_MK248006_Leptolyngbya_cf_frigida_ACT684_96,5	Leptolyngbya	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	598	Both	240	0.0146	Rare	358	0.2221	Average
OTU_282	OTU_282_OK576203_Aliterella_sp_ULC_0197_100	Aliterella	Aliterellaceae	Chroococcidiopsidales	baeocytous	Cosmopolitan	582	Both	579	0.0353	Average	3	0.0019	Rare
OTU_909	OTU_909_KU219716_Phormidiesmis_priestleyi_AB_12-2_98,6	Phormidiesmis	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	582	Both	83	0.0051	Rare	499	0.3096	Average
OTU_260	OTU_260_MH118734_Pseudanabaena_frigida_ULC067_99,5	Pseudanabaena	Pseudanabaenaceae	Synechococcales	thin_filamentous	Cosmopolitan	568	Antarctica	568	0.0346	Average	0	0.0000	Not_present
OTU_210	OTU_210_NR_117190_Calothrix_elsteri_strain_CCALA_953_99,5	Calothrix	Rivulariaceae	Nostocales	heterocystous	Alpine	568	Antarctica	568	0.0346	Average	0	0.0000	Not_present
OTU_212	OTU_212_KT983934_Uncult_Chamaesiphon_100	Chamaesiphon	Chamaesiphonaceae	Synechococcales	unicellular	Cosmopolitan	564	Both	496	0.0302	Average	68	0.0422	Rare
OTU_193	OTU_193_LN849922_Leptolyngbyaceae_cyanobacterium_LD4_4600	Leptolyngbyaceae	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	563	Both	560	0.				

Chapter 2 | 71

OTU_342	OTU_342_KY501891_Uncult_Arthrospira_94,6	<i>uncult_bact</i>	unknown	unknown	unknown	Cosmopolitan	209	Sub_Antarctica	0	0,0000	Not_present	209	0,1297	Average
OTU_349	OTU_349_MT335717_Uncultured_Leptolyngbya_sp._clone_OTU16_98,1	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Alpine	202	Antarctica	202	0,0123	Rare	0	0,0000	Not_present
OTU_395	OTU_395_JQ793008_Uncult_cyano_98,6	<i>Uncult_cyano</i>	unknown	unknown	unknown	Endemic	147	Both	2	0,0001	Rare	145	0,0900	Average
OTU_345	OTU_345_GQ375044_Chroococcus_sp_CCALA_702_97,0	<i>Chroococcus</i>	Chroococcaceae	Chroococcales	unicellular	Endemic	191	Antarctica	191	0,0116	Rare	0	0,0000	Not_present
OTU_495	OTU_495_KF941267_Planktothrix_sp_2H9_99,5	<i>Planktothrix</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Cosmopolitan	191	Antarctica	191	0,0116	Rare	0	0,0000	Not_present
OTU_1166	OTU_1166_KF663711_Leptolyngbya_frigida_I3_98,8	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	184	Both	183	0,0111	Rare	1	0,0006	Rare
OTU_374	OTU_374_MK861909_Tildenella_sp_KL16_100	<i>Tildenella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Alpine	183	Antarctica	183	0,0111	Rare	0	0,0000	Not_present
OTU_389	OTU_389_KC004019_Aphanothece_caldoriarum_var_cavernarum_9_8,4	<i>Aphanothece</i>	Aphanothecaceae	Chroococcales	unicellular	Endemic	182	Antarctica	182	0,0111	Rare	0	0,0000	Not_present
OTU_354	OTU_354_KC311882_Uncult_Microcoleus_98,1	<i>Microcoleus</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Endemic	178	Sub_Antarctica	0	0,0000	Not_present	178	0,1104	Average
OTU_1192	OTU_1192_HM217078_Synechococcus_sp_LEGE_07074_98,9	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Endemic	176	Antarctica	176	0,0107	Rare	0	0,0000	Not_present
OTU_390	OTU_390_OQ079425_Pseudophormidium_sp_P52_99,7	<i>Pseudophormidium</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Endemic	175	Antarctica	175	0,0107	Rare	0	0,0000	Not_present
OTU_393	OTU_393_KY303730_Leptolyngbya_sp_THF_100	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	169	Both	166	0,0101	Rare	3	0,0019	Rare
OTU_608	OTU_608_KX086282_Leptolyngbya_sp_B-CY-NM28_99,7	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	167	Both	162	0,0099	Rare	5	0,0031	Rare
OTU_391	OTU_391_AF218371_Stenomitos_tremulus_98,9	<i>Stenomitos</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	166	Sub_Antarctica	0	0,0000	Not_present	166	0,1030	Average
OTU_1134	OTU_1134_LT634149_Timaviella_circinata_GR4_98,4	<i>Timaviella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Polar	159	Both	157	0,0096	Rare	2	0,0012	Rare
OTU_516	OTU_516_JF832304_Uncult_Chamaesiphon_98,6	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Cosmopolitan	159	Both	154	0,0094	Rare	5	0,0031	Rare
OTU_484	OTU_484_JF832304_Uncult_Chamaesiphon_99,2	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Alpine	155	Both	151	0,0092	Rare	4	0,0025	Rare
OTU_1307	OTU_1307_KY704109_Chamaesiphon_geitleri_No1023_98,4	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	101	Antarctica	101	0,0062	Rare	0	0,0000	Not_present
OTU_463	OTU_463_OR129827_Timaviella_sp_CNP3-B1-c4_16S_98,9	<i>Timaviella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Endemic	152	Antarctica	152	0,0093	Rare	0	0,0000	Not_present
OTU_438	OTU_438_ON037524_Uncultured_Leptolyngbya_sp_clone_122_97,3	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	150	Antarctica	150	0,0091	Rare	0	0,0000	Not_present
OTU_427	OTU_427_AY493597_Calothrix_sp_ANT_LPR2_4_100	<i>Calothrix</i>	Rivulariaceae	Nostocales	heterocystous	Cosmopolitan	148	Antarctica	148	0,0090	Rare	0	0,0000	Not_present
OTU_707	OTU_707_KJ705094_Leptolyngbya_sp_PUPCCC_112_2_98,4	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	147	Antarctica	147	0,0090	Rare	0	0,0000	Not_present
OTU_633	OTU_633_MT174455_Leptolyngbyaceae_96,8	<i>Leptolyngbyaceae</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	147	Antarctica	147	0,0090	Rare	0	0,0000	Not_present
OTU_422	OTU_422_MT488259_Synechococcus_sp_CHAB_TP201820.1_98,6	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Endemic	146	Both	2	0,0001	Rare	144	0,0893	Average
OTU_1007	OTU_1007_NR_172596_Nodosilinea_signiensis_strain_USMFM_98,3	<i>Nodosilinea</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	143	Antarctica	143	0,0087	Rare	0	0,0000	Not_present
OTU_522	OTU_522_KT315945_Anagnostidinema_carotinosum_JR21_100	<i>Anagnostidinema</i>	Coleofasciculaceae	Oscillatoriales	large_filamentous	Alpine	141	Both	113	0,0069	Rare	28	0,0174	Rare
OTU_679	OTU_679_JQ687331_Leptolyngbya_antarctica_CYN65_97,8	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	134	Antarctica	134	0,0082	Rare	0	0,0000	Not_present
OTU_436	OTU_436_KY807918_Leptolyngbya_valderiana_SABC022801_100	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	60	Sub_Antarctica	0	0,0000	Not_present	60	0,0372	Rare
OTU_520	OTU_520_KU161657_Leptolyngbyaceae_99,5	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	126	Antarctica	126	0,0077	Rare	0	0,0000	Not_present
OTU_573	OTU_573_MK861876_Kamptonomia_sp_V9_99,5	<i>Phormidium</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	126	Both	116	0,0071	Rare	10	0,0062	Rare
OTU_451	OTU_451_KM019963_Hydrocoleum_sp_SAG_38_87_94,3	<i>Lyngbya</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	125	Both	1	0,0001	Rare	124	0,0769	Average
OTU_1146	OTU_1146_KU219697_Microcoleus_sp_LENS_2011_47_98,4	<i>Microcoleus</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Endemic	125	Both	124	0,0076	Rare	1	0,0006	Rare
OTU_769	OTU_769_HQ201392_Oscillatoria_sp_KNUA009_97,0	<i>Oscillatoria</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	113	Antarctica	113	0,0069	Rare	0	0,0000	Not_present
OTU_1354	OTU_1354_EU729033_Synechococcus_sp_AECC1334_99,7	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	124	Both	14	0,0009	Rare	110	0,0682	Average
OTU_545	OTU_545_AY493586_Phormidismis_priestleyi_ANT_LACV5_1_100	<i>Phormidismis</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	123	Antarctica	123	0,0075	Rare	0	0,0000	Not_present
OTU_511	OTU_511_MK209097_Nostoc_flagelliforme_CHAB_2816_100	<i>Nostoc</i>	Nostocaceae	Nostocales	heterocystous	Cosmopolitan	123	Both	120	0,0073	Rare	3	0,0019	Rare
OTU_1331	OTU_1331_MK861906_Shackletoniella_antarctica_KL15_94,8	<i>Shackletoniella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Endemic	121	Antarctica	121	0,0074	Rare	0	0,0000	Not_present
OTU_488	OTU_488_KY283068_Phormidismis_sp_ACSS1_170_99,2	<i>Phormidismis</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	121	Antarctica	121	0,0074	Rare	0	0,0000	Not_present
OTU_1250	OTU_1250_AY790861_Leptolyngbya_sp_CCME6119_94,6	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	118	Both	14	0,0009	Rare	104	0,0645	Average
OTU_401	OTU_401_AY768406_Oscillatoria_sp_CCAP_1459_13_98,6	<i>Oscillatoria</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	117	Antarctica	117	0,0071	Rare	0	0,0000	Not_present
OTU_469	OTU_469_EU729079_Plectonema_sp_AECC1345_98,6	<i>Plectonema</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	117	Both	6	0,0004	Rare	111	0,0689	Average
OTU_722	OTU_722_MF781048_Cyanothece_svehlovae_NIVA-CYA_258_97,8	<i>Cyanothece</i>	Cyanothecaceae	Oscillatoriales	unicellular	Endemic	113	Antarctica	113	0,0069	Rare	0	0,0000	Not_present
OTU_1451	OTU_1451_HQ201392_Oscillatoria_sp_KNUA009_97,8	<i>Oscillatoria</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	102	Antarctica	102	0,0062	Rare	0	0,0000	Not_present
OTU_585	OTU_585_AY790844_Leptolyngbya_sp_CCME6034_99,5	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	107	Both	104	0,0063	Rare	3	0,0019	Rare
OTU_568	OTU_568_JX575081_Uncult_Desertifilum_96,2	<i>Desertifilum</i>	Coleofasciculaceae	Oscillatoriales	large_filamentous	Polar	103	Antarctica	103	0,0063	Rare	0	0,0000	Not_present
OTU_685	OTU_685_JQ687331_Leptolyngbya_antarctica_CYN65_97,8	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	94	Antarctica	94	0,0057	Rare	0	0,0000	Not_present
OTU_597	OTU_597_KM153750_uncult_bact_98,4	<i>uncult_bact</i>	unknown	unknown	unknown	Endemic	93	Sub_Antarctica	0	0,0000	Not_present	93	0,0577	Average
OTU_558	OTU_558_MH688842_Hormoscilla_cf_pringsheimii_Us-s-6-2_98,6	<i>Hormoscilla</i>	Gomontiellaceae	Oscillatoriales	large_filamentous	Cosmopolitan	90	Antarctica	90	0,0055	Rare	0	0,0000	Not_present
OTU_530	OTU_530_MH742932_Stenomitos_frigidus_AR3-EA3_100	<i>Stenomitos</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	79	Both	78	0,0048	Rare	1	0,0006	Rare
OTU_547	OTU_547_OQ079417_Aphanocapsa_sp_S121_100	<i>Aphanocapsa</i>	Merismopediaceae	Synechococcales	unicellular	Cosmopolitan	84	Both	83	0,0051	Rare	1	0,0006	Rare
OTU_582	OTU_582_MH453375_Anabaena_sp_B6_94,9	<i>Anabaena</i>	Nostocaceae	Nostocales	heterocystous	Endemic	84	Sub_Antarctica	0	0,0000	Not_present	84	0,0521	Average
OTU_840	OTU_840_UN979956_Leptolyngbya_borchgrevinkii_KOVACIK-ANT_98,1	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	77	Antarctica	77	0,0047	Rare	0	0,0000	Not_present
OTU_632	OTU_632_MF527176_Uncult_Oscillatoriales_98,9	<i>unknown</i>	unknown	Oscillatoriales	large_filamentous	Polar	81	Both	80	0,0049	Rare	1	0,0006	Rare
OTU_1375	OTU_1375_KU219697_Microcoleus_sp_LENS_2011_47_97,0	<i>Microcoleus</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Endemic	80	Antarctica	80	0,0049	Rare	0	0,0000	Not_present
OTU_613	OTU_613_HM217078_Synechococcus_sp_LEGE_07074_98,9	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Endemic	80	Antarctica	80	0,0049	Rare	0	0,0000	Not_present
OTU_979	OTU_979_JQ687331_Leptolyngbya_antarctica_CYN65_98,6	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	77	Antarctica	77	0,0047	Rare	0	0,0000	Not_present
OTU_535	OTU_535_MH427706_Nostoc_sp_SV1-VF1_clone_B_98,9	<i>Nostoc</i>	Nostocaceae	Nostocales	heterocystous	Endemic	76	Antarctica	76	0,0046	Rare	0	0,0000	Not_present
OTU_711	OTU_711_KF359686_Nostoc_sp_Nephroma_arcticum_UK564_cyano_biont_99,7	<i>Nostoc</i>	Nostocaceae	Nostocales	heterocystous	Cosmopolitan	76	Antarctica	76	0,0046	Rare	0	0,0000	Not_present

OTU_62	OTU_62_MF527187_Uncult_Leptolyngbya_96.2	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Polar	75	Both	73	0,0044	Rare	2	0,0012	Rare
OTU_576	OTU_576_KY704155_Uncult_Chamaesiphon_98.1	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	75	Both	1	0,0001	Rare	74	0,0459	Rare
OTU_670	OTU_670_KP688591_Scytolyngbya_timoelontis_XS01_100	<i>Scytolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	71	Both	68	0,0041	Rare	3	0,0019	Rare
OTU_634	OTU_634_AJ781039_Snowella_litoralis_OTU35S07_99.5	<i>Snowella</i>	Coelosphaeriaceae	Synechococcales	unicellular	Cosmopolitan	42	Sub_Antarctica	0	0,0000	Not_present	42	0,0261	Rare
OTU_577	OTU_577_KY704115_Chamaesiphon_cf_incrustans_str_Ch_fontanille_95.1	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	70	Sub_Antarctica	0	0,0000	Not_present	70	0,0434	Rare
OTU_565	OTU_565_OK382067_Leptolyngbya_sp_SC-S3D_100	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	65	Antarctica	65	0,0040	Rare	0	0,0000	Not_present
OTU_778	OTU_778_KR923206_Uncult_cyano_100	<i>Uncult_cyano</i>	unknow	unknow	unknow	Cosmopolitan	65	Antarctica	65	0,0040	Rare	0	0,0000	Not_present
OTU_1132	OTU_1132_GQ859652_Leptolyngbya_tenuis_PMC304_07_99.2	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	64	Both	63	0,0038	Rare	1	0,0006	Rare
OTU_904	OTU_904_MF527193_Uncult_Chroococcidiopsis_sp_clone_OTU_57_98.7	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Cosmopolitan	61	Antarctica	61	0,0037	Rare	0	0,0000	Not_present
OTU_324	OTU_324_KY704115_Chamaesiphon_cf_incrustans_str_Ch_Fontanille_94.8	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	56	Both	51	0,0031	Rare	5	0,0031	Rare
OTU_936	OTU_936_MT733179_Leptolyngbya_sp_CHAB4063_clone_1_99.4	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	60	Antarctica	60	0,0037	Rare	0	0,0000	Not_present
OTU_1263	OTU_1263_HM241099_Uncult_Chroococcidiopsis_99.7	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Cosmopolitan	60	Both	54	0,0033	Rare	6	0,0037	Rare
OTU_429	OTU_429_DQ519736_Uncult_Synechococcus_100	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	34	Both	1	0,0001	Rare	33	0,0205	Rare
OTU_1149	OTU_1149_KJ705094_Leptolyngbya_sp_PUPCCC_112_2_97.6	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	57	Antarctica	57	0,0035	Rare	0	0,0000	Not_present
OTU_1324	OTU_1324_LT600742_Nodosilinea_sp_Tapi_15_97.3	<i>Nodosilinea</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	56	Antarctica	56	0,0034	Rare	0	0,0000	Not_present
OTU_1243	OTU_1243_ON803467_Stenomitos_frigidus_BEa_1253B_97.6	<i>Stenomitos</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	55	Antarctica	55	0,0033	Rare	0	0,0000	Not_present
OTU_642	OTU_642_FM177503_Synechocystis_aquatilis_1LT32S04_94.0	<i>Synechocystis</i>	Merismopediaceae	Synechococcales	unicellular	Endemic	54	Antarctica	54	0,0033	Rare	0	0,0000	Not_present
OTU_1019	OTU_1019_KF631399_Anabaena_sp_XSPORK7B_99.7	<i>Anabaena</i>	Nostocaceae	Nostocales	heterocystous	Cosmopolitan	52	Both	51	0,0031	Rare	1	0,0006	Rare
OTU_749	OTU_749_JQ310280_Uncult_cyano_98.9	<i>Uncult_cyano</i>	unknow	unknow	unknow	Endemic	51	Sub_Antarctica	0	0,0000	Not_present	51	0,0316	Rare
OTU_854	OTU_854_MG652769_Gloeobacter_95.1	<i>Gloeobacter</i>	Gloeobacteraceae	Gloeobacterales	unicellular	Cosmopolitan	50	Antarctica	50	0,0030	Rare	0	0,0000	Not_present
OTU_814	OTU_814_KF856486_Uncult_cyano_98.0	<i>Uncult_cyano</i>	unknow	unknow	unknow	Endemic	49	Both	3	0,0002	Rare	46	0,0285	Rare
OTU_878	OTU_878_LC325262_Pseudocapsa_sp_Ryu8-6_96.8	<i>Pseudocapsa</i>	Chroococcaceae	Chroococcales	unicellular	Polar	46	Both	27	0,0016	Rare	19	0,0118	Rare
OTU_775	OTU_775_KT354214_Synechocystis_fuscopigmentosa_CCALA_191_98.6	<i>Synechocystis</i>	Merismopediaceae	Synechococcales	unicellular	Endemic	46	Sub_Antarctica	0	0,0000	Not_present	46	0,0285	Rare
OTU_908	OTU_908_KX367890_Uncult_Synechococcaceae_97.6	unknow	Synechococcaceae	Synechococcales	unicellular	Endemic	46	Sub_Antarctica	0	0,0000	Not_present	46	0,0285	Rare
OTU_1217	OTU_1217_DQ264223_Leptolyngbya_sp_1ES25S1_99.7	<i>Leptolyngbyaceae</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	44	Antarctica	44	0,0027	Rare	0	0,0000	Not_present
OTU_1129	OTU_1129_KY078780_Nodosilinea_sp_GSE-PSE-MK55-09B_97.6	<i>Nodosilinea</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	42	Antarctica	42	0,0026	Rare	0	0,0000	Not_present
OTU_834	OTU_834_KY704115_Chamaesiphon_cf_incrustans_str_Ch_fontanille_94.4	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	41	Sub_Antarctica	0	0,0000	Not_present	41	0,0254	Rare
OTU_1180	OTU_1180_KT347094_Phormidium_pseudopriestleyi_FRX01_97.0	<i>Phormidium</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Endemic	39	Antarctica	39	0,0024	Rare	0	0,0000	Not_present
OTU_665	OTU_665_NR_177027_Timaviella_circinata_strain_personal::GR4_9_8.9	<i>Timaviella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Polar	37	Antarctica	37	0,0023	Rare	0	0,0000	Not_present
OTU_743	OTU_743_KY704115_Chamaesiphon_cf_incrustans_str_Ch_fontanille_95.7	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	36	Antarctica	36	0,0022	Rare	0	0,0000	Not_present
OTU_1057	OTU_1057_MH688842_Hormoscilla_cf_pringsheimii_Us-s-6-2_99.7	<i>Hormoscilla</i>	Gomontelliaceae	Oscillatoriales	large_filamentous	Cosmopolitan	35	Both	24	0,0015	Rare	11	0,0068	Rare
OTU_986	OTU_986_KC004019_Aphanothece_caldoriarum_var_cavernarum_9_8.9	<i>Aphanothece</i>	Aphanothecaceae	Chroococcales	unicellular	Alpine	34	Antarctica	34	0,0021	Rare	0	0,0000	Not_present
OTU_1078	OTU_1078_JQ701219_Uncult_Synechococcus_95.7	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Alpine	34	Sub_Antarctica	0	0,0000	Not_present	34	0,0211	Rare
OTU_895	OTU_895_KY704150_Uncult_Chamaesiphon_98.6	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	34	Sub_Antarctica	0	0,0000	Not_present	34	0,0211	Rare
OTU_1126	OTU_1126_JQ687331_Leptolyngbya_antarctica_CYN65_98.7	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	33	Antarctica	33	0,0020	Rare	0	0,0000	Not_present
OTU_1077	OTU_1077_JQ701262_Uncult_Chroococcales_94.9	unknow	unknow	Chroococcales	unicellular	Endemic	33	Sub_Antarctica	0	0,0000	Not_present	33	0,0205	Rare
OTU_1184	OTU_1184_KC525099_Chroococcidiopsis_sp_UFS-A4UI-NPMV4-B4_99.5	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Cosmopolitan	32	Antarctica	32	0,0019	Rare	0	0,0000	Not_present
OTU_782	OTU_782_KF417430_Stenomitos_rutilans_HA7619-LM2_99.2	<i>Stenomitos</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	32	Antarctica	32	0,0019	Rare	0	0,0000	Not_present
OTU_1063	OTU_1063_MN477097_Uncultured_Calothrix_sp_clone_em.359_96.8	<i>Calothrix</i>	Rivulariaceae	Nostocales	heterocystous	Alpine	20	Sub_Antarctica	0	0,0000	Not_present	20	0,0124	Rare
OTU_1305	OTU_1305_MT335717_Uncultured_Leptolyngbya_sp_clone_OTU16_97.6	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	31	Antarctica	31	0,0019	Rare	0	0,0000	Not_present
OTU_1020	OTU_1020_KY662381_Nostoc_sp_ARC_64_98.4	<i>Nostoc</i>	Nostocaceae	Nostocales	heterocystous	Endemic	29	Antarctica	29	0,0018	Rare	0	0,0000	Not_present
OTU_1167	OTU_1167_MN477087_Uncultured_Leptolyngbyaceae_clone_em.66_96.7	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	29	Antarctica	29	0,0018	Rare	0	0,0000	Not_present
OTU_848	OTU_848_LC716288_Synechococcus_sp_NIES-982_97	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Endemic	29	Antarctica	29	0,0018	Rare	0	0,0000	Not_present
OTU_1249	OTU_1249_KT867211_Stigonema_elegans_WYS5-2_99.7	<i>Stigonema</i>	Stigonemataceae	Nostocales	branching	Cosmopolitan	28	Both	24	0,0015	Rare	4	0,0025	Rare
OTU_1156	OTU_1156_KX670262_Cylindrospermum_majus_AUS-JR_MT_NT-097_96.5	<i>Cylindrospermum</i>	Nostocaceae	Nostocales	heterocystous	Endemic	28	Sub_Antarctica	0	0,0000	Not_present	28	0,0174	Rare
OTU_1365	OTU_1365_EU852497_Leptolyngbya_cf_foveolarum_TM3S129_99.5	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	25	Antarctica	25	0,0015	Rare	0	0,0000	Not_present
OTU_1136	OTU_1136_MW776414_Sphaerospermopsis_sp_BACA0031_98.7	<i>Sphaerospermopsis</i>	Aphanizomenonaceae	Nostocales	heterocystous	Endemic	27	Sub_Antarctica	0	0,0000	Not_present	27	0,0168	Rare
OTU_1137	OTU_1137_EU624415_Phormidium_sp_SAG_61_90_99.7	<i>Phormidium</i>	Oscillatoriaceae	Oscillatoriales	large_filamentous	Cosmopolitan	26	Antarctica	26	0,0016	Rare	0	0,0000	Not_present
OTU_1377	OTU_1377_MK938353_Uncult_Chroococcidiopsis_sp_clone_OTU54	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Endemic	25	Antarctica	25	0,0015	Rare	0	0,0000	Not_present

OTU_1256	OTU_1256_MF527146_Uncult_Microcoleus_95,7	<i>Microcoleus</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Polar	25	Antarctica	25	0,0015	Rare	0	0,0000	Not_present
OTU_969	OTU_969_AB075995_Pseudanabaena_sp_PCC_7403_96,2	<i>Pseudanabaena</i>	Pseudanabaenaceae	Synechococcales	thin_filamentous	Endemic	25	Sub_Antarctica	0	0,0000	Not_present	25	0,0155	Rare
OTU_1045	OTU_1045_MT488259_Synechococcus_sp_CHAB_TP201820.1_clone_06_98,4	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Endemic	23	Sub_Antarctica	0	0,0000	Not_present	23	0,0143	Rare
OTU_1412	OTU_1412_MK211230_Pseudophormidium_battersii_KZ-16-2_98,4	<i>Pseudophormidium</i>	Microcoleaceae	Oscillatoriales	large_filamentous	Endemic	24	Antarctica	24	0,0015	Rare	0	0,0000	Not_present
OTU_1008	OTU_1008_KY704115_Chamaesiphon_cf_incrustans_str_Ch_fontanille_95,4	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	23	Sub_Antarctica	0	0,0000	Not_present	23	0,0143	Rare
OTU_1239	OTU_1239_MF527165_Uncult_Chroococcidiopsis_96,2	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Endemic	22	Antarctica	22	0,0013	Rare	0	0,0000	Not_present
OTU_1174	OTU_1174_KF704343_Nostoc_sp_Krosnia_crystallifera_LG_M788_99,7	<i>Nostoc</i>	Nostocaceae	Nostocales	heterocytous	Cosmopolitan	22	Antarctica	22	0,0013	Rare	0	0,0000	Not_present
OTU_1095	OTU_1095_MK861868_Drouetiella_sp_V16_99,7	<i>Drouetiella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Cosmopolitan	22	Both	6	0,0004	Rare	16	0,0099	Rare
OTU_1064	OTU_1064_KT983915_Uncult_Chamaesiphon_98,4	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Polar	22	Both	1	0,0001	Rare	21	0,0130	Rare
OTU_1150	OTU_1150_MG652769_Gloeobacter_99,2	<i>Gloeobacter</i>	Gloeobacteraceae	Gloeobacterales	unicellular	Cosmopolitan	21	Antarctica	21	0,0013	Rare	0	0,0000	Not_present
OTU_1201	OTU_1201_KC004019_Aphanothece_caldoriarum_var_cavernarum_97,8	<i>Aphanothece</i>	Aphanothaceae	Chroococcales	unicellular	Endemic	20	Antarctica	20	0,0012	Rare	0	0,0000	Not_present
OTU_1260	OTU_1260_MT488259_Synechococcus_sp_CHAB_TP201820.1_98,6	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	20	Antarctica	20	0,0012	Rare	0	0,0000	Not_present
OTU_1330	OTU_1330_OM732250_Chalicogloea_sp_BACA0589_99,5	<i>Chalicogloea</i>	Chroococcaceae	Chroococcales	unicellular	Cosmopolitan	20	Antarctica	20	0,0012	Rare	0	0,0000	Not_present
OTU_1500	OTU_1500_JQ327934_Uncult_Chroococcidiopsis_98,9	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Endemic	19	Antarctica	19	0,0012	Rare	0	0,0000	Not_present
OTU_1442	OTU_1442_KY498228_Tildenella_torsiva_strain_UHER1998_13D_100	<i>Tildenella</i>	Oculatellaceae	Synechococcales	thin_filamentous	Cosmopolitan	19	Antarctica	19	0,0012	Rare	0	0,0000	Not_present
OTU_172	OTU_172_KY704115_Chamaesiphon_cf_incrustans_str_Ch_fontanille_96,8	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Endemic	19	Sub_Antarctica	0	0,0000	Not_present	19	0,0118	Rare
OTU_1050	OTU_1050_MH688842_Hormosilla_cf_pringsheimii_Us-s-6-2_98,9	<i>Hormosilla</i>	Gomontiellaceae	Oscillatoriales	large_filamentous	Cosmopolitan	18	Antarctica	18	0,0011	Rare	0	0,0000	Not_present
OTU_1334	OTU_1334_MK248010_Synechococcus_sp_ATE729_96,8	<i>Synechococcus</i>	Synechococcaceae	Synechococcales	unicellular	Cosmopolitan	18	Antarctica	18	0,0011	Rare	0	0,0000	Not_present
OTU_1238	OTU_1238_JQ700544_Uncult_Chamaesiphon_97,3	<i>Chamaesiphon</i>	Chamaesiphonaceae	Synechococcales	unicellular	Cosmopolitan	18	Sub_Antarctica	0	0,0000	Not_present	18	0,0112	Rare
OTU_1388	OTU_1388_DQ393595_Pseudanabaena_constantiae_97,6	<i>Pseudanabaena</i>	Pseudanabaenaceae	Synechococcales	thin_filamentous	Cosmopolitan	18	Sub_Antarctica	0	0,0000	Not_present	18	0,0112	Rare
OTU_1196	OTU_1196_OR046508_Chroococcales_cyanobacterium_LEGE_1760_7_99,5	unknow	unknow	Chroococcales	unicellular	Cosmopolitan	15	Both	1	0,0001	Rare	14	0,0087	Rare
OTU_1387	OTU_1387_MN152980_Stenomitrospira_sp_HA6792-KK3_99,2	<i>Stenomitrospira</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Cosmopolitan	16	Both	1	0,0001	Rare	15	0,0093	Rare
OTU_1346	OTU_1346_LC103290_uncult_bact_97,3	<i>uncult_bact</i>	unknow	unknow	unknow	Endemic	15	Antarctica	15	0,0009	Rare	0	0,0000	Not_present
OTU_1318	OTU_1318_MH118734_Pseudanabaena_frigida_ULC067_99,5	<i>Pseudanabaena</i>	Pseudanabaenaceae	Synechococcales	thin_filamentous	Cosmopolitan	15	Antarctica	15	0,0009	Rare	0	0,0000	Not_present
OTU_1483	OTU_1483_MF527193_Uncult_Chroococcidiopsis_99,7	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Cosmopolitan	15	Antarctica	15	0,0009	Rare	0	0,0000	Not_present
OTU_44	OTU_44_MN796298_Snowella_sp_CHAB6606_100	<i>Snowella</i>	Coelosphaeriaceae	Synechococcales	unicellular	Cosmopolitan	14	Both	12	0,0007	Rare	2	0,0012	Rare
OTU_1441	OTU_1441_OQ509454_Uncult_Chroococcidiopsis_sp_clone_8_98,6	<i>Chroococcidiopsis</i>	Chroococcidiopsidaceae	Chroococcidiopsidales	baeocytous	Endemic	14	Antarctica	14	0,0009	Rare	0	0,0000	Not_present
OTU_1488	OTU_1488_KU309579_Leptolyngbya_boryana_CCALA_082_97,5	<i>Leptolyngbya</i>	Leptolyngbyaceae	Synechococcales	thin_filamentous	Endemic	13	Antarctica	13	0,0008	Rare	0	0,0000	Not_present
OTU_804	OTU_804_KF264594_Placoma_regulare_UCFM_PR_99	<i>Placoma</i>	Chroococcaceae	Chroococcales	unicellular	Cosmopolitan	13	Both	12	0,0007	Rare	1	0,0006	Rare
OTU_96	OTU_96_GU935356_Pseudanabaena_sp_Sai010_100	<i>Pseudanabaena</i>	Pseudanabaenaceae	Synechococcales	thin_filamentous	Cosmopolitan	11	Both	10	0,0006	Rare	1	0,0006	Rare
OTU_1433	OTU_1433_FR798926_Chroococcus_sp_2T05h_99,7	<i>Chroococcus</i>	Chroococcaceae	Chroococcales	unicellular	Cosmopolitan	10	Antarctica	10	0,0006	Rare	0	0,0000	Not_present

Table S2.3. PERMANOVA analysis between the 2 bioregions of Antarctica (Maritime and Continental Antarctica) and the sub-Antarctic islands.

	Df	SumsOfSqs	MeanSqs	F.Model	R2	p value	p adjusted	Signif
continental vs maritime	1	0,8855031/		2,592106	0,03429069	0,006	0,018.	
continental vs sub-Antarctic	1	2,4296007/		6,978071	0,08116106	0,001	0,003*	
maritime vs sub-Antarctic	1	0,9883339/		2,740822	0,12052432	0,002	0,006*	

Table S2.4. Variation partitioning showing the relative importance of spatial (PCNMs) and environmental (pH and conductivity) parameters on the structure of cyanobacterial communities.

Resp. variable (model type)	Env. Variables	Spatial filters (PCNMs)	Fraction of variation as partial adj. R ²				
			[E]	[S]	[E S]	[S E]	[E+S]
Antarctica + Sub-Ant. Islands (71 samples)	Cond + pH	1, 2, 3, 4, 6, 9	4.9	14.8	4.0	13.8	18.7

Table S2.5 Number of OTUs, proportion of OTUs (%), number of reads (#) and proportion of reads (% #) classed by geographic distribution (cosmopolitan, alpine, polar and endemic) for the different categories of OTUs.

Group	Cosmo politan				Alpine				Polar				Endemic			
	(%)	#	(% #)		(%)	#	(% #)		(%)	#	(% #)		(%)	#	(% #)	
Total (310 OTUs)	142	45,8	986905	54,7	27	8,7	619124	34,3	15	4,8	27354	1,5	126	40,6	169834	9,4
Antarctica (279 OTUs)	133	47,7	850165	51,8	25	9	615696	37,5	15	5,4	26629	1,6	106	38	149545	9,1
sub-Antarctic islands (177 OTUs)	92	52	136740	84,8	18	10,2	3428	2,1	7	4	725	0,4	60	33,9	20289	12,6
Abundant Antarctica (31 OTUs)	20	64,5	755126	51,4	6	19,4	597607	40,7	1	3,2	23506	1,6	4	12,9	92702	6,3
Abundant sub-Antarctic islands (20 OTUs)	14	70	130535	90,1	1	5	1219	0,8	0	0	0	0	5	25	13173	9,1
Rare Antarctica (162 OTUs)	70	43,2	6501	39,2	9	5,6	1198	7,2	11	6,8	1296	7,8	72	44,4	7596	45,8
Rare sub-Antarctic islands (109 OTUs)	58	53,2	685	41	12	11	176	10,5	6	5,5	52	3,1	33	30,3	758	45,4

Chapter 3

From single-family dominance to balanced assemblages: substrate-driven soil cyanobacterial communities in the Sør Rondane Mountains, East Antarctica

Valentina Savaglia^{1,2}, Benoit Durieu¹, Bjorn Tytgat², Marie Harmel¹, Josef Elster³, Wim Vyverman², Elie Verleyen², Annick Wilmotte¹

¹InBioS, University of Liège, 4000 Liège, Belgium

²Laboratory of Protistology & Aquatic Ecology, Ghent University, 9000 Ghent, Belgium

³Centre for Polar Ecology, Faculty of Science, University of South Bohemia, České Budějovice and Institute of Botany, AS CR, Třeboň, Czechia

Manuscript in preparation for submission to **FEMS Microbiology Ecology**

Author contribution of BD:

BD performed DNA extractions, PCR-amplifications, libraries preparation, bioinformatic analyses, a part of the statistical analyses (mock community and replicates analysis) and helped in writing the manuscript. BD and VS equally contributed to the work.

3.1. Abstract

Cyanobacteria are key primary producers in Antarctic terrestrial ecosystems, yet little is known about their diversity and microhabitat distribution across soils derived from different parent rock types. Here, we investigated cyanobacterial community structure across surface soils in the western Sør Rondane Mountains, East Antarctica. We sampled 81 samples of topsoils ranging from bare gravel to biocrust-covered surfaces across four distinct substrate types: marble, granite, gneiss, and moraines. Using cyanobacteria-specific primers targeting the V3–V4 region of the 16S rRNA gene and updated taxonomic assignments via the CyanoSeq database, we evaluated community patterns alongside soil chemical and physical properties. Ordination analyses revealed that these different soil substrates harbored distinct cyanobacterial communities, with variations in community structure significantly driven by soil pH, total nitrogen, nitrates, and electrical conductivity. While marble exhibited the highest overall ASV richness, granitic soils hosted remarkably balanced and taxonomically diverse assemblages co-dominated by filamentous (Gomontiellaceae, Oculatellaceae, Microcoleaceae, Leptolyngbyaceae), unicellular (Cyanothecaceae), and heterocystous (*Nostocaceae*) taxa. In contrast, other substrates exhibited strong single-family/genus dominance. Hyper-arid Dry Valley moraines were overwhelmingly dominated by desiccation-tolerant Gomontiellaceae (*Hormoscilla* sp.), whereas Yûboku Valley moraines shifted toward non-heterocystous filamentous families (Oculatellaceae and Leptolyngbyaceae). Similarly, marble bedrock was heavily dominated by Chroococcidiopsidaceae (*Aliterella* sp.) and gneiss bedrock by Cyanothecaceae (*Cyanotheca* sp.). A substantial proportion of sequences remained unclassified at the family or genus level, emphasizing the need for improved Antarctic representation in reference databases. Overall, our study demonstrates that lithology and microenvironmental heterogeneity shape polar cyanobacterial assembly, providing an essential biodiversity baseline to support future conservation and monitoring within Antarctic Specially Protected Area No. 179 (ASPA 179).

3.2. Introduction

Cyanobacteria occur worldwide in a variety of terrestrial and aquatic ecosystems (Vincent 2009), including extreme environments, such as hot springs, barren rocks and desert soils (Whitton and Potts, 2000). The cold deserts of Antarctica are among the most extreme biotopes on Earth (Brooks et al. 2019), characterized by sub-zero temperatures, frequent freeze-thaw cycles, extreme photosynthetically active radiation (PAR) and ultra-violet radiation (UVR), and severe

nutrient limitation (Vincent 2000). In these ice-free terrestrial habitats – characterized by exposed rocks and oligotrophic mineral soils (Ugolini and Bockheim 2008, Lee et al. 2017) derived from bedrock weathering – cyanobacteria are often the dominant photoautotrophs, including some taxa also capable of nitrogen fixation (Broady, 2005; Pointing et al., 2015; Savaglia et al., 2024; Pushkareva et al., 2024). By consequence, cyanobacteria act as true ecosystem engineers in Antarctic terrestrial habitats (Christmas et al., 2018). Despite their ecological importance, the diversity and environmental drivers shaping the community structure of terrestrial cyanobacteria remains far less understood than those in Antarctic lacustrine habitats (e.g. Vincent et al., 1993; Taton et al., 2003; 2006a; 2006b; de los Ríos et al., 2004; Jungblut et al., 2005; Wood et al., 2008; Zhang et al., 2015; Pessi et al., 2018; Durieu et al., 2025).

In sheltered and more humid micro-habitats, cyanobacteria are free-living or form lichen symbioses, facilitating the development of biological soil crusts (biocrusts; Weber et al., 2022). Instead, under hyper-arid conditions – such as the moraine soils of the McMurdo Dry Valleys – cyanobacteria may be absent or restricted to few, often unicellular taxa (Pointing et al., 2015), occupying endolithic, chasmoendolithic, or hypolithic niches within or beneath the lithic matrix (Chan et al., 2012). Survival in these extreme niches relies on physiological and biochemical adaptations, including the production of exopolysaccharides (EPS), mycosporine-like amino acids (MAAs), and protective pigments such as scytonemin and carotenoids (Tamaru et al., 2005; Effendi et al., 2022; Dextro et al., 2023; Roncero-Ramos et al., 2024; Van Goethem et al., 2025), as well as vertical motility in filamentous biocrust species to optimize light exposure (Bebout and Garcia-Pichel, 1995).

Most of our knowledge regarding Antarctic cyanobacterial communities relies on molecular surveys using general bacterial 16S rRNA gene primers (but see Taton et al., 2003; 2006a; 2006b; de los Ríos et al., 2004; Jungblut et al., 2005; Wood et al., 2008; Zhang et al., 2015). In fact, this also holds true for cyanobacteria studies worldwide. Because general primers preferentially amplify dominant bacteria, cyanobacterial sequences often represent a minor fraction of the dataset, especially in hyper-arid regions, limiting taxonomic depth of rare or low-abundant groups. Furthermore, standard taxonomic databases (e.g. SILVA or Greengenes) rely on outdated classifications that fail to reflect recent revisions in cyanobacterial polyphasic taxonomy. These limitations can be overcome by combining the use of cyanobacterial-specific primers (Nübel et al., 1997), with modern, curated databases such as CyanoSeq (Lefler et al., 2023) which provides updated, high-resolution cyanobacterial taxonomic assignments (Strunecký et al., 2023).

In the western Sør Rondane Mountains (Dronning Maud Land, East Antarctica), ice-free nunataks and moraine valleys exhibit distinct geological histories, bedrock mineralogies, and

microclimates (Tytgat et al., 2016). A recent broad-scale bacterial survey of this region using general 16S primers revealed that substrate type strongly influences bacterial community structure: heterotrophic Actinomycetota dominated dry moraine soils, whereas Cyanobacteriota dominated granitic and marble bedrocks, with distinct candidate keystone taxa such as *Aliterella* sp. in marble and *Nostoc* sp. in granite (Savaglia et al., 2024). However, targeted cyanobacterial surveys in this region remain scarce. While a pilot study previously used cyanobacteria-specific primers across a few nunataks (Pushkareva et al., 2018), fine-scale community structure and substrate-specific indicator taxa encompassing the full spectrum of local soil types remain unresolved.

Here, we performed an in-depth study of the cyanobacterial communities across 81 soil samples from the western Sør Rondane Mountains (East Antarctica) using cyanobacteria-specific 16S rRNA primers, the CyanoSeq database, and the delineation of taxa as amplicon sequence variants (ASVs) through a robust bioinformatic workflow (RASPAM; Harmel et al., 2025), as well as a comprehensive dataset of soil abiotic variables (TOC, EC, TN, TP, PO_4^{3-} , NO_3^- , NH_4^+ , and soil dry weight as a proxy of water content). We hypothesized that targeted cyanobacterial amplification combined with CyanoSeq taxonomic curation would uncover significantly higher taxonomic richness, reveal rare biosphere members, and resolve polyphyletic lineages previously obscured by general bacterial surveys (methodological hypothesis) and that substrate mineralogy and micro-topology would structure the cyanobacterial diversity and structures, with more moisture-retentive, less porous granitic soils supporting diverse, filamentous, atmospheric N_2 fixing taxa (e.g. Nostocaceae), and hyper-arid, alkaline marble and moraine soils supporting more specialized, stress-tolerant, unicellular taxa (e.g. Chroococcidiopsaceae; ecological hypothesis).

3.3. Materials & Methods

3.3.1. Sampling and site description

Detailed information regarding sampling locations and DNA extractions was described in a previous study (Savaglia et al., 2024). Briefly, the samples were collected during two expeditions in 2018 and 2019. They came from 5 different nunataks, 1 ridge and 3 valleys encompassing 4 different kinds of substrates (**Fig. 3.1**).

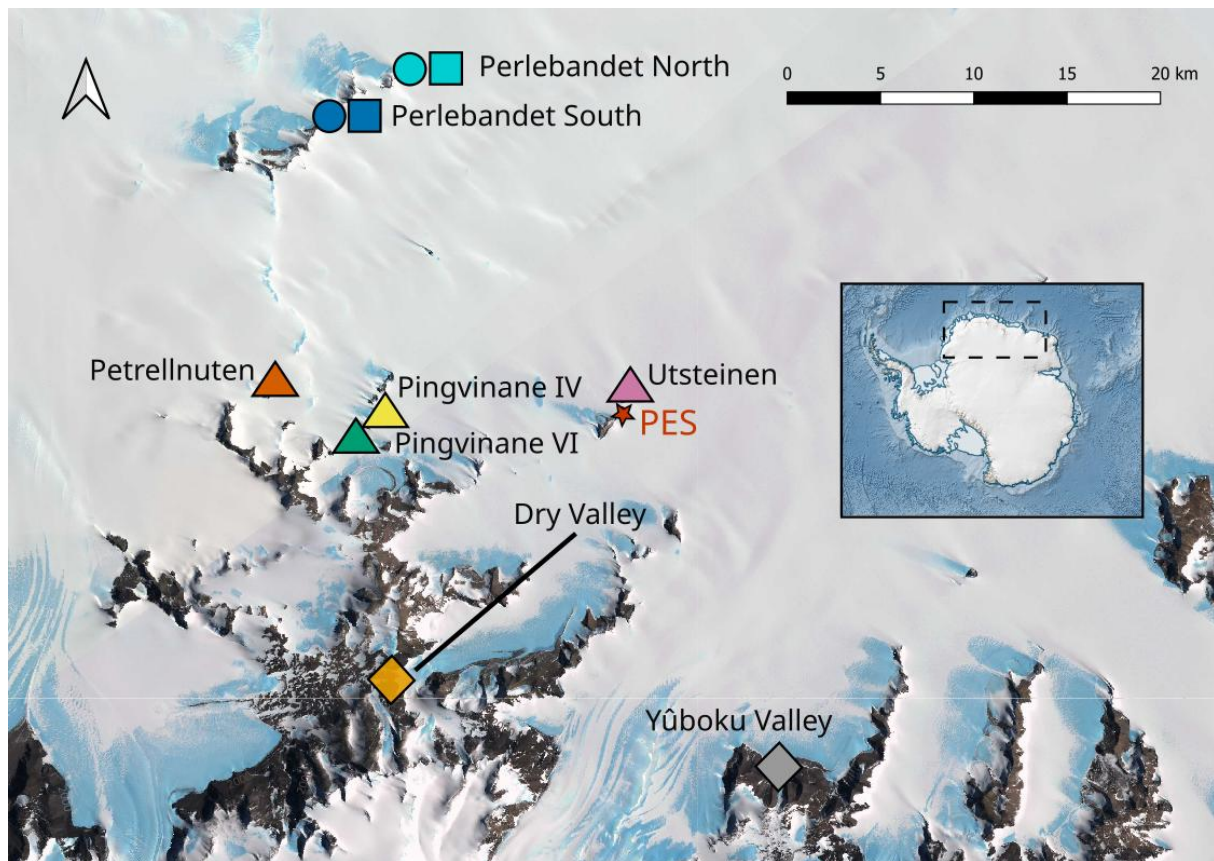


Fig. 3.1. Map of the Sør Rondane Mountains showing sampled sites. The region is indicated by the dashed box in the inset map of Antarctica. Colors represent different sampling sites, while symbols indicate substrate types: circles (gneiss), triangles (granite), squares (marble), and diamonds (moraine). The orange star highlights Princess Elisabeth Station (PES). Detailed sample information is provided in **Table S3.1**. Contains modified Copernicus Sentinel-2 data (2020) and Landsat Image Mosaic of Antarctica (LIMA).

3.3.2. DNA extraction and sequencing

High-quality cyanobacterial 16S rRNA gene libraries were successfully constructed for 81 soil samples. Each sample was first thoroughly homogenized and between 0.2 to 1g was used for DNA extraction based on the amount of biomass or mineral sample required with the Qiagen DNeasy Powersoil kit (Valencia, CA). PCRs targeting the V3-V4 hypervariable region of the 16S rRNA gene were performed using the cyanobacteria-specific primers CYA359F and CYA781Rb (Nübel et al., 1997) following the protocol described by Pessi et al. (2016). The PCR mixture consisted of 2.5µl buffer with 2.5µl BSA (1 mg mL⁻¹), 0.5µl dNTP (200µM), 1.25µl of each primer (0.2µM), 0.2 µl of 1 U µl⁻¹ SUPER TAQ plus DNA polymerase (HT Biotechnology, Cambridge, UK), 2µl template DNA and completed with filtered water (Sigma-Aldrich) to reach a final volume of 25µl. PCR amplification was carried out with an initial denaturation step at 94 °C for 5 min, followed by 37 cycles of 94 °C for 1 min, 1 min annealing at 57 °C and 1 min at 72 °C and an elongation step at 72

°C for 5 min. After a final elongation of 20 min at 72 °C, PCR products were visualized on TBE 0.8% agarose gel. If amplification was successful, PCR duplicates for each sample were pooled to minimize potential biases. The pooled PCR products were subsequently purified (NucleoSpin® Gel and PCR Clean-up, Macherey-Nagel GmbH & Co, Düren, Germany) according to the manufacturer's instructions. The DNA concentration of the cleaned PCR products was measured using Quant-iT PicoGreen dsDNA Assay Kit, Life Technologies, Carlsbad, CA, USA), and the samples were equimolarly pooled based on their molecular weight and DNA concentrations of at least 6 ng μl^{-1} for a total quantity of 200ng for each sample. Amplicon libraries were sequenced on an Illumina Miseq platform (2×300 bp paired-end, v3 chemistry) at Genewiz (Azenta Life Sciences, Leipzig, Germany) where ligation of sequencing adapters to the amplicons and sequencing were performed. Each run contained 2 replicates of a mock community containing 6 cyanobacterial strains of different species (**Table S3.3**), plus each of the latter strains separately in the same concentration used for the mock community, in order to check the effective quantity of each strain of the mock community), 2 negative controls (each composed by a mixture of PCR blanks randomly selected) to check the overall run quality and benchmark processing variables, as well as 24 randomly chosen PCR replicate samples.

3.3.3. Data processing

Amplicon sequencing data targeting the V3-4 region of the 16S rRNA gene were processed using the RASPAM workflow (Harmel et al., 2025) in ASV mode. Analyses were performed with DADA2 v1.22.0 (Callahan et al., 2016) within the R version 4.5.2 (R Core Team, 2021), using the core functions `plotQualityProfile`, `filterAndTrim`, `learnErrors`, `derepFastq`, `dada`, `mergePairs`, and `assignTaxonomy`. Quality profiles were inspected to determine optimal trimming and filtering parameters, resulting in the use of `truncLenF` = 290, `truncLenR` = 240, `trimLeftF` = 5, `trimLeftR` = 5, `maxEE` = 2 (forward) and 5 (reverse), a minimum overlap of 12 bp, and zero mismatches allowed during merging. Chimera removal was performed using the consensus method implemented in DADA2 (Callahan et al., 2016). ASV tables were further filtered according to abundance thresholds, applying both min occurrences per ASV (across samples) = 2 and min occurrence of ASV in the sample = 1000, to retain high-confidence sequences. Taxonomic assignment of the resulting ASVs was performed using CyanoSeq v1.3 with added species-level annotations database (Lefler et al. 2023) `minBoot`=80 parameter. No lineage-based filtering was applied to preserve the full taxonomic diversity.

3.3.4. Environmental data collection

Environmental factors including elevation, pH (H₂O), soil total organic carbon (TOC, %), soil dry weight (%), total phosphorus (TP, %) total nitrogen (TN, %), electric conductivity (EC, $\mu\text{S cm}^{-1}$), and concentrations of N-NO₃⁻ (mg kg⁻¹), N-NH₄⁺ (mg kg⁻¹), and P-PO₄³⁻ (mg kg⁻¹), were conducted following established methods at the Institute of Botany, AS CR in Třeboň (Czech Republic) as described and summarized in Savaglia et al. (2024).

3.3.5. Statistical analysis

All statistical calculations were conducted in R 4.5.2 (R Core Team 2021).

ASVs represented by only one sequence per sample, mitochondrial, chloroplast-related sequences and non-Cyanobacteriota ASVs were removed. After inspection of the mock community data (**Table S3.2**), blanks and technical replicates were removed from the ASV table, retaining only the replicate with the highest read depth for downstream analyses (**Fig. S3.1**).

To assess differences in mean local alpha-diversity (i.e. ASV richness) between sites and substrate types, a negative binomial generalized linear model was applied following Tytgat et al. (2023), accounting for differences in sequencing depth (log-transformed) and community evenness (defined as the Shannon index divided by the natural logarithm of richness) across samples. To evaluate consistency with general bacterial profiling, these diversity estimates were compared against cyanobacterial sequences extracted from the 16S rRNA gene dataset generated with V1–V3 universal bacterial primers in our earlier study (Savaglia et al., 2024). Only samples present in both datasets were included in comparative analyses (n = 80). Due to non-normality and primer-specific differences in absolute estimates, concordance between datasets was evaluated using Spearman's rank correlation via the `cor.test()` function and visualized with `corrplot` (v0.95; **Fig. S3.2**). Log-transformed richness values were used for visualization purposes only.

To evaluate differences in community structure across samples, beta-diversity analyses were conducted. The ASV table was first Hellinger-transformed (i.e., square root of relative abundances) to down-weight highly abundant taxa. Community composition patterns were visualized using Principal Coordinates Analysis (PCoA) and Non-metric Multidimensional Scaling (nMDS) ordinations generated with the `ordinate()` function in the `phyloseq` package (v1.54.2; McMurdie and Holmes, 2013).

Phylotypes (ASVs) significantly associated with specific substrate types were identified using the Indicator Value index (IndVal) via the `multipatt()` function in the `indicspecies` package (version 1.8; De Cáceres and Legendre, 2009). The calculated indicator value statistic (“stat” column, ranging from 0 to 1) measures the strength of association between an ASV and a substrate group, combining specificity (exclusivity to a given substrate) and fidelity (frequency of occurrence within that substrate). Statistical significance was evaluated using 999 permutations, with taxa considered significant indicators at $p \leq 0.05$.

Diagnostic plots, alongside skewness and kurtosis metrics (moments package v0.14; Komsta and Novomestky, 2015), were evaluated to determine necessary data transformations for environmental variables. Normality and homoscedasticity were formally assessed using Shapiro–Wilk and Bartlett’s tests (stats package; R Core Team, 2021). To attenuate right-skewness, TOC, N-NO₃⁻, N-NH₄⁺, TN, TP, dry weight, and electrical conductivity were $\log(x+1)$ transformed prior to analysis. Collinearity among transformed parameters was evaluated using pairwise Pearson correlation coefficients ($|r|$) and stepwise Variance Inflation Factors (VIF). Pearson correlations were calculated using the base R `cor()` function and visualized with `corrplot` (v0.95; Wei and Simko, 2021; **Fig. S3.4**). Stepwise VIF values were calculated using the `vifstep()` function in the `usdm` package (v2.1-7; Naimi et al., 2014). No severe collinearity was detected ($|r| < 0.6$ and “VIF” < 5); thus, all environmental variables were retained for downstream analyses. Non-parametric Kruskal–Wallis rank-sum tests (`kruskal.test()`) were performed to evaluate differences in geochemical properties across substrate types (**Table S3.3**).

The influence of spatial structure across multiple scales on cyanobacterial community composition was tested using principal coordinates of neighbour matrices (PCNM) following Borcard and Legendre (2002) and Sakaeva et al. (2016). First, sample coordinates were used to construct a pairwise geographic distance matrix using the `geosphere` R package (v1.5-14; Hijmans, 2021). PCNM eigenvectors were then computed using the `pcnm()` function in the `vegan` package (v2.5-7; Oksanen et al., 2020). PCNMs significantly correlated with environmental parameters ($p \leq 0.05$) were identified. From this subset, stepwise linear model selection (Venables and Ripley, 2002; Sokol et al., 2013) was performed to select the most parsimonious set of spatial predictors linking environmental and substrate variables to spatial structure.

To assess the proportion of community variation explained by specific drivers, distance-based redundancy analysis (db-RDA; McArdle and Anderson, 2001) was performed using a Bray–Curtis dissimilarity matrix (with square-root correction to avoid negative eigenvalues). Community variation was modelled as a function of environmental parameters (pH, N-NO₃⁻, N-NH₄⁺, TN, TP, dry weight, electrical conductivity), spatial variables (selected PCNMs), and substrate types

(gneiss, granite, marble, moraine). Forward stepwise model selection (Blanchet et al., 2008) based on adjusted R² values and alpha significance thresholds (Peres-Neto et al., 2006) was applied to identify the optimal set of non-redundant variables that best explained variations in community structure.

3.4. Results

3.4.1. Community structure and diversity in relation to substrate and sampling sites

Illumina sequencing generated 2,697,299 raw reads. The mock community analysis showed consistent taxonomic composition and relative abundances between the mock community and individual strain samples (**Table S3.2**). Replicate samples have similar taxonomic profiles, and differences between samples is significantly higher than between replicates of the same samples (**Fig. S3.1**). Following bioinformatic processing, quality filtering, and the removal of replicate samples and non-target sequences, a total of 1,088,899 high-quality cyanobacterial reads (representing ~40% of initial reads across 81 samples) were retained for downstream analyses. These reads clustered into 295 Cyanobacteriota ASVs. The community was predominantly composed of Gomontiellaceae (23%), Chroococcidiopsidaceae (18%), Cyanothecaceae (18%), Oculatellaceae (11%), and Leptolyngbyaceae (9%). Unclassified cyanobacterial families accounted for 8% of total reads, while Microcoleaceae, Nostocaceae, Chamaesiphonaceae, and Tolypothrichaceae represented 5%, 4%, 2%, and 1%, respectively. The remaining families each constituted less than 1% of total reads, including Prochlorococcaceae, Cymatolegaceae, Microcystaceae, Oscillatoriaceae, Aphanizomenonaceae, Wilmottiaceae, Stigonemataceae, Pseudanabaenaceae, and Pleurocapsaceae (**Table S3.4**).

The relative abundance of cyanobacterial families varied markedly across substrate types (**Fig. 3.2, Tables S3.4 and S3.5**). Gomontiellaceae was the most abundant family overall (23% of total reads) and dominated the moraine Dry Valley samples, where it comprised 75% of all reads. Gomontiellaceae was also prominent in granite bedrock, representing 17% of granite reads. Chroococcidiopsidaceae was the second most abundant family overall (18% of total reads); the vast majority of its sequences (86%) were recovered from marble, where it formed the dominant taxon (63% of marble reads). Cyanothecaceae constituted 18% of total reads and occurred across all substrates; however, over half of all Cyanothecaceae sequences (51%) originated from

gneiss, where the family accounted for 81% of the community. Oculatellaceae (11%) and Leptolyngbyaceae (9%) exhibited similar spatial distributions, being primarily restricted to granite (accounting for 42% and 43% of their respective total reads) and moraine Yûboku Valley (47% and 39%, respectively), while remaining minor components (<3% each) in marble, gneiss, and moraine Dry Valley.

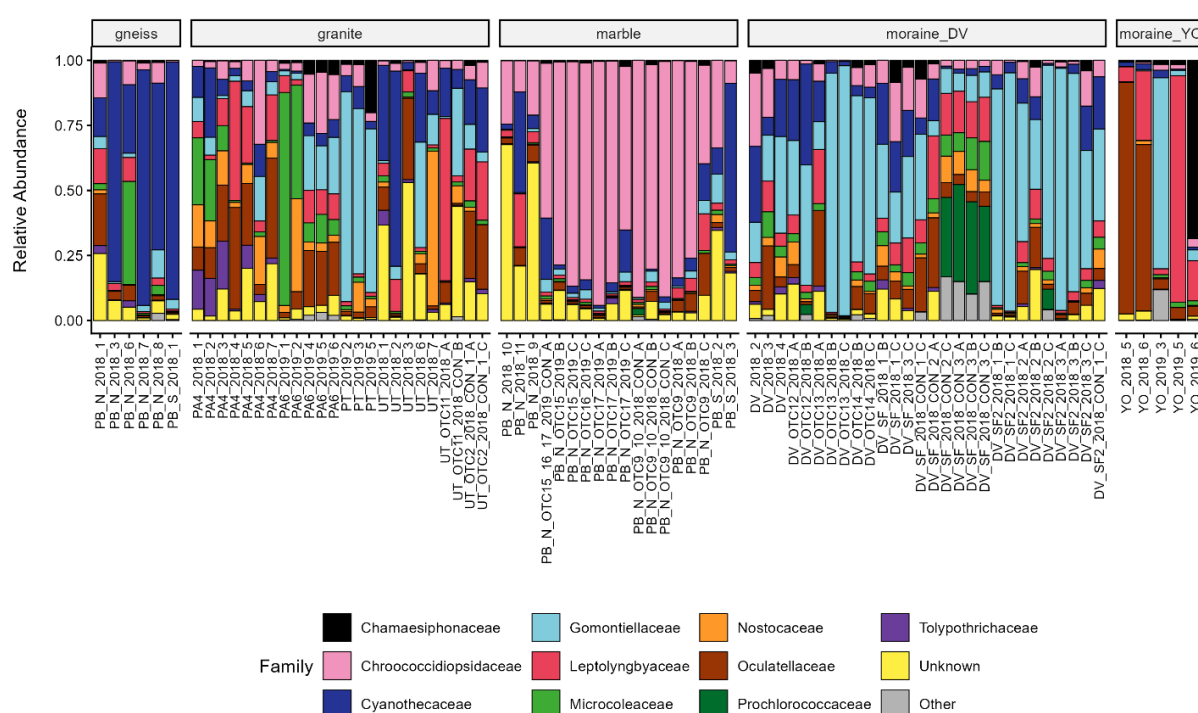


Fig. 3.2. Relative abundance of cyanobacterial families per sample, grouped by substrate type. Samples are grouped by substrate (gneiss, granite, marble, Dry Valley moraine [moraine_DV], and Yûboku Valley moraine [moraine_YO]) and arranged alphabetically by sample ID. Colors represent individual cyanobacterial families (see legend); low-abundance taxa are combined into "Other", and unassigned sequences are labeled as "Unknown".

Significant indicator taxa were identified across all substrate types except for moraine Dry Valley (**Table S3.6**). Marble substrate was strongly characterized by indicator ASVs belonging to *Aliterella* (ASV_3, stat = 0.919, $p = 0.001$) and Chroococcidiopsidaceae (ASV_208 and ASV_190; $p \leq 0.016$). Gneiss substrate was distinguished by indicator ASVs including *Cyanothece* (ASV_7, stat = 0.833, $p = 0.001$) and an unclassified Cyanobacteriota phylotype (ASV_285, stat = 0.960, $p = 0.001$). Granitic substrate exhibited specific indicator ASVs belonging to *Nostoc* (ASV_13, stat = 0.875, $p = 0.002$), *Microcoleaceae* (ASV_16), and *Toxopsis* (ASV_141). Finally, Moraine Yûboku

Valley harbored a diverse suite of indicator taxa, led by Leptolyngbyaceae (ASV_6, stat = 0.973, $p = 0.001$), Oculatellaceae (ASV_9, stat = 0.919, $p = 0.009$), Timaviella (ASV_29), Shackletoniella (ASV_108), Chamaesiphon (ASV_72).

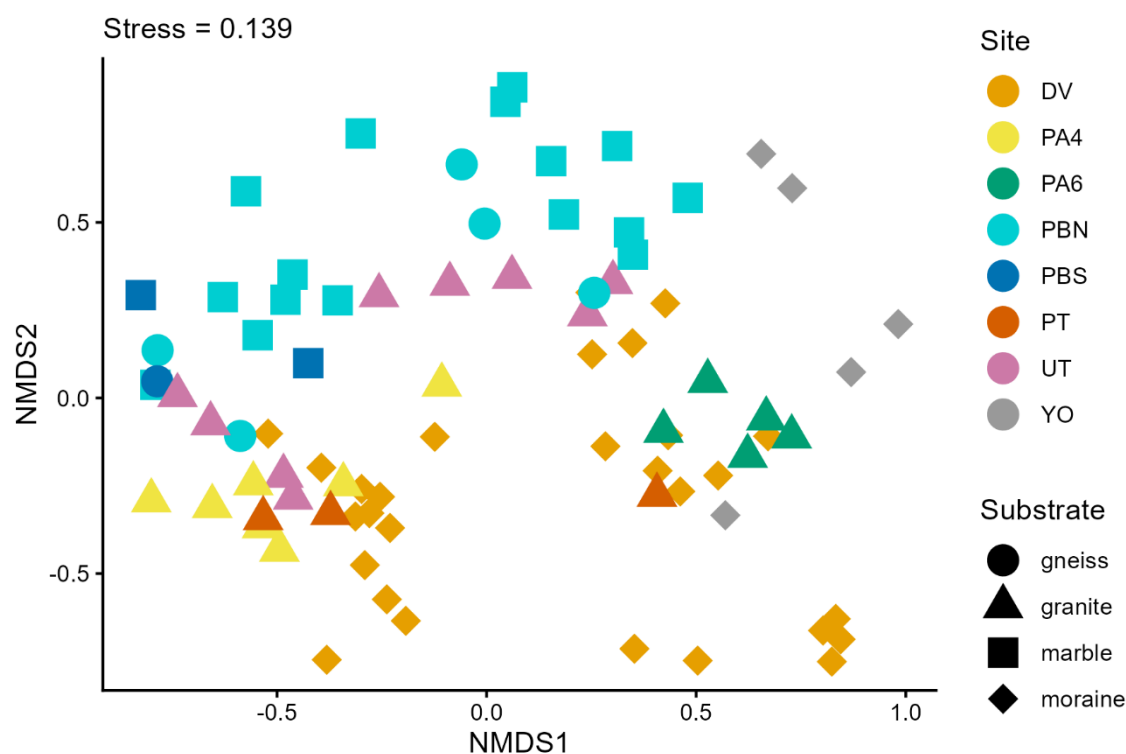


Fig. 3.3. Non-metric multidimensional scaling (nMDS) plot of 81 samples of cyanobacterial communities. Circles represent gneiss, triangles granite, squares marble and diamonds moraine samples. Colors are according to the different kinds of substrate.

Ordination of Hellinger-transformed Bray–Curtis dissimilarity matrices (**Fig. 3.3** and **Fig. S3.5**) revealed that microbial community structure was driven by an interplay between substrate type and geographic site. Marble samples from different nunataks (PBN and PBS) consistently formed a distinct, cohesive cluster in upper ordination space, demonstrating strong substrate-level grouping regardless of location. In contrast, granitic samples exhibited clear site-dependent structuring: while granite communities from PA4, UT, and PT clustered tightly together, granite samples from PA6 separated completely and grouped alongside gneiss and Dry Valley moraine samples. Furthermore, moraine communities varied spatially, with Yûboku Valley (YO) samples occupying a distinct, isolated position on the periphery of the ordination space compared to Dry Valley moraines.

Cyanobacterial ASV richness differed significantly across the majority of both substrate types and sampling sites (**Fig. S3.2** and **S3.3; Table S3.6**). At the substrate level, marble bedrock harbored the highest ASV richness, which was significantly greater than that of Dry Valley moraines ($p < 0.05$) and granitic soils. Gneiss also supported significantly higher richness than moraine soils of Dry Valley, whereas Yûboku Valley moraines exhibited intermediate richness. Site-level richness closely mirrored these substrate patterns: PBN – which alongside PBS, hosts all marble and gneiss samples – displayed significantly higher richness than DV, PA4, and UT ($p < 0.05$). Richness across all remaining sites (PBS, PA6, PT, and YO) was intermediate and did not differ significantly from either group.

To evaluate whether cyanobacterial community patterns were influenced by sequencing strategy, we compared alpha diversity metrics derived from targeted cyanobacterial profiling (16S rRNA gene V3–V4) against general bacterial profiling (16S rRNA gene V1–V3; **Fig. S3.7**) from the exact sample set ($n = 80$). Both approaches demonstrated strong, statistically significant positive correlations for ASV richness (Spearman $\rho = 0.63$, $p < 0.001$) and community evenness (Spearman $\rho = 0.59$, $p < 0.001$). Notably, cyanobacterial-specific primers consistently recovered higher community evenness whilst ASV richness was similar across samples. These results confirm that the use of these primers provides a more valid and sensitive approach for accurately capturing cyanobacterial community structure and distribution, as they potentially avoid amplification bias from dominant general bacteria and better recover low-abundance cyanobacterial taxa.

3.4.2. Community structure in relation to environmental factors, site and substrate

Distance-based redundancy analysis (db-RDA) of Hellinger-transformed Bray-Curtis dissimilarity matrices demonstrated that environmental factors, spatial variables (PCNMs), and substrate type all significantly structured cyanobacterial community composition ($p = 0.002$ for each category; **Table S3.7**). Forward stepwise variable selection revealed that the best-fit environmental model (encompassing pH, TN, N-NO_3^- and conductivity) accounted for 26% of the total community variation, with pH (13%) and total nitrogen (TN) exerting the strongest influences (18% combined). Spatial structure explained 32% of total variation across seven significant spatial eigenvectors (PCNM1-4 and PCNM8-10). Large-scale spatial patterns (PCNM1-3) dominated this spatial component, driving 22% of total community variance, whereas fine-scale vectors (PCNM8-10) accounted for localized variation. Finally, substrate type explained 31% of total community variation across four significant categories (marble, Moraine Yûboku Valley, Moraine Dry Valley,

and granite), with marble and Moraine Yûboku Valley together accounting for the vast majority (25%) of substrate-driven variance.

3.5. Discussion

Our findings demonstrate that terrestrial cyanobacterial communities in the Sør Rondane Mountains largely differ among the investigated substrate types and, in some cases, also at the sampling site level, at different taxonomy resolution. While substrate geochemistry (such as pH, total nitrogen, and phosphate) exerts a strong baseline filter on community composition and richness, local microenvironmental factors – including exposure to katabatic winds, biocrust development, and local moisture conditions also influenced by the porosity of each substrate – further structure these assemblages.

The application of the recently released CyanoSeq database greatly enhanced our taxonomic resolution, successfully assigning order- and family-level classifications to the vast majority of encountered ASVs, linking them with the modern evolving taxonomic framework (Strunecký et al., 2023). Furthermore, comparison with general bacterial profiling (Savaglia et al., 2024) confirmed that cyanobacterial diversity patterns were methodologically robust across primer sets (**Fig. S3.7**). Nonetheless, the higher community evenness recovered by cyanobacteria-specific primers indicates that this targeted approach provides a more valid and sensitive representation of community structure, likely by avoiding amplification bias from dominant general bacteria and better capturing low-abundance cyanobacterial taxa.

3.5.1. Chroococcidiopsaceae dominates marble samples

Marble samples exhibited strong substrate specificity, consistently forming a distinct community cluster across sites (PBN and PBS) and harboring a high dominance of the unicellular family Chroococcidiopsidaceae, primarily represented by the genus *Aliterella*. Although marble soils harbored the highest ASV richness (**Fig. S3.3**), this diversity was heavily skewed toward a single family (Chroococcidiopsidaceae, up to 74%). Rather than a broad, multi-family assemblage, marble diversity is characterized by strong microdiversity within the genus *Aliterella*, where multiple specialized sequence variants co-occur to exploit the alkaline chasmolithic microhabitat. Members of Chroococcidiopsidaceae are well-known endolithic and hypolithic keystones in polar and hot deserts, particularly in lithic substrates that filter harmful UV rays while permitting photosynthetically active radiation (PAR) to reach the biota (Broady, 2005; Pointing et

al., 2009; Chan et al., 2012; Jung et al., 2021; Van Goethem et al., 2025). Our indicator species analysis confirmed several *Aliterella* ASVs as strong, exclusive indicators for marble bedrock (**Table S3.6**), corroborating co-occurrence network analyses from general bacterial surveys where *Aliterella* was recognized as a putative keystone taxon (Savaglia et al., 2024). Furthermore, NCBI BLAST analysis revealed that unclassified Chroococcidiopsidaceae indicators (ASV_208) matched uncultured Antarctic lithic sequences (e.g. Pointing et al., 2009) at high identity (99.21%). Although *Chroococcidiopsis* lineages are conventionally associated with quartz hypoliths, the translucency of coarse marble calcite grains combined with the elevated pH of carbonate rock provides an ideal microhabitat, allowing light penetration while offering shelter from extreme surface conditions.

This strong association is likely driven by the high alkalinity and translucency of marble soils, a pattern consistent with urban studies on marble statues that also observed a predominance of *Aliterella* (Timoncini et al., 2022). Members of *Aliterella* withstand severe desiccation and osmotic stress in fluctuating environments through extracellular polymeric substance (EPS) production, trehalose, and sucrose accumulation (Azua-Bustos et al., 2014; Van Goethem et al., 2025). First described by Rigonato et al. (2016), *Aliterella* encompasses both aquatic and terrestrial species – such as *A. chasmolithica* from the Atacama Desert – yet they could in fact represent different genera (Rigonato et al., 2016; Jung et al., 2020 and 2021). To date, only eight species have been formally cultivated or proposed worldwide (Rigonato et al., 2016; Zhang et al., 2018; Jung et al., 2020, 2021; Abdullin et al., 2021; Van Goethem et al., 2025), including specialized polar isolates such as *A. bergstromii* from Antarctic endoliths. Interestingly, two described species (*A. compacta* and *A. gigantea*) function as cyanobionts in lichen symbioses with *Peltula* fungi (Jung et al., 2021). Given that well-developed black lichen communities were frequently observed on our marble samples, the *Aliterella* phylotypes detected here may similarly form symbiotic lichen associations in addition to maintaining a free-living chasmolithic existence. Other taxa that were particularly present within the investigated marble samples but were not identified with the general bacterial primers (Savaglia et al., 2024) belonged to the *Cyanothece* genus (12 %), from the Cyanothecaceae family (Gomontiellales order). This genus is only represented by unicellular phylotypes, of which some are able to fix nitrogen (Sherman et al., 2010).

3.5.2. *Cyanothecaceae* dominates gneiss samples

Gneiss bedrock was characterized by the striking dominance of the unicellular genus *Cyanothece* (Cyanothecaceae, Gomontiellales). Significant sequence yields (>1000 reads per sample) were obtained from Perlebandet North samples featuring visible lichen cover, whereas highly mineralized, bare gneiss soils yielded lower read counts (<1000 reads per sample) that were thus discarded. Despite this, community composition across the investigated gneiss samples was remarkably consistent.

The dominance of *Cyanothece* is probably linked to its key ecological role as a primary producer and potential nitrogen fixer in hyper-oligotrophic Antarctic soils (Sherman et al., 2010). In Antarctica, the only *Cyanothece* strain that has been isolated from lichen vegetation is *Cyanothece aeruginosa* NIVA-CYA 258/1 (Mareš et al., 2019). This genus was also reported from soils with visible ‘algal’ crusts in Victoria Valley, McMurdo Dry Valleys (Broady, 2005). While *Cyanothece* was absent in earlier gneiss surveys from bare, unvegetated soils in the Teltet nunatak (not sampled in the present study), phylotypes belonging to this genus were previously found in biocrusts encountered in the granitic Pingvinane 4th nunatak (Pushkareva et al., 2018), further supporting a lichen-symbiotic lifestyle or specific microhabitat preference that warrants further physiological characterization.

This dominance of members of the Cyanothecaceae family in gneiss samples had not been observed in the earlier study conducted with general bacterial primers, that only highlighted the Chroococcidiopsaceae family as the dominant family (Savaglia et al., 2024). Overall, any *Cyanothece* phylotypes were encountered within the dataset of our previous survey. Interestingly, within the SILVA database, the only few sequences belonging to the *Cyanothece* genus are misassigned to the Microcystaceae or the Thermosynechococcaceae families, and none of these families were present on the dataset of our previous survey (Savaglia et al., 2024; in which Silva v138.1 database was used).

3.5.3. *Gomontiellaceae* dominates moraine samples of the Dry Valley

Moraines from the Dry Valley (DV) harbored extremely low cyanobacterial read abundances and displayed the lowest ASV richness among all investigated substrate types (**Fig. S3.3**). Exposed to intense katabatic winds and severe desiccation, cyanobacteria at this site are largely restricted to subsurface gravel layers, where Cyanobacteriota represented a minor component of the overall bacterial community (Savaglia et al., 2024).

These hyper-arid soils were overwhelmingly dominated by the non-heterocystous filamentous family Gomontiellaceae (formerly included within Oscillatoriales; Komárek et al., 2014; Bohunická et al., 2015), composed mainly of the genus *Hormoscilla* and, to a lesser extent, *Crinalium*. Gomontiellaceae encompasses specialized, rarely reported cyanobacteria and only in low abundances, adapted to high-stress environments worldwide, such as alpine regions in Peru, Nepal, and Tibet, as well as coastal sand dunes (Bohunická et al., 2015; Mikhailyuk et al., 2019). According to the SILVA database, the *Hormoscilla* genus did not exist, and the *Crinalium* genus belonged to the Chroococcidiopsaceae family. However, in Savaglia et al. (2024) the Cyanobacteriota phylum represented only 3% of the total reads within moraine soils, with ~ 1% belonging to the Chroococcidiopsaceae family, that probably represented misclassified sequences. Only very few sequences belonging to the genus *Crinalium* were previously found in Taylor Valley (McMurdo Dry Valleys, Michaud et al., 2012), in a lake in the Larsemann Hills (Snowbowl Lake; Pessi et al., 2023) and soils of our sampled region in a previous survey (Tytgat et al., 2016), as well as in cryoconite holes (Singh et al., 2021; Broady, 2005). Although *Hormoscilla* and *Crinalium* share high 16S rRNA sequence identity (>99%), well above standard genus-level thresholds (Stackebrandt & Goebel, 1994), their distinct ecological and morphological traits justify their separation (Bohunická et al., 2015).

3.5.4. High morphological diversity and taxonomic evenness characterize granitic samples

Granitic soils hosted diverse cyanobacterial assemblages co-dominated by Gomontiellaceae, Oculatellaceae, Cyanothecaceae, Microcoleaceae, Leptolyngbyaceae and Nostocaceae (**Fig. 3.2** and **Table S3.4**). While granitic substrates exhibited high richness – consistent with earlier regional surveys at both the bacterial (Tytgat et al., 2016) and cyanobacterial levels (Pushkareva et al., 2018) – our broader sampling revealed that granitic bedrock exhibits intermediate ASV richness, significantly lower than marble and gneiss (**Fig. 3.2** and **Table S3.3**). Despite this moderate richness, granitic soils displayed clear site-level divergence in ordination space (**Fig. 3.3** and **Fig. S3.5**), probably driven by site-specific physical differences and micro-habitat conditions. However, in contrast to marble, granitic soils exhibited much greater taxonomic evenness and structural diversity at higher taxonomic ranks. Rather than being dominated by a single group, granitic communities were evenly represented across these six major cyanobacterial families, reflecting the broad array of ecological niches provided by the heterogeneous physical structure of granite. On these substrates, biocrusts with well-developed

lichen communities are visibly established, a pattern reflected also by the relatively high TOC values (Chilton et al., 2018). In the present study, more locations and more substrate types were added compared to earlier surveys (Tytgat et al., 2016; Pushkareva et al., 2018), which allowed a more nuanced view on the cyanobacterial diversity in the region.

Granite communities from Pingvinane 4th nunatak (PA4; yellow triangles) and Utsteinen (UT) formed a cohesive cluster on the left side of NMDS space (**Fig. 3.3**), associated with high total organic carbon (TOC) and visible biocrusts (Chilton et al., 2018). The coarse texture, cracks, and high water-retention capacity of granite boulders provide stable microhabitats and protection from desiccation (Namsaraev et al., 2010; Caster et al., 2021).

The 4th nunatak of Pingvinane range (PA4) and Utsteinen (UT) formed a cohesive cluster on the left side of NMDS space (**Fig. 3.3**). The coarse texture, cracks, and high water-retention capacity of granite boulders provide stable microhabitats and protection from desiccation (Namsaraev et al., 2010; Caster et al., 2021). Furthermore, chemical weathering of granite bedrock may supply available phosphorus and potassium which can favour key functional taxa across the substrate (Stewart and Alexander, 1971), including unicellular *Cyanothece*, filamentous Oscillatoriales (*Phormidesmis*, *Hormoscilla*), and diazotrophs (*Nostoc*) (**Table S3.5**). Indeed, granite had the highest phosphate content (**Table S3.1**) which could be an advantage for nitrogen-fixing cyanobacteria that need higher amounts compared to other organisms in order to perform diazotrophy. Filamentous taxa such as *Phormidesmis* (*P. priestleyi*) and *Leptolyngbya* possess cold-tolerant ancestral lineages and are widely distributed across Arctic and Antarctic habitats (Taton et al., 2006; Jungblut et al., 2009; Christmas et al., 2015; 2018; Mai et al., 2018). The prominence of *Timaviella* aligns with its known adaptation to low-light microhabitats (Sciuto et al., 2017), while *Nostoc* acts as a foundational keystone diazotroph, supplying bioavailable nitrogen to the broader soil ecosystem (Belnap et al., 2003; Büdel & Colesie, 2014; Lee et al., 2016; Van Goethem et al., 2017; Ortiz et al., 2020), as confirmed by our indicator species and network analyses.

Conversely, granite samples from Pingvinane 6th nunatak (PA6) separated cleanly along NMDS1 towards the right side of ordination space (**Fig. 3.3**). This nunatak was characterized by well-developed microalgal biocrusts, suggesting that competitive exclusion by eukaryotic phototrophs may be responsible for the lower raw cyanobacterial diversity at this site (**Fig. S3.2**). Petrellnuten granite (PT) showed notable intra-site divergence in ordination space (**Fig. 3.3** and **Fig. S3.5**), with samples being heavily dominated by Gomontiellaceae (primarily *Hormoscilla*; **Fig. 3.2**, **Tables S3.4** and **S3.5**), which was also dominant in the Dry Valley samples. This could be

explained by the exposed geographic position near the polar plateau of the East Antarctic Ice Sheet.

Many phylotypes within granitic soils, similarly to other substrate types (i.e., marble and moraine of Yûboku Valley, see **Table S3.4** and **3.5**) remained unclassified at the genus level (~20%) and the family level (~8%) and may comprehend potentially endemic lineages. This is not uncommon for desert microbial communities (Niederberger et al., 2008; Lee et al., 2012; Makhalanyane et al., 2013; Van Goethem et al., 2016) and other Antarctic environments, such as the benthic microbial mats, where Pessi et al. (2018) also encountered a high percentage of “novel” cyanobacterial sequences.

3.6. Conclusion

This study demonstrates that different substrate type and potential local microhabitat heterogeneity support distinct cyanobacterial communities across the nunataks and valleys of the western Sør Rondane Mountains. While marble soils harbored the highest overall ASV richness, granitic substrates were distinguished by their exceptional taxonomic evenness and structural diversity, supporting balanced assemblages of filamentous, unicellular, and heterocystous cyanobacteria. In contrast, the other substrates were characterized by strong single-family/genus dominance (e.g., Chroococcidiopsidaceae on marble bedrock, Cyanothecaceae on gneiss bedrock, and Gomontiellaceae on Dry Valley moraines). Significant site-level divergence among granitic soils further highlights the critical role of physical micro-refugia, moisture retention, and localized weathering in structuring polar micro-ecosystems.

Crucially, a substantial proportion of sequences across all substrates remained unclassified at family and genus levels, underscoring that Antarctic cyanobacteria remain significantly underrepresented in reference databases. This highlights the urgent need for targeted isolation efforts, the use of specific primers, and updated reference databases to improve taxonomic assignment of cyanobacterial sequences. Ultimately, these findings provide an essential cyanobacterial biodiversity baseline to inform future ecological monitoring and conservation efforts within Antarctic Specially Protected Area (ASPA) No. 179 (Antarctic Treaty Consultative Meeting, 2023).

3.7. Acknowledgments

This work was funded by the Federal Belgian Science Policy Office (BELSPO) under the BRAIN-BE program MICROBIAN project (BR/165/A1/MICROBIAN). The authors thank the crew of the Princess Elisabeth Antarctica (PES) station for the operation of the base and providing the necessary support during the field activities, in particular Sam Lambrechts and Quinten Vanhellemont for helping in collecting the samples, Alain Hubert for leading the Belgian Antarctic Research Expedition (BELARE) and our field guides for the 2018 campaign, Pierre-Yves Terrettaz, and the 2019 campaign, Raphaël Mayoraz. ESA/Copernicus is thanked for the free and open access to Sentinel-2 imagery used in Fig.

3.1. VS and BD had a PhD FRIA fellowship from the FRS-FNRS. AW is Senior Research Associate of the FRS-FNRS.

This paper is a contribution to the Scientific Committee on Antarctic Research (SCAR) Scientific Research Program ‘Integrated science to inform Antarctic and Southern Ocean Conservation’ (Ant-ICON).

3.8. Author contributions

Conceived and designed the study: EV, WV, BT, AW. Performed the sampling: VS, EV. Performed the laboratory analysis: BD. Analyzed the data: BD, VS. Minor data analysis: MH. Supervision: EV, AW. Secured the funding WV, AW and EV. Wrote the paper: VS. All authors reviewed and commented on previous versions of the manuscript and approved the final version. BD and VS equally contributed to the work.

3.9. References

- Antarctic Treaty Consultative Meeting, 2023. Antarctic Specially Protected Area No. 179 (Parts of Western Sør Rondane Mountains, Dronning Maud Land, East Antarctica): Management plan, Measure 18 (2023). Secretariat of the Antarctic Treaty.
- Azua-Bustos, A., Zúñiga, J., Arenas-Fajardo, C., Orellana, M., Salas, L. and Rafael, V., 2014. *Gloeocapsopsis* AAB1, an extremely desiccation-tolerant cyanobacterium isolated from the Atacama Desert. *Extremophiles*, 18, pp.61-74.
- Bebout, B.M. and Garcia-Pichel, F., 1995. UV B-induced vertical migrations of cyanobacteria in a microbial mat. *Applied and Environmental Microbiology*, 61(12), pp.4215-4222.
- Belnap, J., Büdel, B., Lange, O.L., 2003. Biological soil crusts: characteristics and distribution. In: Belnap, J., Lange, O.L. (Eds.), *Biological Soil Crusts: Structure, Function, and Management*. Springer-Verlag, Berlin, pp. 3–30.
- Blanchet, F.G., Legendre, P. and Borcard, D., 2008. Forward selection of explanatory variables. *Ecology*, 89(9), pp.2623-2632.
- Bohunická, M., Mareš, J., Hrouzek, P., Urajová, P., Lukeš, M., Šmarda, J., Komárek, J., Gaysina, L.A. and Strunecký, O., 2015. A combined morphological, ultrastructural, molecular, and biochemical study of the peculiar family Gomontiellaceae (Oscillatoriales) reveals a new cylindrospermopsin-producing clade of cyanobacteria. *Journal of Phycology*, 51(6), pp.1040-1054.
- Broady, P.A., 1996. Diversity, distribution and dispersal of Antarctic terrestrial algae. *Biodiversity & Conservation*, 5, pp.1307-1335.
- Broady, P.A., 2005. The distribution of terrestrial and hydro-terrestrial algal associations at three contrasting locations in southern Victoria Land, Antarctica. *Algol Stud*, 118(1), pp.95-112.
- Brooks, S.T., Jabour, J., van den Hoff, J., Bergstrom, D.M., 2019. Our footprint on Antarctica competes with nature for rare ice-free land. *Nat. Sustain.* 2, 185–190.
- Büdel, B. and Colesie, C., 2014. Biological soil crusts. *Antarctic Terrestrial Microbiology: Physical and Biological Properties of Antarctic Soils*, pp.131-161.
- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A. and Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature methods*, 13(7), pp.581-583.
- Čapková, K., Hauer, T., Řeháková, K. and Doležal, J., 2016. Some like it high! Phylogenetic diversity of high-elevation cyanobacterial community from biological soil crusts of western Himalaya. *Microbial ecology*, 71, pp.113-123.
- Caster, J., Sankey, T.T., Sankey, J.B., Bowker, M.A., Buscombe, D., Duniway, M.C., Barger, N., Faist, A. and Joyal, T., 2021. Biocrust and the soil surface: Influence of climate, disturbance, and biocrust recovery on soil surface roughness. *Geoderma*, 403, p.115369.

Cavacini, 2001. Soil algae from northern Victoria Land (Antarctica). *Polar Bioscience*, 14 (2001), pp. 45-60

Chan, Y., Lacap, D.C., Lau, M.C., Ha, K.Y., Warren-Rhodes, K.A., Cockell, C.S., Cowan, D.A., McKay, C.P. and Pointing, S.B., 2012. Hypolithic microbial communities: between a rock and a hard place. *Environmental microbiology*, 14(9), pp.2272-2282.

Chilton, A.M., Neilan, B.A. and Eldridge, D.J., 2018. Biocrust morphology is linked to marked differences in microbial community composition. *Plant and soil*, 429, pp.65-75.

Choe, Y.H., Kim, M. and Lee, Y.K., 2021. Distinct microbial communities in adjacent rock and soil substrates on a high Arctic polar desert. *Frontiers in Microbiology*, 11, p.607396.

Christmas, N.A., Anesio, A.M. and Sánchez-Baracaldo, P., 2015. Multiple adaptations to polar and alpine environments within cyanobacteria: a phylogenomic and Bayesian approach. *Frontiers in microbiology*, 6, p.1070.

Christmas, N.A., Anesio, A.M. and Sanchez-Baracaldo, P., 2018. The future of genomics in polar and alpine cyanobacteria. *FEMS Microbiology Ecology*, 94(4), p.fiy032.

Cockell, C.S. and Stokes, M.D., 2004. Widespread colonization by polar hypoliths. *Nature*, 431(7007), pp.414-414.

De Cáceres, M., Legendre, P. and Moretti, M., 2010. Improving indicator species analysis by combining groups of sites. *Oikos*, 119(10), pp.1674-1684.

de Los Ríos, A., Ascaso, C., Wierzbos, J., Fernández-Valiente, E. and Quesada, A., 2004. Microstructural characterization of cyanobacterial mats from the McMurdo Ice Shelf, Antarctica. *Applied and Environmental Microbiology*, 70(1), pp.569-580.

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P. and Andersen, G.L., 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and environmental microbiology*, 72(7), pp.5069-5072.

Dextro, R.B., Delbaje, E., Freitas, P.N.N., Geraldès, V., Pinto, E., Long, P.F. and Fiore, M.F. (2023), Environmental adaptations by the intertidal Antarctic cyanobacterium *Halotia branconii* CENA392 as revealed using long-read genome sequencing. *Limnol. Oceanogr. Lett.*, 8: pp. 770-780.

Durieu, B., Savaglia, V., Lara, Y., Lambion, A., Pessi, I.S., Vyverman, W., Verleyen, E. and Wilmotte, A. (2025), (Sub-)Antarctic endemic cyanobacteria from benthic mats are rare and have restricted geographic distributions. *Ecography*, 2025: e07489.

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods*, 10(10), pp.996-998.

Effendi, D.B., Sakamoto, T., Ohtani, S., Awai, K. and Kanesaki, Y., 2022. Possible involvement of extracellular polymeric substrates of Antarctic cyanobacterium *Nostoc* sp. strain SO-36 in adaptation to harsh environments. *Journal of Plant Research*, 135(6), pp.771-784.

Fernandes, V.M., Machado de Lima, N.M., Roush, D., Rudgers, J., Collins, S.L. and Garcia-Pichel, F., 2018. Exposure to predicted precipitation patterns decreases population size and alters community structure of cyanobacteria in biological soil crusts from the Chihuahuan Desert. *Environmental Microbiology*, 20(1), pp.259-269.

Fernández-Carazo, R., Namsaraev, Z., Mano, M.J., Ertz, D. and Wilmotte, A., 2012. Cyanobacterial diversity for an anthropogenic impact assessment in the Sør Rondane Mountains area, Antarctica. *Antarctic Science*, 24(3), pp.229-242.

Garcia-Pichel, F., Loza, V., Marusenko, Y., Mateo, P. and Potrafka, R.M., 2013. Temperature drives the continental-scale distribution of key microbes in topsoil communities. *Science*, 340(6140), pp.1574-1577.

Greening, C., Islam, Z.F. and Bay, S.K., 2022. Hydrogen is a major lifeline for aerobic bacteria. *Trends in Microbiology*, 30(4), pp.330-337.

Harmel, M., Durieu, B., Savaglia, V., Tytgat, B., Baurain, D., Wilmotte, A., Verleyen, E., & Cornet, L. (2025). A FAIR Amplicon Sequencing Workflow for Long-term Environmental Monitoring (p. 2025.12.04.692289). *bioRxiv*.

Hijmans, R. J. (2021). Introduction to the "geosphere" package (Version 1.5-14).

Jung, P., Brust, K., Schultz, M., Büdel, B., Donner, A. and Lakatos, M., 2021. Opening the gap: rare lichens with rare cyanobionts—unexpected cyanobiont diversity in cyanobacterial lichens of the order Lichinales. *Frontiers in Microbiology*, 12, p.728378.

Jung, P., Mikhailyuk, T., Emrich, D., Baumann, K., Dultz, S. and Büdel, B., 2020. Shifting boundaries: Ecological and geographical range extension based on three new species in the cyanobacterial genera *Cyanocohniella*, *Oculatella*, and *Aliterella*. *Journal of phycology*, 56(5), pp.1216-1231.

Jung, P., Schermer, M., Briegel-Williams, L., Baumann, K., Leinweber, P., Karsten, U., Lehnert, L., Achilles, S., Bendix, J. and Büdel, B., 2019. Water availability shapes edaphic and lithic cyanobacterial communities in the Atacama Desert. *Journal of phycology*, 55(6), pp.1306-1318.

Jungblut, A.D., Allen, M.A., Burns, B.P. and Neilan, B.A., 2009. Lipid biomarker analysis of cyanobacteria-dominated microbial mats in meltwater ponds on the McMurdo Ice Shelf, Antarctica. *Organic Geochemistry*, 40(2), pp.258-269.

Jungblut, A.D., Hawes, I., Mountfort, D., Hitzfeld, B., Dietrich, D.R., Burns, B.P. and Neilan, B.A., 2005. Diversity within cyanobacterial mat communities in variable salinity meltwater ponds of McMurdo Ice Shelf, Antarctica. *Environmental microbiology*, 7(4), pp.519-529.

Komsta, L. and Novomestky, F., 2015. Moments, cumulants, skewness, kurtosis and related tests. R package version, 14(1).

Lee, C.K., Barbier, B.A., Bottos, E.M., McDonald, I.R. and Cary, S.C., 2012. The inter-valley soil comparative survey: the ecology of Dry Valley edaphic microbial communities. The ISME journal, 6(5), pp.1046-1057.

Lee, J.R., Raymond, B., Bracegirdle, T.J., Chadès, I., Fuller, R.A., Shaw, J.D. and Terauds, A., 2017. Climate change drives expansion of Antarctic ice-free habitat. Nature, 547(7661), pp.49-54.

Lee, K.C., Archer, S.D., Boyle, R.H., Lacap-Bugler, D.C., Belnap, J. and Pointing, S.B., 2016. Niche filtering of bacteria in soil and rock habitats of the Colorado Plateau Desert, Utah, USA. Frontiers in Microbiology, 7, p.1489.

Lefler, F.W., Berthold, D.E. and Dail Laughinghouse IV, H., 2023. CyanoSeq: a database of cyanobacterial 16S rRNA sequences with curated taxonomy. Journal of Phycology.

Mai, T., Johansen, J.R., Pietrasiak, N., Bohunicka, M. and Martin, M.P., 2018. Revision of the Synechococcales (Cyanobacteria) through recognition of four families including Oculatellaceae fam. nov. and Trichocoleaceae fam. nov. and six new genera containing 14 species. Phytotaxa.

Makhalanyane, T.P., Valverde, A., Birkeland, N.K., Cary, S.C., Marla Tuffin, I. and Cowan, D.A., 2013. Evidence for successional development in Antarctic hypolithic bacterial communities. The ISME Journal, 7(11), pp.2080-2090.

Mareš, J., Johansen, J.R., Hauer, T., Zima Jr, J., Ventura, S., Cuzman, O., Tiribilli, B. and Kaštovský, J., 2019. Taxonomic resolution of the genus *Cyanothece* (Chroococcales, Cyanobacteria), with a treatment on *Gloeothece* and three new genera, *Crocospaera*, *Rippkaea*, and *Zehria*. Journal of Phycology, 55(3), pp.578-610.

McArdle, B.H. and Anderson, M.J., 2001. Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology, 82(1), pp.290-297.

McMurdie, P.J. and Holmes, S., 2013. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PloS one, 8(4), p.e61217.

Mikhailyuk, T., Vinogradova, O., Holzinger, A., Glaser, K., Samolov, E. and Karsten, U., 2019. New record of the rare genus *Crinalium* Crow (Oscillatoriales, Cyanobacteria) from sand dunes of the Baltic Sea, Germany: epitypification and emendation of *Crinalium magnum* Fritsch et John based on an integrative approach. Phytotaxa, 400(3), p.165.

Muñoz-Martín, M.Á., Becerra-Absalón, I., Perona, E., Fernández-Valbuena, L., Garcia-Pichel, F. and Mateo, P., 2019. Cyanobacterial biocrust diversity in Mediterranean ecosystems along a latitudinal and climatic gradient. New Phytologist, 221(1), pp.123-141.

Namsaraev, Z., Mano, M.J., Fernandez, R. and Wilmotte, A., 2010. Biogeography of terrestrial cyanobacteria from Antarctic ice-free areas. Annals of Glaciology, 51(56), pp.171-177.

Niederberger, T.D., McDonald, I.R., Hacker, A.L., Soo, R.M., Barrett, J.E., Wall, D.H. and Cary, S.C., 2008. Microbial community composition in soils of Northern Victoria Land, Antarctica. *Environmental microbiology*, 10(7), pp.1713-1724.

Nübel, U., Garcia-Pichel, F. and Muyzer, G., 1997. PCR primers to amplify 16S rRNA genes from cyanobacteria. *Applied and environmental microbiology*, 63(8), pp.3327-3332.

Obbels, D., Verleyen, E., Mano, M.J., Namsaraev, Z., Sweetlove, M., Tytgat, B., Fernandez-Carazo, R., De Wever, A., D'hondt, S., Ertz, D. and Elster, J., 2016. Bacterial and eukaryotic biodiversity patterns in terrestrial and aquatic habitats in the Sør Rondane Mountains, Dronning Maud Land, East Antarctica. *FEMS microbiology ecology*, 92(6).

Ortiz, M., Bosch, J., Coclet, C., Johnson, J., Lebre, P., Salawu-Rotimi, A., Vikram, S., Makhalanyane, T. and Cowan, D., 2020. Microbial nitrogen cycling in Antarctic soils. *Microorganisms*, 8(9), p.1442.

Paulson, J.N., Pop, M. and Bravo, H.C., 2013. metagenomeSeq: Statistical analysis for sparse high-throughput sequencing. *Bioconductor package*, 1(0), p.191.

Peres-Neto, P. R., Legendre, P., Dray, S., & Borcard, D. (2006). Variation partitioning of species data matrices: estimation and comparison of fractions. *Ecology*, 87(10), 2614-2625.

Pessi, I.S., Lara, Y., Durieu, B., Maalouf, P.D.C., Verleyen, E. and Wilmotte, A., 2018. Community structure and distribution of benthic cyanobacteria in Antarctic lacustrine microbial mats. *FEMS Microbiology Ecology*, 94(5), p.fiy042.

Pessi, I.S., Maalouf, P.D.C., Laughinghouse IV, H.D., Baurain, D. and Wilmotte, A., 2016. On the use of high-throughput sequencing for the study of cyanobacterial diversity in Antarctic aquatic mats. *Journal of Phycology*, 52(3), pp.356-368.

Pessi, I.S., Popin, R.V., Durieu, B., Lara, Y., Savaglia, V., Roncero-Ramos, B., Hultman, J., Verleyen, E., Vyverman, W. and Wilmotte, A., 2023. Novel diversity of polar Cyanobacteria revealed by genome-resolved metagenomics. *bioRxiv*, pp.2023-02.

Pointing, S.B., Büdel, B., Convey, P., Gillman, L.N., Koerner, C., Leuzinger, S. and Vincent, W.F., 2015. Biogeography of photoautotrophs in the high polar biome. *Frontiers in plant science*, 6, p.692.

Pointing, S.B., Chan, Y., Lacap, D.C., Lau, M.C., Jurgens, J.A. and Farrell, R.L., 2009. Highly specialized microbial diversity in hyper-arid polar desert. *Proceedings of the National Academy of Sciences*, 106(47), pp.19964-19969.

Pushkareva, E., Pessi, I.S., Namsaraev, Z., Mano, M.J., Elster, J. and Wilmotte, A., 2018. Cyanobacteria inhabiting biological soil crusts of a polar desert: Sør Rondane Mountains, Antarctica. *Systematic and applied microbiology*, 41(4), pp.363-373.

Pushkareva, E., Elster, J., Kudoh, S., Imura, S. and Becker, B. (2024) Microbial community composition of terrestrial habitats in East Antarctica with a focus on microphototrophs. *Front. Microbiol.* 14:1323148.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J. and Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research*, 41(D1), pp.D590-D596.

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

Rego, A., Raio, F., Martins, T.P., Ribeiro, H., Sousa, A.G., Séneca, J., Baptista, M.S., Lee, C.K., Cary, S.C., Ramos, V. and Carvalho, M.F., 2019. Actinobacteria and cyanobacteria diversity in terrestrial antarctic microenvironments evaluated by culture-dependent and independent methods. *Frontiers in microbiology*, 10, p.1018.

Rigonato, J., Gama, W.A., Alvarenga, D.O., Branco, L.H.Z., Brandini, F.P., Genuario, D.B. and Fiore, M.F., 2016. *Aliterella atlantica* gen. nov., sp. nov., and *Aliterella antarctica* sp. nov., novel members of coccoid Cyanobacteria. *International Journal of Systematic and Evolutionary Microbiology*, 66(8), pp.2853-2861.

Rodríguez-Caballero, E., Cantón, Y., Chamizo, S., Afana, A. and Solé-Benet, A., 2012. Effects of biological soil crusts on surface roughness and implications for runoff and erosion. *Geomorphology*, 145, pp.81-89.

Roncero-Ramos, B., Savaglia, V., Durieu, B., Van de Vreken, I., Richel, A., & Wilmotte, A. (2024). Ecophysiological and genomic approaches to cyanobacterial hardening for restoration. *Journal of Phycology*, 60, pp. 465–482.

Sakaeva, A., Sokol, E. R., Kohler, T. J., Stanish, L. F., Spaulding, S. A., Howkins, A., ... & McKnight, D. M. (2016). Evidence for dispersal and habitat controls on pond diatom communities from the McMurdo Sound Region of Antarctica. *Polar Biology*, 39(12), 2441-2456.

Savaglia, V., Lambrechts, S., Tytgat, B., Vanhellemont, Q., Elster, J., Willems, A., Wilmotte, A., Verleyen, E. and Vyverman, W. (2024) Geology defines microbiome structure and composition in nunataks and valleys of the Sør Rondane Mountains, East Antarctica. *Front. Microbiol.* 15:1316633.

Schulze-Makuch, D., Wagner, D., Kounaves, S.P., Mangelsdorf, K., Devine, K.G., de Vera, J.P., Schmitt-Kopplin, P., Grossart, H.P., Parro, V., Kaupenjohann, M. and Galy, A., 2018. Transitory microbial habitat in the hyperarid Atacama Desert. *Proceedings of the National Academy of Sciences*, 115(11), pp.2670-2675.

Sciuto, K., Moschin, E. and Moro, I., 2017. Cryptic cyanobacterial diversity in the Giant Cave (Trieste, Italy): the new genus *Timaviella* (Leptolyngbyaceae). *Cryptogamie, Algologie*, 38(4), pp.285-323.

Sherman, L.A., Min, H., Toepel, J. and Pakrasi, H.B., 2010. Better living through Cyanothecae—unicellular diazotrophic cyanobacteria with highly versatile metabolic systems. In Recent advances in phototrophic prokaryotes (pp. 275-290). Springer New York.

Sokol, E. R., Herbold, C. W., Lee, C. K., Cary, S. C., & Barrett, J. E. (2013). Local and regional influences over soil microbial metacommunities in the Transantarctic Mountains. *Ecosphere*, 4(11), 1-24.

Stackebrandt, E. and Goebel, B.M., 1994. Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International journal of systematic and evolutionary microbiology*, 44(4), pp.846-849.

Stewart, W.D.P. and Alexander, G., 1971. Phosphorus availability and nitrogenase activity in aquatic blue-green algae. *Freshwater Biology*, 1(4), pp.389-404.

Strunecký, O., Ivanova, A.P. and Mareš, J., 2023. An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. *Journal of Phycology*, 59(1), pp.12-51.

Suganuma, Y., Miura, H., Zondervan, A. and Okuno, J.I., 2014. East Antarctic deglaciation and the link to global cooling during the Quaternary: Evidence from glacial geomorphology and ¹⁰Be surface exposure dating of the Sør Rondane Mountains, Dronning Maud Land. *Quaternary Science Reviews*, 97, pp.102-120.

Tamaru, Y., Takani, Y., Yoshida, T. and Sakamoto, T., 2005. Crucial role of extracellular polysaccharides in desiccation and freezing tolerance in the terrestrial cyanobacterium *Nostoc commune*. *Applied and Environmental Microbiology*, 71(11), pp.7327-7333.

Taton, A., Grubisic, S., Balthasart, P., Hodgson, D.A., Laybourn-Parry, J. and Wilmotte, A., 2006a. Biogeographical distribution and ecological ranges of benthic cyanobacteria in East Antarctic lakes. *FEMS microbiology ecology*, 57(2), pp.272-289.

Taton, A., Grubisic, S., Brambilla, E., De Wit, R. and Wilmotte, A., 2003. Cyanobacterial diversity in natural and artificial microbial mats of Lake Fryxell (McMurdo Dry Valleys, Antarctica): a morphological and molecular approach. *Applied and environmental microbiology*, 69(9), pp.5157-5169.

Taton, A., Grubisic, S., Ertz, D., Hodgson, D.A., Piccardi, R., Biondi, N., Tredici, M.R., Mainini, M., Losi, D., Marinelli, F. and Wilmotte, A., 2006b. Polyphasic study of Antarctic cyanobacterial strains 1. *Journal of Phycology*, 42(6), pp.1257-1270.

Timoncini, A., Costantini, F., Bernardi, E., Martini, C., Mugnai, F., Mancuso, F.P., Sassoni, E., Ospitali, F. and Chiavari, C., 2022. Insight on bacteria communities in outdoor bronze and marble artefacts in a changing environment. *Science of The Total Environment*, 850, p.157804.

Tytgat, B., Verleyen, E., Sweetlove, M., D'hondt, S., Clercx, P., Van Ranst, E., Peeters, K., Roberts, S., Namsaraev, Z., Wilmotte, A. and Vyverman, W., 2016. Bacterial community composition in

relation to bedrock type and macrobiota in soils from the Sør Rondane Mountains, East Antarctica. *FEMS Microbiology Ecology*, 92(9).

Ugolini, F.C. and Bockheim, J.G., 2008. Antarctic soils and soil formation in a changing environment: a review. *Geoderma*, 144(1-2), pp.1-8.

Van Goethem, M.W., Makhalanyane, T.P., Cowan, D.A. and Valverde, A., 2017. Cyanobacteria and Alphaproteobacteria may facilitate cooperative interactions in niche communities. *Frontiers in microbiology*, 8, p.2099.

Van Goethem, M.W., Vikram, S., Cowan, D.A. and Makhalanyane, T.P., 2025. Comparative genomics reveals adaptive traits in novel Antarctic lithic cyanobacteria. *BMC genomics*, 26(1), pp.1-15.

Venables, W.R. and Ripley, B., 2002. BD (2002). *Modern Applied Statistics with S*. New York: Springer Science & Business Media, 200, pp.183-206.

Vincent, W., 1993. Antarctic cyanobacteria: light, nutrients and photosynthesis in the microbial mat environment. *J. Phycol.*, 29, pp.745-755.

Vincent, W.F., 2000. Evolutionary origins of Antarctic microbiota: invasion, selection and endemism. *Antarctic Science*, 12(3), pp.374-385.

Vincent, W.F., 2009. Cyanobacteria. In: *Encyclopedia of Inland Waters* (ed. Likens, G.E.). Elsevier Inc., Oxford, pp. 226–232.

Weber, B., Belnap, J., Büdel, B., Antoninka, A.J., Barger, N.N., Chaudhary, V.B., Darrouzet-Nardi, A., Eldridge, D.J., Faist, A.M., Ferrenberg, S. and Havrilla, C.A., 2022. What is a biocrust? A refined, contemporary definition for a broadening research community. *Biological Reviews*, 97(5), pp.1768-1785.

Wei, T., Simko, V., Levy, M., Xie, Y., Jin, Y. and Zemla, J., 2017. Package ‘corrplot’. *Statistician*, 56(316), p.e24.

Whitton, B.A. and Potts, M., 2012. Introduction to the cyanobacteria. *Ecology of cyanobacteria II: their diversity in space and time*, pp.1-13.

Whitton, B.A. and Potts, M., 2012. Introduction to the cyanobacteria. *Ecology of cyanobacteria II: their diversity in space and time*, pp.1-13.

Williams, L., Loewen-Schneider, K., Maier, S. and Büdel, B., 2016. Cyanobacterial diversity of western European biological soil crusts along a latitudinal gradient. *FEMS microbiology ecology*, 92(10).

Wood, S.A., Mountfort, D., Selwood, A.I., Holland, P.T., Puddick, J. and Cary, S.C., 2008. Widespread distribution and identification of eight novel microcystins in Antarctic cyanobacterial mats. *Applied and Environmental Microbiology*, 74(23), pp.7243-7251.

Zhang, L., Jungblut, A.D., Hawes, I., Andersen, D.T., Sumner, D.Y. and Mackey, T.J., 2015. Cyanobacterial diversity in benthic mats of the McMurdo Dry Valley lakes, Antarctica. *Polar Biology*, 38, pp.1097-1110.

Zhang, Q.I., Zheng, L., Li, T., Li, R. and Song, L., 2018. *Aliterella shaanxiensis* (Aliterellaceae), a new coccoid cyanobacterial species from China. *Phytotaxa*, 374(3), pp.211-2.

3.10. Supplementary information

Table S3.1. Environmental data.

Sample	Latitude	Longitude	Substrate	Substrate detail	pH	EC ($\mu\text{S cm}^{-1}$)	TOC (%)	N_NH ₄ (mg kg ⁻¹)	N_NO ₃ (mg kg ⁻¹)	DW (%)	P_PO ₄ (mg kg ⁻¹)	TN (%)	TP (%)	Elevation (m a.s.l.)	# cyano reads
DV_2018_2	-72.11599	23.157045	moraine_DV	(Mo)_Sch_D_Gbh_To	6.36	119.5	0.46	1.645	0.133	99.55	16.177	0.02	0.08	1712	2020
DV_2018_3	-72.11476	23.163443	moraine_DV	(Mo)_Sch_D_Gbh_To	6.33	260	0.74	1.31	0.077	99.51	28.958	0.02	0.07	1694	1077
DV_2018_4	-72.11468	23.163328	moraine_DV	(Mo)_Sch_D_Gbh_To	6.44	78.9	0.49	1.085	0.021	99.64	24.585	0.03	0.07	1694	1258
DV_OTC12_2018_A	-72.11223	23.184898	moraine_DV	(Mo)_Sch_D_Gbh_To	6.14	47.6	0.49	0.637	0.125	99.92	44.325	0.01	0.07	1666	1086
DV_OTC12_2018_B	-72.11223	23.184898	moraine_DV	(Mo)_Sch_D_Gbh_To	6.29	140	0.3	0.922	0.135	99.87	40.959	0.01	0.09	1666	24501
DV_OTC13_2018_A	-72.11219	23.184841	moraine_DV	(Mo)_Sch_D_Gbh_To	6.1	41.5	0.39	0.855	0.091	99.91	52.72	0.02	0.13	1667	1461
DV_OTC13_2018_B	-72.11219	23.184841	moraine_DV	(Mo)_Sch_D_Gbh_To	6.78	57.2	0.7	1.342	0.202	99.88	43.947	0.02	0.1	1667	26079
DV_OTC13_2018_C	-72.11219	23.184841	moraine_DV	(Mo)_Sch_D_Gbh_To	6.56	50.8	0.64	1.047	0.082	99.91	37.159	0.01	0.07	1667	47812
DV_OTC14_2018_B	-72.11219	23.184841	moraine_DV	(Mo)_Sch_D_Gbh_To	6.5	264	0.4	0.927	0.646	99.89	41.475	0.06	0.15	1667	7959
DV_OTC14_2018_C	-72.11219	23.184841	moraine_DV	(Mo)_Sch_D_Gbh_To	6.46	46.9	0.3	1.275	0.211	99.92	43.192	0.05	0.15	1667	6181
DV_SF_2018_1_B	-72.11258	23.18555	moraine_DV	(Mo)_Sch_D_Gbh_To	6.45	59.2	0.34	0.802	0.143	99.74	37.011	0.02	0.09	1667	1741
DV_SF_2018_1_C	-72.11258	23.18555	moraine_DV	(Mo)_Sch_D_Gbh_To	6.4	52.7	0.57	1.101	0.026	99.82	37.628	0.03	0.07	1667	1267
DV_SF_2018_3_C	-72.11258	23.18555	moraine_DV	(Mo)_Sch_D_Gbh_To	6.49	62.7	0.49	1.217	0.046	99.63	37.329	0.02	0.08	1667	3359
DV_SF_2018_CON_1_C	-72.11262	23.185282	moraine_DV	(Mo)_Sch_D_Gbh_To	6.24	33.7	0.38	0.377	0.092	99.34	40.899	0.01	0.1	1668	1541
DV_SF_2018_CON_2_A	-72.11278	23.185186	moraine_DV	(Mo)_Sch_D_Gbh_To	5.75	125.6	0.85	0.259	0.031	99.84	61.912	0.03	0.12	1668	1224
DV_SF_2018_CON_2_C	-72.11278	23.185186	moraine_DV	(Mo)_Sch_D_Gbh_To	5.55	38.9	0.35	0.631	0.031	98.91	62.933	0.03	0.12	1668	2291
DV_SF_2018_CON_3_A	-72.11271	23.184937	moraine_DV	(Mo)_Sch_D_Gbh_To	6.14	37.4	0.38	0.477	0.634	99.72	53.264	0.02	0.09	1668	2002
DV_SF_2018_CON_3_B	-72.11271	23.184937	moraine_DV	(Mo)_Sch_D_Gbh_To	5.81	22.3	0.3	0.504	0.232	99.19	63.504	0.02	0.12	1668	1157
DV_SF_2018_CON_3_C	-72.11271	23.184937	moraine_DV	(Mo)_Sch_D_Gbh_To	5.88	16.9	0.43	0.733	0.056	99.19	42.924	0.02	0.12	1668	2116
DV_SF2_2018_1_A	-72.11214	23.185244	moraine_DV	(Mo)_Sch_D_Gbh_To	6.44	26.4	0.44	1.234	0.066	99.34	44.511	0.01	0.05	1666	15589
DV_SF2_2018_1_C	-72.11214	23.185244	moraine_DV	(Mo)_Sch_D_Gbh_To	7.8	74.9	0.57	2.735	0.619	99.89	55.622	0.04	0.12	1666	12135
DV_SF2_2018_2_A	-72.112114	23.185379	moraine_DV	(Mo)_Sch_D_Gbh_To	6.41	29	0.73	1.082	0.081	99.52	48.976	0.02	0.11	1666	2508
DV_SF2_2018_2_B	-72.112114	23.185379	moraine_DV	(Mo)_Sch_D_Gbh_To	6.51	32.8	0.58	0.636	0.076	99.87	50.217	0.03	0.11	1666	2032
DV_SF2_2018_2_C	-72.112114	23.185379	moraine_DV	(Mo)_Sch_D_Gbh_To	6.59	100.4	0.53	0.831	0.549	99.88	42.732	0.04	0.1	1666	15788

DV_SF2_2018_3_A	-72.112114	23.185379	moraine_DV	(Mo)_Sch_D_Gbh_To	5.91	19.3	0.48	0.663	0.051	99.88	62.873	0.02	0.13	1666	29693
DV_SF2_2018_3_B	-72.112114	23.185379	moraine_DV	(Mo)_Sch_D_Gbh_To	6.29	108.7	0.43	0.482	0.188	99.81	51.853	0.02	0.1	1666	6841
DV_SF2_2018_3_C	-72.112114	23.185379	moraine_DV	(Mo)_Sch_D_Gbh_To	5.98	18.2	0.46	0.889	0.076	99.75	53.056	0.01	0.12	1666	1387
DV_SF2_2018_CON_1_C	-72.11206	23.184668	moraine_DV	(Mo)_Sch_D_Gbh_To	6.85	185.5	0.67	0.558	0.091	99.68	56.133	0.02	0.12	1667	1688
PA4_2018_1	-72.00135	22.999208	granite	(Gr)	6.12	22.9	0.76	2.753	0.088	99.75	48.568	0.06	0.12	1409	27269
PA4_2018_2	-72.00131	22.999477	granite	(Gr)	6.06	19.4	0.88	0.989	0.089	99.82	55.398	0.08	0.09	1410	30221
PA4_2018_3	-72.00144	22.999647	granite	(Gr)	6.41	22.4	1.46	3.359	0.389	99.45	57.074	0.09	0.13	1410	4971
PA4_2018_4	-72.00144	22.999647	granite	(Gr)	6.58	20.7	1.05	2.004	0.223	99.36	61.305	0.03	0.05	1410	32024
PA4_2018_5	-72.00149	23.000029	granite	(Gr)	6.21	21.1	0.43	1.379	0.111	99.45	67.494	0.14	0.1	1411	24096
PA4_2018_6	-72.00143	23.000105	granite	(Gr)	6.22	15.4	1.13	1.464	0.116	99.25	40.294	0.03	0.06	1412	4054
PA4_2018_7	-72.00139	23.000048	granite	(Gr)	6.45	47.4	1.54	8.585	0.108	99.74	55.434	0.1	0.05	1412	25013
PA6_2019_1	-72.010506	22.989674	granite	(Gr)	7.31	62.1	0.78	2.017	0.898	99.41	86.474	0.06	0.09	1412	17071
PA6_2019_2	-72.01036	22.991528	granite	(Gr)	7.04	193.4	1.47	1.859	0.289	98.25	73.589	0.07	0.07	1416	25056
PA6_2019_4	-72.010056	22.993479	granite	(Gr)	6.97	24.8	0.31	7.283	0.312	99.75	114.824	0.02	0.09	1416	1751
PA6_2019_5	-72.01002	22.993422	granite	(Gr)	7.57	65.6	0.32	8.557	1.053	99.85	95.352	0.03	0.06	1416	2016
PA6_2019_6	-72.0101	22.993536	granite	(Gr)	6.52	108.8	3.4	24.04	5.464	99.35	119.427	0.46	0.02	1417	1005
PB_N_2018_1	-71.84361	22.838146	gneiss	(Gbg)	5.88	414.2	9.23	20.526	0.04	98.25	38.751	0.23	0.4	1256	1402
PB_N_2018_10	-71.84405	22.834051	marble	(M)	9.09	149.7	1.48	0.89	0.506	99.74	7.968	0.27	0.21	1287	22408
PB_N_2018_11	-71.84405	22.834051	marble	(M)	8.98	56.5	0.79	0.437	0.036	99.68	8.81	0.15	0.12	1287	22218
PB_N_2018_3	-71.84353	22.838034	gneiss	(Gbg)	5.01	163.1	10.95	46.657	0.136	98.21	48.899	0.28	0.17	1255	10687
PB_N_2018_6	-71.84395	22.835169	gneiss	(Gbg)	5.63	91.7	4.6	12.926	0.036	99.15	34.056	0.2	0.41	1279	7732
PB_N_2018_7	-71.843895	22.835245	gneiss	(Gbg)	5.82	49.2	2.52	13.117	0.131	99.22	18.639	0.16	0.16	1277	35269
PB_N_2018_8	-71.84399	22.834904	gneiss	(Gbg)	8.38	89.4	0.68	3.906	0.964	99.02	23.818	0.07	0.81	1282	20527
PB_N_2018_9	-71.84405	22.834051	marble	(M)	7.4	336	5.62	4.553	0.351	99.45	18.515	0.36	0.23	1287	6834
PB_N_OTC15_16_17_2019_CON_A	-71.84423	22.831173	marble	(M)	9.09	208.1	0.48	0.926	0.27	99.95	6.899	0.04	0.03	1310	10878
PB_N_OTC15_2019_B	-71.84423	22.831173	marble	(M)	9.21	58.2	0.5	0.292	0.05	99.78	6.686	0.06	0.07	1310	34486
PB_N_OTC15_2019_C	-71.84423	22.831173	marble	(M)	9.34	60.2	0.46	0.096	0.031	99.85	6.537	0.13	0.08	1310	12796
PB_N_OTC16_2019_C	-71.84423	22.831173	marble	(M)	9.08	90.7	0.6	0.286	0.827	99.89	6.445	0.06	0.06	1310	23142
PB_N_OTC17_2019_A	-71.84423	22.831173	marble	(M)	8.73	72.5	0.59	1.081	0.025	99.74	7.534	0.08	0.13	1310	11150
PB_N_OTC17_2019_B	-71.84423	22.831173	marble	(M)	8.93	57.6	0.38	0.997	0.022	99.91	6.749	0.06	0.04	1310	18004

PB_N_OTC17_2019_C	-71.84423	22.831173	marble	(M)	8.88	66.4	0.69	1.354	0.015	99.74	6.906	0.08	0.07	1310	5316
PB_N_OTC9_10_2018_CON_A	-71.84405	22.834051	marble	(M)	9.42	117.9	1.19	0.321	0.035	99.91	6.918	0.24	0.12	1287	37705
PB_N_OTC9_10_2018_CON_B	-71.84405	22.834051	marble	(M)	9.17	73.6	0.58	0.59	0.06	99.95	7.173	0.27	0.16	1287	13828
PB_N_OTC9_10_2018_CON_C	-71.84405	22.834051	marble	(M)	9.45	117.3	1.14	0.282	0.05	99.78	7.104	0.21	0.1	1287	12201
PB_N_OTC9_2018_A	-71.84405	22.834051	marble	(M)	9.46	160.4	1.3	0.761	0.062	99.91	8.042	0.23	0.15	1287	7913
PB_N_OTC9_2018_B	-71.84405	22.834051	marble	(M)	9.49	185.5	1.49	0.392	0.151	99.78	8.109	0.17	0.12	1287	3497
PB_N_OTC9_2018_C	-71.84405	22.834051	marble	(M)	9.55	142.6	1.56	0.748	0.045	99.82	9.184	0.26	0.12	1287	1330
PB_S_2018_1	-71.8797	22.746996	gneiss	(Gbh)	6.45	17.6	1.41	1.648	0.088	99.27	17.073	0.09	0.16	1349	45191
PB_S_2018_2	-71.87994	22.746687	marble	(M)	9.01	63.7	0.53	0.299	0.173	99.35	14.094	0.04	0.1	1354	7021
PB_S_2018_3	-71.88009	22.74659	marble	(M)	8.94	50.7	0.78	0.333	0.09	99.41	14.28	0.04	0.04	1357	22034
PT_2019_2	-72.007385	22.832005	granite	(Gr)	6.13	70.4	1.28	2.444	3.837	99.57	60.359	0.11	0.14	1469	5671
PT_2019_3	-72.007416	22.831738	granite	(Gr)	7.11	31.8	0.45	3.258	0.428	99.47	42.777	0.04	0.06	1472	5830
PT_2019_5	-72.00757	22.833204	granite	(Gr)	7.01	14.9	0.25	0.721	0.28	99.9	36.502	0.03	0.07	1460	35996
UT_2018_1	-71.9456	23.345005	granite	(Gr)	5.89	21.1	2.46	7.048	0.303	99.55	64.229	0.15	0.11	1362	6332
UT_2018_2	-71.94543	23.345041	granite	(Gr)	6.28	20.9	1.83	2.835	0.077	99.28	26.373	0.16	0.15	1358	18612
UT_2018_3	-71.94559	23.345139	granite	(Gr)	5.24	30.1	2.15	5.929	0.084	99.36	151.121	0.22	0.12	1361	18555
UT_2018_6	-71.94626	23.3456	granite	(Gr)	5.95	30.3	1.1	1.607	0.076	99.75	39.742	0.25	0.12	1372	18738
UT_2018_7	-71.94624	23.345734	granite	(Gr)	5.87	14.2	0.59	1.151	0.146	99.81	40.212	0.07	0.04	1370	10779
UT_OTC11_2018_A	-71.945465	23.344967	granite	(Gr)	5.92	58.1	2.49	2.059	0.091	99.72	23.943	0.35	0.17	1359	10370
UT_OTC11_2018_CON_B	-71.945465	23.344967	granite	(Gr)	6.73	36.6	0.89	1.334	0.115	99.89	43.154	0.09	0.09	1359	20052
UT_OTC2_2018_CON_1_A	-71.946075	23.345505	granite	(Gr)	5.83	329.1	2.67	30.354	0.497	99.25	96.667	0.34	0.24	1369	1801
UT_OTC2_2018_CON_1_C	-71.946075	23.345505	granite	(Gr)	5.97	273.1	2.13	22.79	1.773	99.72	50.768	0.43	0.2	1369	1043
YO_2018_5	-72.081375	23.793673	moraine_YO	(Mo)_Sch_To	8.45	110.8	1.19	0.627	0.197	99.74	15.318	0.05	0.1	1360	42897
YO_2018_6	-72.08111	23.793453	moraine_YO	(Mo)_Sch_To	8.53	60.2	1.38	0.407	0.296	99.82	16.151	0.03	0.04	1360	22075
YO_2019_3	-72.08115	23.793186	moraine_YO	(Mo)_Sch_To	8.81	90.5	0.54	0.256	0.236	99.81	22.712	0.04	0.09	1361	17072
YO_2019_5	-72.08337	23.79094	moraine_YO	(Mo)_Sch_To	8.03	4750	2.47	1.04	290	98.25	20.732	0.09	0.07	1354	32190
YO_2019_6	-72.08344	23.790731	moraine_YO	(Mo)_Sch_To	8.96	93.7	0.57	0.286	2.38	99.65	27.897	0.02	0.08	1357	8977

DV: Dry Valley; PB_N: Perlebandet North (I nunatak); PB_S: Perlebandet South (III nunatak); PA: Pingvinane North (III nunatak); PA_S: Pingvinane South (VI nunatak); PT: Petrellnuten; UT: Utsteinen; YO: Yûboku Valley. Mo_DV: moraines of Dry Valley; mo_YO: moraines of Yûboku Valley. (Mo)_Amp: Moraine with orthopyroxene

amphibolite; (Mo)_Sch_D_Gbh_To: Moraine with amphibole schist, diorite and quartz, biotite-hornblende gneiss and metatonalite; (M): Marble and skarns; (Gbg): Sillimanite-garnet-biotite gneiss, partly containing cordierite; (Gr): Granite and pegmatite; (Gbh): Biotite-hornblende gneiss and hornblende gneiss containing garnet in places; (Mo)_Sch_To: Moraine with amphibole schist and metatonalite. EC: electric conductivity; TOC: total carbon; DW: dry weight; TN: total nitrogen; TP: total phosphorus.

Table S3.2. Mock community.

ASV_ID	Control	Mock_comm.	ULC093	ULC100	ULC112	ULC166	ULC420	ULC469	Taxonomy
ASV_1	0	2746	820	1064	1857	825	831	38141	Prochlorococcaceae
ASV_21	0	2669	111	40565	183	98	68	280	<i>Anagnostidinema pseudacutissimum</i>
ASV_22	0	2611	38318	187	558	134	164	166	<i>Laspinema lumbricale</i>
ASV_12	0	2392	262	133	83999	241	222	236	<i>Heteroleibleinia</i>
ASV_16	0	2353	164	167	220	23028	108	121	Microcoleaceae
ASV_13	0	2320	254	163	482	176	14938	150	Nostoc

Table S3.3. Kruskal-Wallis rank sum test showing the effect of substrate type on environmental parameters.

	df	χ^2	P
pH	4	44.841	<0.001
TOC	4	17.693	<0.05
N-NO ₃	4	12.617	<0.05
P-PO ₄	3	46.136	<0.001
N-NH ₄	4	36.576	<0.001
TN	4	25.693	<0.001
TP	4	13.887	<0.001
DW	4	21.477	<0.001
EC	4	23.886	<0.001

EC: Electric conductivity; TOC: Total carbon; TN: Total nitrogen; TP: Total phosphorus.

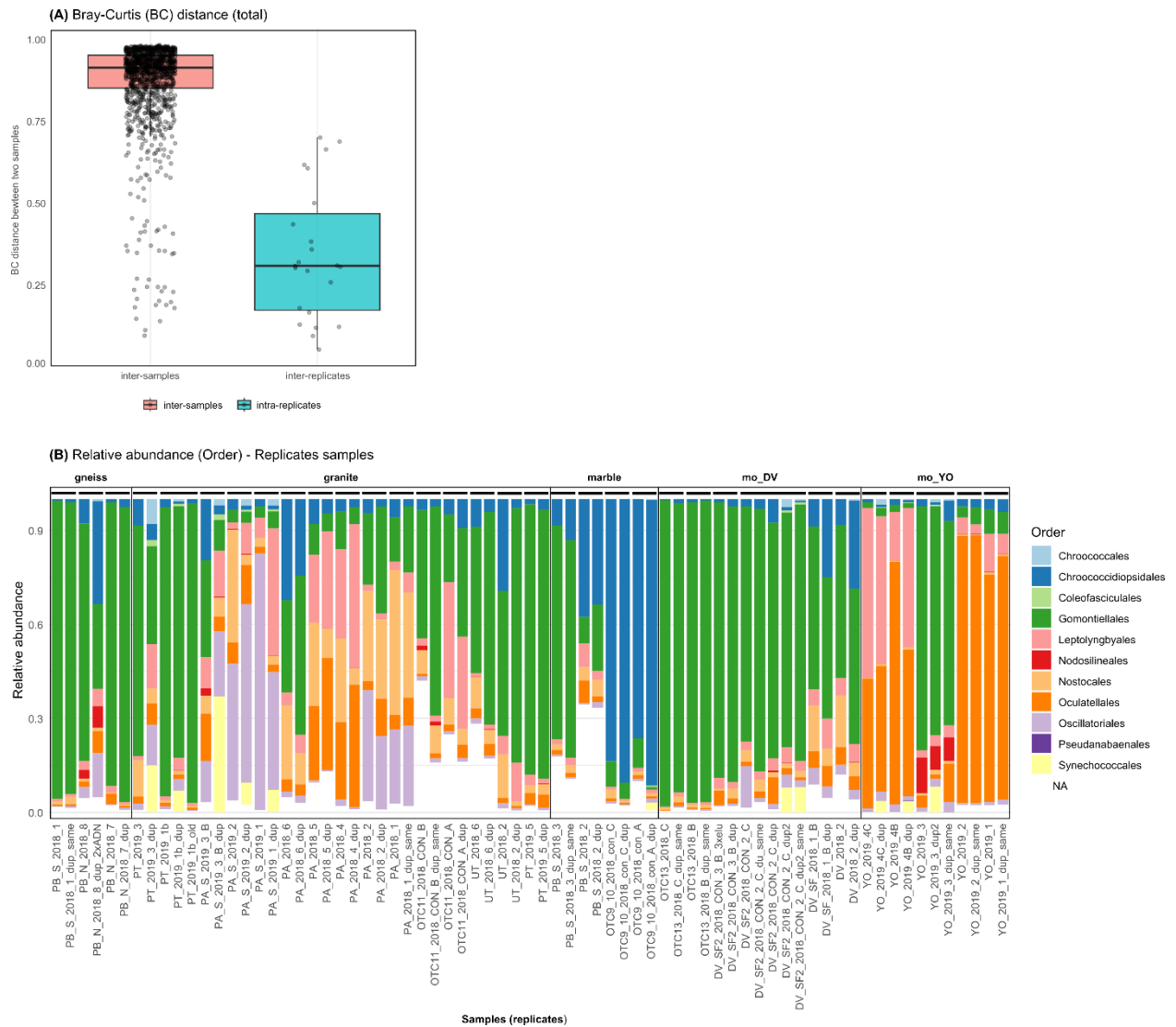


Fig. S3.1. Validation of technical replicate concordance and pipeline reliability. **(A)** Boxplots showing pairwise Bray–Curtis dissimilarity distances between distinct samples (inter-samples) versus technical duplicates (inter-replicates). The significantly lower dissimilarity among replicates confirms experimental and sequencing consistency. **(B)** Stacked bar charts showing relative taxonomic abundance across all paired duplicate samples (e.g. sample vs. sample_dup), illustrating near-identical community composition between technical pairs. A single representative replicate with the highest read depth was retained for downstream analyses.

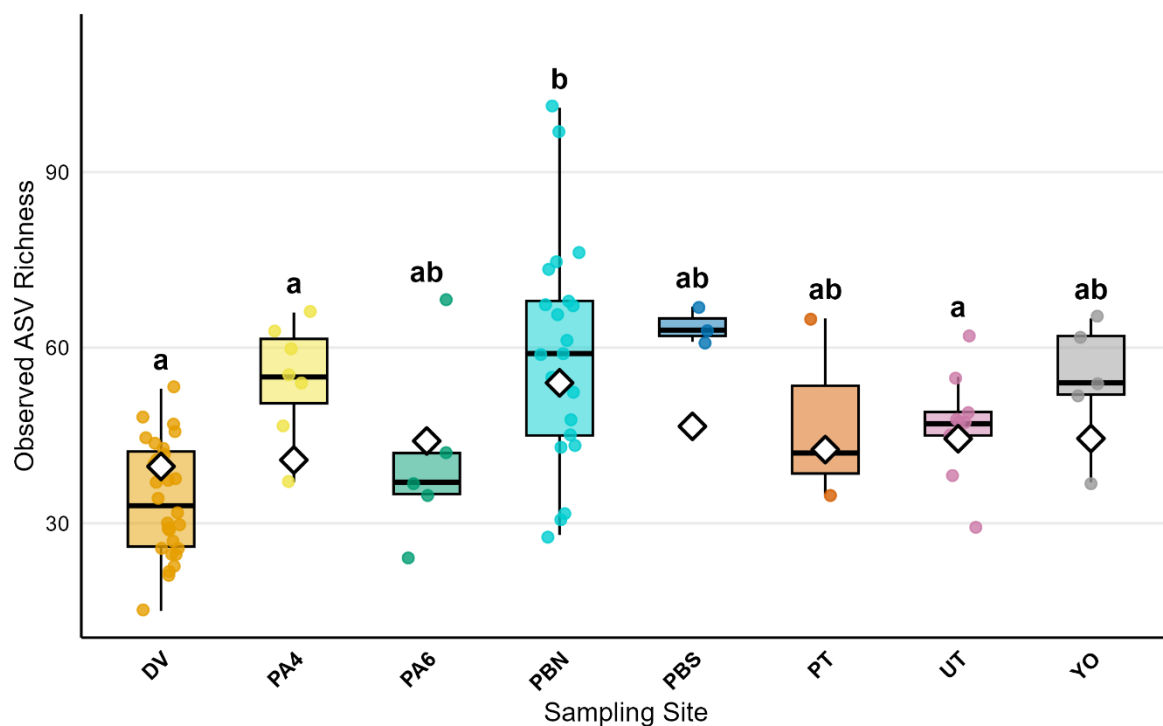


Fig. S3.2. Boxplots of cyanobacterial richness per sampled site (n=81). Diamonds represent GLM adjusted means; circles data points. Letters correspond to the significant differences tested via Tukey test. DV = moraines of Dry Valley; PA4 = granitic Pingvinane 4th nunatak; PA6 = granitic Pingvinane 6th nunatak; PBN = Perlebandet N nunatak; PBS = Perlebandet S nunatak; PT = granitic Petrellnuten nunatak; UT = granitic Utsteinen ridge; YO = moraines of Yûboku-dan Valley.

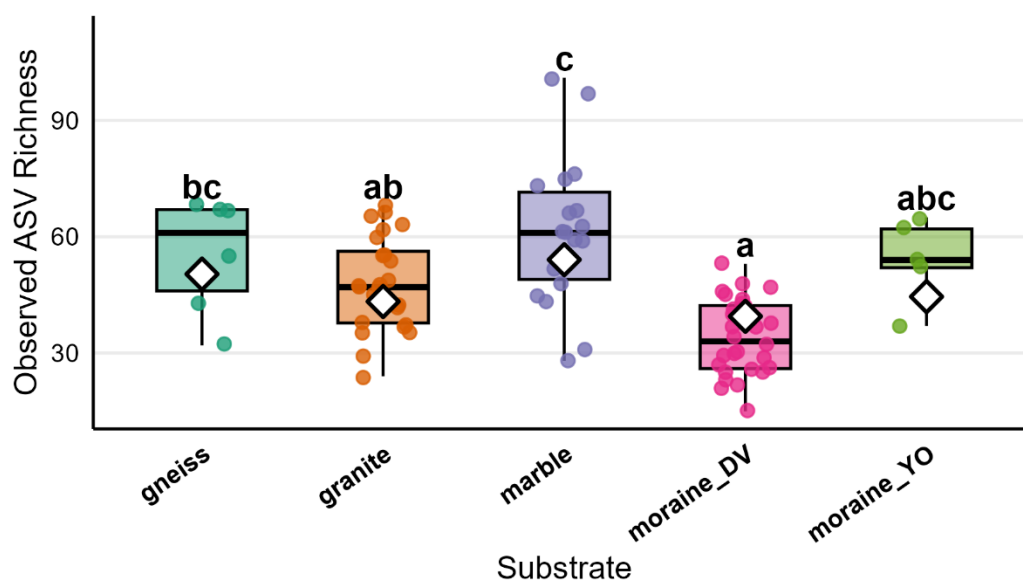


Fig. S3.3. Boxplots of cyanobacterial richness per substrate type. Diamonds represent GLM adjusted means; circles data points. Letters correspond to the significant differences tested via Tukey test.

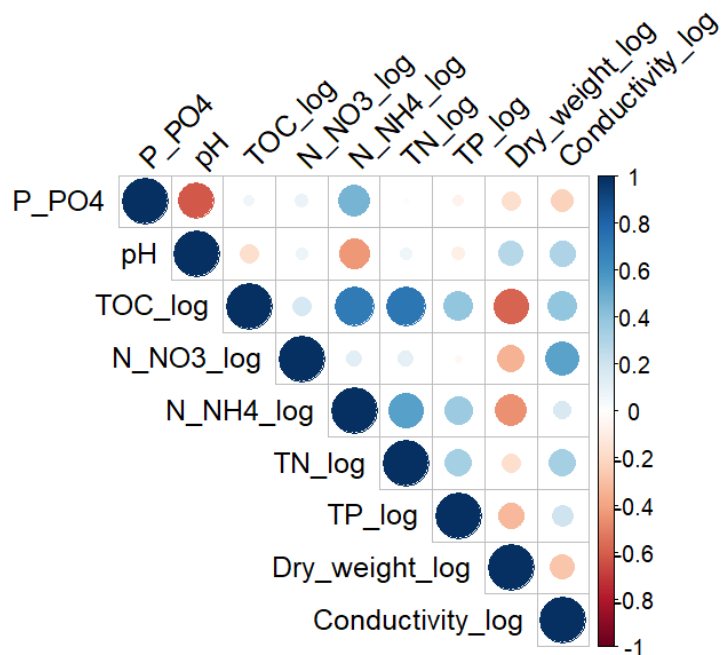


Fig. S3.4. Correlogram showing Pearson's correlation coefficients of all geochemical variables (transformed when necessary). TOC: total organic carbon; N_NH4: N-NH₄⁺; N_NO3: N-NO₃⁻; DW: dry weight; TN: total nitrogen; TP: total phosphorus; P_PO4: P-PO₄³⁻.

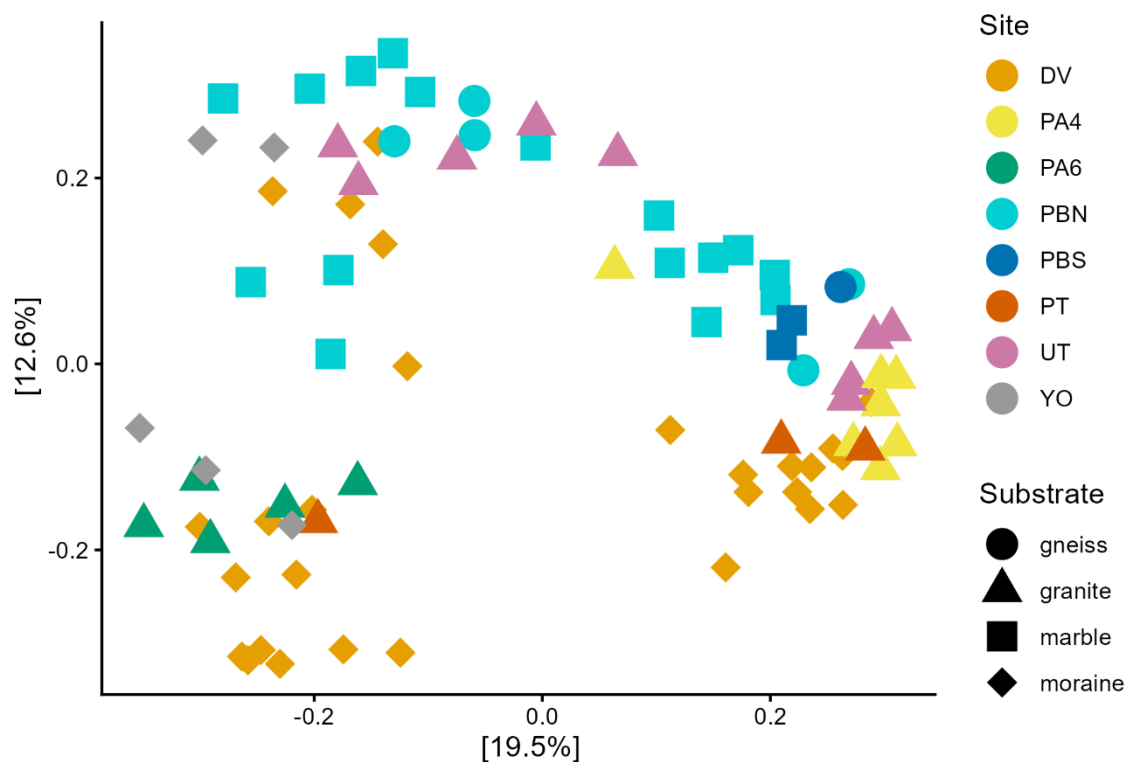


Fig. S3.5. Principal Coordinates Analysis (PCoA) plot at the ASV level of Bray-Curtis distances for each sample (n=81). Symbols are colored by sampling site, with different shapes representing the substrate type (see legend on the right).

Key Substrate Indicator Taxa

Showing dominant taxa driving single-substrate separation (IndVal, $p \leq 0.05$)

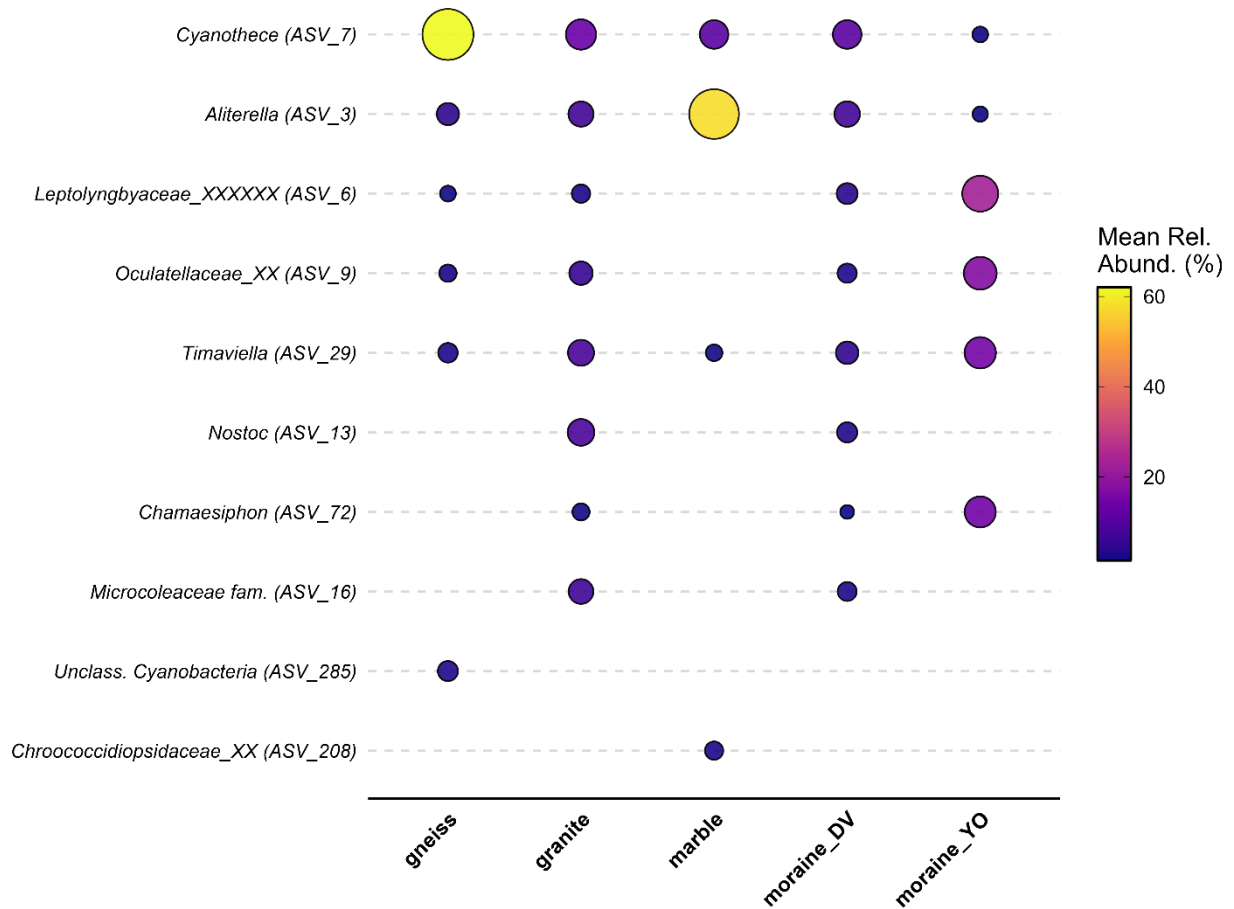


Fig. S3.6. Substrate-specific cyanobacterial indicator ASVs. Bubble plot displaying dominant indicator taxa ($p \leq 0.05$, IndVal test) associated with specific geological substrates. Y-axis lists candidate ASVs with their corresponding taxonomic assignment; X-axis denotes substrate types. Circle size and color intensity indicate mean relative abundance (%), highlighting distinct compositional specialization across different rock and moraine substrates.

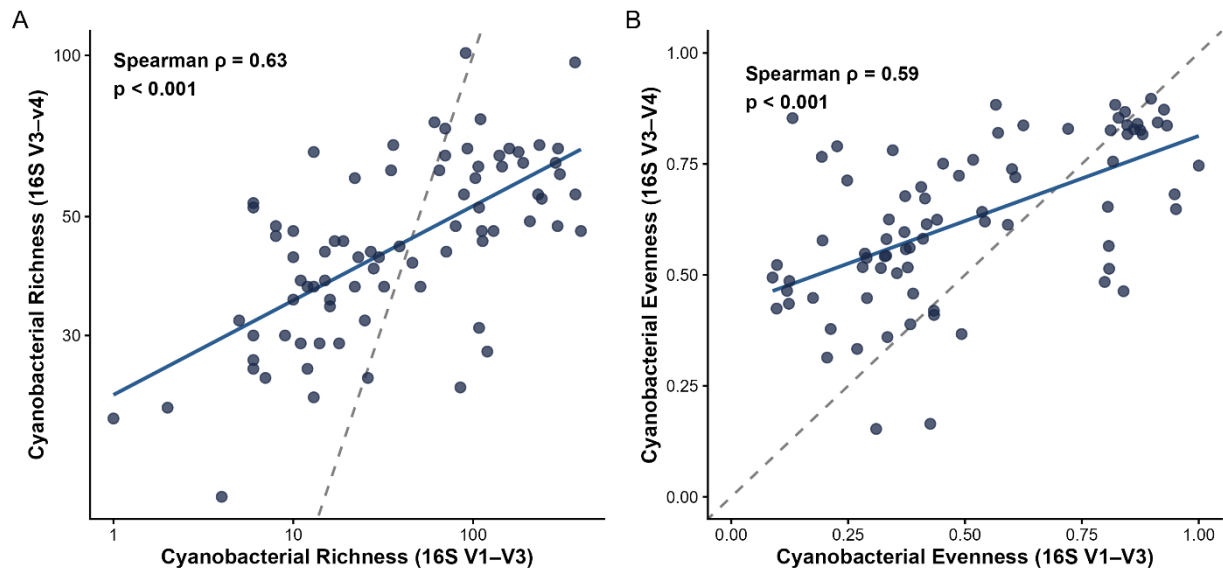


Fig. S3.7. Comparison of cyanobacterial alpha diversity between general bacterial (16S V1–V3) and cyanobacteria-specific (16S V3–V4) primer profiling. Pairwise correlations for (A) cyanobacterial ASV richness (log-scaled) and (B) community evenness in paired samples. Metrics on the x -axis were derived from general bacterial 16S rRNA gene profiling used in Savaglia et al. (2024) study, that were targeting the V1–V3 hypervariable regions using primers pA and BKL1 (Edwards et al., 1989; Cleenwerck et al., 2007). Metrics on the y -axis were derived from targeted cyanobacterial profiling of the V3–V4 hypervariable regions using primers CYA359F and CYA781Rb (Nübel et al., 1997). Solid blue lines represent linear regression trends; dashed grey lines indicate the 1:1 line of identity ($y = x$). Both richness (Spearman $\rho = 0.63$, $p < 0.001$) and evenness (Spearman $\rho = 0.59$, $p < 0.001$) display strong, statistically significant positive correlations, confirming that cyanobacterial diversity patterns are robust and consistent across both profiling strategies.

Table S3.4. Family taxonomic assignment of the ASVs with cumulative abundance calculated for each substrate type.

All substrates			Gneiss			Granite		
Family	read counts	% of reads	Family	read counts	% of reads	Family	read counts	% of reads
Gomontiellaceae	250538	23.008%	Cyanothecaceae	97622	80.808%	Gomontiellaceae	58513	16.798%
Chroococcidiopsidaceae	199722	18.342%	Gomontiellaceae	5002	4.140%	Oculatellaceae	48861	14.027%
Cyanothecaceae	192887	17.714%	Microcoleaceae	4175	3.456%	Cyanothecaceae	44013	12.636%
Oculatellaceae	117865	10.824%	Chroococcidiopsidaceae	3751	3.105%	Microcoleaceae	42380	12.167%
Leptolyngbyaceae	97799	8.981%	Cyanophyceae (unclassified Family)	2653	2.196%	Leptolyngbyaceae	42082	12.081%
Cyanophyceae (unclassified Family)	82875	7.611%	Leptolyngbyaceae	2346	1.942%	Cyanophyceae (unclassified Family)	35412	10.166%
Microcoleaceae	53179	4.884%	Oculatellaceae	1881	1.557%	Nostocaceae	33071	9.494%
Nostocaceae	38205	3.509%	Cyanobacteriota (unclassified Family)	859	0.711%	Chroococcidiopsidaceae	15019	4.312%
Chamaesiphonaceae	19351	1.777%	Chamaesiphonaceae	626	0.518%	Tolypothrichaceae	13367	3.837%
Tolypothrichaceae	15447	1.419%	Cymatolegaceae	592	0.490%	Chamaesiphonaceae	8942	2.567%
Nostocales (unclassified Family)	7027	0.645%	Tolypothrichaceae	448	0.371%	Nostocales (unclassified Family)	5583	1.603%
Prochlorococcaceae	5569	0.511%	Nostocaceae	411	0.340%	Cymatolegaceae	518	0.149%
Cymatolegaceae	3432	0.315%	Nostocales (unclassified Family)	278	0.230%	Aphanizomenonaceae	296	0.085%
Microcystaceae	1263	0.116%	Stigonemataceae	161	0.133%	Cyanobacteriota (unclassified Family)	221	0.063%
Cyanobacteriota (unclassified Family)	1142	0.105%	Pleurocapsaceae	3	0.002%	Microcystaceae	42	0.012%
Oscillatoriaceae	1047	0.096%				Pleurocapsaceae	6	0.002%
Aphanizomenonaceae	651	0.060%						
Wilmottiaceae	649	0.060%						
Stigonemataceae	161	0.015%						
Pseudanabaenaceae	54	0.005%						
Pleurocapsaceae	36	0.003%						

Marble			Moraine Dry Valley			Moraine Yûboku Valley		
Family	read counts	% of reads	Family	read counts	% of reads	Family	read counts	% of reads
Chroococcidiopsidaceae	172627	63.289%	Gomontiellaceae	167268	74.742%	Oculatellaceae	54763	44.447%
Cyanophyceae								
(unclassified Family)	40075	14.692%	Cyanothecaceae	17974	8.032%	Leptolyngbyaceae	38133	30.949%
Cyanothecaceae	31354	11.495%	Chroococcidiopsidaceae	6763	3.022%	Gomontiellaceae	13763	11.170%
Leptolyngbyaceae	8693	3.187%	Leptolyngbyaceae	6545	2.925%	Chamaesiphonaceae	7242	5.878%
Oculatellaceae	6280	2.302%	Oculatellaceae	6080	2.717%	Cymatolegaceae	2005	1.627%
						Cyanophyceae		
Gomontiellaceae	5992	2.197%	Prochlorococcaceae	4637	2.072%	(unclassified Family)	1973	1.601%
Microcoleaceae	1766	0.647%	Microcoleaceae	3771	1.685%	Cyanothecaceae	1924	1.562%
Nostocaceae	1377	0.505%	Nostocaceae	2847	1.272%	Chroococcidiopsidaceae	1562	1.268%
			Cyanophyceae					
Chamaesiphonaceae	1375	0.504%	(unclassified Family)	2762	1.234%	Microcoleaceae	1087	0.882%
Tolypothrichaceae	1042	0.382%	Chamaesiphonaceae	1166	0.521%	Nostocaceae	499	0.405%
Prochlorococcaceae	932	0.342%	Microcystaceae	1055	0.471%	Aphanizomenonaceae	99	0.080%
Nostocales (unclassified Family)	450	0.165%	Oscillatoriaceae	813	0.363%	Pseudanabaenaceae	54	0.044%
			Nostocales (unclassified Family)	681	0.304%			
Oscillatoriaceae	234	0.086%				Tolypothrichaceae	52	0.042%
						Nostocales (unclassified Family)	35	0.028%
Wilmottiaceae	184	0.067%	Tolypothrichaceae	538	0.240%			
Microcystaceae	146	0.054%	Wilmottiaceae	465	0.208%	Microcystaceae	20	0.016%
Aphanizomenonaceae	100	0.037%	Cymatolegaceae	234	0.105%			
Cymatolegaceae	83	0.030%	Aphanizomenonaceae	156	0.070%			
Cyanobacteriota								
(unclassified Family)	45	0.016%	Pleurocapsaceae	21	0.009%			
			Cyanobacteriota					
Pleurocapsaceae	6	0.002%	(unclassified Family)	17	0.008%			

Table S3.5. Genus taxonomic assignment of the ASVs with cumulative abundance calculated for each substrate type.

All substrates			Gneiss			Granite		
Genus	read counts	% of reads	Genus	read counts	% of reads	Genus	read counts	% of reads
<i>Hormoscilla</i>	224497	20.617%	<i>Cyanothece</i>	97622	80.808%	<i>Cyanothece</i>	44013	12.636%
<i>Cyanothece</i>	192887	17.714%	<i>Hormoscilla</i>	4534	3.753%	<i>Phormidesmis</i>	38318	11.001%
<i>Aliterella</i>	187087	17.181%	<i>Aliterella</i>	3453	2.858%	<i>Hormoscilla</i>	37850	10.866%
<i>Cyanophyceae (unclassified Genus)</i>	82875	7.611%	<i>Tychonema</i>	3278	2.713%	<i>Cyanophyceae (unclassified Genus)</i>	35412	10.166%
<i>Timaviella</i>	60862	5.589%	<i>Cyanophyceae (unclassified Genus)</i>	2653	2.196%	<i>Nostoc</i>	33071	9.494%
<i>Oculatellaceae_XX</i>	55261	5.075%	<i>Phormidesmis</i>	1655	1.370%	<i>Microcoleaceae (unclassified Genus)</i>	30417	8.732%
<i>Phormidesmis</i>	51156	4.698%	<i>Timaviella</i>	1219	1.009%	<i>Timaviella</i>	29758	8.543%
<i>Leptolyngbyaceae_XXXXXX</i>	41857	3.844%	<i>Cyanobacteriota (unclassified Genus)</i>	859	0.711%	<i>Oculatellaceae_XX</i>	18911	5.429%
<i>Nostoc</i>	38205	3.509%	<i>Microcoleus</i>	686	0.568%	<i>Gomontiellaceae (unclassified Genus)</i>	13647	3.918%
<i>Microcoleaceae (unclassified Genus)</i>	34491	3.168%	<i>Chamaesiphon</i>	626	0.518%	<i>Aliterella</i>	13425	3.854%
<i>Chamaesiphon</i>	19351	1.777%	<i>Oculatellaceae_XX</i>	596	0.493%	<i>Tolypothrichaceae (unclassified Genus)</i>	10593	3.041%
<i>Gomontiellaceae (unclassified Genus)</i>	18530	1.702%	<i>Cymatolege</i>	592	0.490%	<i>Tychonema</i>	9484	2.723%
<i>Tychonema</i>	14823	1.361%	<i>Tolypothrichaceae (unclassified Genus)</i>	434	0.359%	<i>Chamaesiphon</i>	8942	2.567%
<i>Tolypothrichaceae (unclassified Genus)</i>	12592	1.156%	<i>Leptolyngbyaceae_XXXXXX</i>	424	0.351%	<i>Crinalium</i>	7016	2.014%
<i>Crinalium</i>	7511	0.690%	<i>Nostoc</i>	411	0.340%	<i>Nostocales (unclassified Genus)</i>	5583	1.603%
<i>Chroococcidiopsidaceae_XX</i>	7496	0.688%	<i>Gomontiellaceae (unclassified Genus)</i>	304	0.252%	<i>Toxopsis</i>	2774	0.796%
<i>Nostocales (unclassified Genus)</i>	7027	0.645%	<i>Nostocales (unclassified Genus)</i>	278	0.230%	<i>Microcoleus</i>	2479	0.712%

<i>Prochlorococcaceae</i>								
(unclassified Genus)	5569	0.511%	<i>Stenomitos</i>	267	0.221%	<i>Leptolyngbyaceae_XXXXXX</i>	2165	0.622%
<i>Chroococcidiopsidaceae</i>			<i>Microcoleaceae</i>					
(unclassified Genus)	5139	0.472%	(unclassified Genus)	211	0.175%	<i>Stenomitos</i>	1392	0.400%
<i>Microcoleus</i>	3865	0.355%	<i>Chroococcidiopsidaceae</i>			<i>Chroococcidiopsidaceae</i>		
<i>Cymatolege</i>	3432	0.315%	(unclassified Genus)	197	0.163%	(unclassified Genus)	880	0.253%
<i>Toxopsis</i>	2855	0.262%	<i>Crinalium</i>	164	0.136%	<i>Chroococcidiopsidaceae_XX</i>	714	0.205%
<i>Stenomitos</i>	1926	0.177%	<i>Stigonema</i>	161	0.133%	<i>Cymatolege</i>	518	0.149%
			<i>Chroococcidiopsidaceae_XX</i>	101	0.084%	<i>Dolichospermum</i>	296	0.085%
<i>Shackletoniella</i>	1742	0.160%	<i>Shackletoniella</i>	66	0.055%	<i>Cyanobacteriota</i>		
<i>Apatinema</i>	1419	0.130%	<i>Toxopsis</i>	14	0.012%	(unclassified Genus)	221	0.063%
<i>Heteroleibleinia</i>	1405	0.129%	<i>Dermocarpella</i>	3	0.002%	<i>Shackletoniella</i>	192	0.055%
<i>Microcystis</i>	1263	0.116%				<i>Apatinema</i>	104	0.030%
<i>Cyanobacteriota</i>						<i>Heteroleibleinia</i>	67	0.019%
(unclassified Genus)	1142	0.105%						
<i>Laspinema</i>	1047	0.096%				<i>Microcystis</i>	42	0.012%
<i>Anagnostidinema</i>	649	0.060%				<i>Myxacorys</i>	36	0.010%
<i>Dolichospermum</i>	646	0.059%				<i>Dermocarpella</i>	6	0.002%
<i>Stigonema</i>	161	0.015%						
<i>Pseudanabaenaceae_X</i>	54	0.005%						
<i>Dermocarpella</i>	36	0.003%						
<i>Myxacorys</i>	36	0.003%						
<i>Aphanizomenon</i>	5	0.000%						

Marble			Moraine Dry Valley			Moraine Yûboku Valley		
Genus	read counts	% of reads	Genus	read counts	% of reads	Genus	read counts	% of reads
<i>Aliterella</i>	162171	59.455%	<i>Hormoscilla</i>	164275	73.405%	Leptolyngbyaceae_XXXXXX	34123	27.695%
<i>Cyanophyceae (unclassified Genus)</i>	40075	14.692%	<i>Cyanothece</i>	17974	8.032%	<i>Oculatellaceae_XX</i>	30653	24.878%
<i>Cyanothece</i>	31354	11.495%	<i>Aliterella</i>	6551	2.927%	<i>Timaviella</i>	22889	18.577%
<i>Chroococcidiopsidaceae_XX</i>	6605	2.422%	<i>Prochlorococcaceae (unclassified Genus)</i>	4637	2.072%	<i>Hormoscilla</i>	12462	10.114%
<i>Phormidesmis</i>	6293	2.307%	<i>Timaviella</i>	3792	1.694%	<i>Chamaesiphon</i>	7242	5.878%
<i>Hormoscilla</i>	5376	1.971%	<i>Leptolyngbyaceae_XXXXXX</i>	3098	1.384%	<i>Phormidesmis</i>	2803	2.275%
<i>Chroococcidiopsidaceae (unclassified Genus)</i>	3851	1.412%	<i>Gomontiellaceae (unclassified Genus)</i>	2853	1.275%	<i>Cymatolege</i>	2005	1.627%
<i>Timaviella</i>	3204	1.175%	<i>Nostoc</i>	2847	1.272%	<i>Cyanophyceae (unclassified Genus)</i>	1973	1.601%
<i>Oculatellaceae_XX</i>	2923	1.072%	<i>Cyanophyceae (unclassified Genus)</i>	2762	1.234%	<i>Cyanothece</i>	1924	1.562%
<i>Leptolyngbyaceae_XXXXXX</i>	2047	0.750%	<i>Microcoleaceae (unclassified Genus)</i>	2346	1.048%	<i>Aliterella</i>	1487	1.207%
<i>Nostoc</i>	1377	0.505%	<i>Oculatellaceae_XX</i>	2178	0.973%	<i>Gomontiellaceae (unclassified Genus)</i>	1301	1.056%
<i>Chamaesiphon</i>	1375	0.504%	<i>Phormidesmis</i>	2087	0.933%	<i>Shackletoniella</i>	1221	0.991%
<i>Microcoleaceae (unclassified Genus)</i>	1122	0.411%	<i>Heteroleibleinia</i>	1169	0.522%	<i>Apatinema</i>	1152	0.935%
<i>Tolypothrichaceae (unclassified Genus)</i>	1013	0.371%	<i>Chamaesiphon</i>	1166	0.521%	<i>Tychonema</i>	590	0.479%
<i>Prochlorococcaceae (unclassified Genus)</i>	932	0.342%	<i>Microcystis</i>	1055	0.471%	<i>Nostoc</i>	499	0.405%
<i>Tychonema</i>	471	0.173%	<i>Tychonema</i>	1000	0.447%	<i>Microcoleaceae (unclassified Genus)</i>	395	0.321%
<i>Nostocales (unclassified Genus)</i>	450	0.165%	<i>Laspinema</i>	813	0.363%	<i>Microcoleus</i>	102	0.083%
<i>Gomontiellaceae (unclassified Genus)</i>	425	0.156%	<i>Nostocales (unclassified Genus)</i>	681	0.304%	<i>Dolichospermum</i>	99	0.080%

<i>Laspinema</i>	234	0.086%	<i>Tolypothrichaceae</i> (unclassified Genus)	500	0.223%	<i>Chroococcidiopsidaceae</i> (unclassified Genus)	75	0.061%
<i>Crinalium</i>	191	0.070%	<i>Anagnostidinema</i>	465	0.208%	<i>Stenomitos</i>	55	0.045%
<i>Anagnostidinema</i>	184	0.067%	<i>Microcoleus</i>	425	0.190%	<i>Pseudanabaenaceae_X</i>	54	0.044%
<i>Microcoleus</i>	173	0.063%	<i>Cymatolege</i>	234	0.105%	<i>Tolypothrichaceae</i> (unclassified Genus)	52	0.042%
<i>Heteroleibleinia</i>	169	0.062%	<i>Dolichospermum</i>	151	0.067%	<i>Nostocales (unclassified</i> <i>Genus)</i>	35	0.028%
<i>Shackletoniella</i>	153	0.056%	<i>Crinalium</i>	140	0.063%	<i>Microcystis</i>	20	0.016%
<i>Microcystis</i>	146	0.054%	<i>Chroococcidiopsidaceae</i> (unclassified Genus)	136	0.061%	<i>Anagnostidinema</i>	0	0.000%
<i>Stenomitos</i>	135	0.049%	<i>Apatinema</i>	114	0.051%			
<i>Dolichospermum</i>	100	0.037%	<i>Shackletoniella</i>	110	0.049%			
<i>Cymatolege</i>	83	0.030%	<i>Stenomitos</i>	77	0.034%			
<i>Apatinema</i>	49	0.018%	<i>Chroococcidiopsidaceae_XX</i>	76	0.034%			
<i>Cyanobacteriota</i> (unclassified Genus)	45	0.016%	<i>Toxopsis</i>	38	0.017%			
<i>Toxopsis</i>	29	0.011%	<i>Dermocarpella</i>	21	0.009%			
<i>Dermocarpella</i>	6	0.002%	<i>Cyanobacteriota</i> (unclassified Genus)	17	0.008%			
			<i>Aphanizomenon</i>	5	0.002%			

Table S3.6. Cyanobacterial indicator ASVs significantly associated with specific substrate types identified via Indicator Value analysis (IndVal). Taxonomic assignment, indicator index, test statistic (stat), and permutation p-values are shown for ASVs exhibiting significant substrate specificity ($p \leq 0.05$). Binary columns indicate substrate association (1 = presence, 0 = absence). Substrate abbreviations: mo_DV = moraine Dry Valley; mo_YO = moraine Yûboku Valley.

ASV	Genus (or Family)	index	stat	p.value	gneiss	granite	marble	mo_DV	mo_YO
ASV_6	Leptolyngbyaceae	5	0.972916984	0.001	0	0	0	0	1
ASV_285	Cyanobacteriota	1	0.959578918	0.001	1	0	0	0	0
ASV_3	<i>Aliterella</i>	3	0.91865272	0.001	0	0	1	0	0
ASV_208	Chroococcidiopsidaceae	3	0.885219382	0.001	0	0	1	0	0
ASV_7	<i>Cyanothece</i>	1	0.882892039	0.001	1	0	0	0	0
ASV_13	<i>Nostoc</i>	2	0.875089929	0.002	0	1	0	0	0
ASV_29	<i>Timaviella</i>	5	0.850116969	0.005	0	0	0	0	1
ASV_108	<i>Shackletoniella</i>	5	0.841890689	0.005	0	0	0	0	1
ASV_228	<i>Apatinema</i>	5	0.756543725	0.007	0	0	0	0	1
ASV_557	Pseudanabaenaceae	5	0.632455532	0.007	0	0	0	0	1
ASV_9	Oculatellaceae	5	0.919053992	0.009	0	0	0	0	1
ASV_190	Chroococcidiopsidaceae	3	0.713754055	0.016	0	0	1	0	0
ASV_16	Microcoleaceae	2	0.852006954	0.02	0	1	0	0	0
ASV_72	<i>Chamaesiphon</i>	5	0.841924302	0.026	0	0	0	0	1
ASV_141	<i>Toxopsis</i>	2	0.719660387	0.031	0	1	0	0	0
ASV_271	<i>Dolichospermum</i>	5	0.525123643	0.042	0	0	0	0	1

Table S3.7. Relative influence of physicochemical, spatial and bedrock type variables on bacterial communities using a dbRDA and a forward, stepwise variable selection.

Environmental variables	Adj. R2	F	df	P
pH	0.13	10.5	1	0.002
TN	0.18	5.34	1	0.002
N-NO ₃ ⁻	0.22	3.98	1	0.002
EC	0.24	2.42	1	0.002
All	0.26	NA	NA	NA

PCNM	Adj. R2	F	df	P
PCNM1	0.13	10.6	1	0.002
PCNM3	0.18	5.12	1	0.002
PCNM2	0.22	4.02	1	0.002
PCNM10	0.25	3.72	1	0.002
PCNM4	0.27	2.55	1	0.002
PCNM8	0.30	2.62	1	0.002
PCNM9	0.31	2.30	1	0.002
All	0.32	NA	NA	NA

Bedrock type	Adj. R2	F	df	P
marble	0.18	15.4	1	0.002
mo_YO	0.25	6.96	1	0.002
mo_DV	0.29	4.66	1	0.002
granite	0.31	3.20	1	0.002
All	0.31	NA	NA	NA

Research chapters - Part two

Endemism is a key concept from a nature conservation perspective. Defining organisms or communities as endemic provides strong arguments for the development of conservation guidelines (Griffiths, 2012). In this context, refining the concept of endemism makes it possible to raise the level of conservation requirements.

Based on this rationale, we decided to use genomics to better characterize Antarctic cyanobacteria and highlight their genetic specificities. In this second part, we will focus on two specific case studies of cyanobacteria whose genera are considered “cosmopolitan” on a global scale and would therefore not necessarily be regarded as priority “keystone taxa” for conservation. However, we hypothesize that, even when genera are distributed across broad geographic scales, their ecotypes may harbour biological specificities associated with the environment they inhabit. Studying Antarctic ecotypes in greater detail through comparative genomics should enable the characterization of their genetic specificities.

In **Chapter 4**, we compare the Antarctic strain *Phormidesmis priestleyi* ULC007 with non-Antarctic cyanobacteria using a comparative genomic approach based on a predefined set of genes of interest (GOIs). In **Chapter 5**, we investigate a subgroup of picocyanobacteria, subcluster 5.2, which includes both Antarctic and non-Antarctic strains, using a discovery-based approach to identify genes clusters (GCs) specific to the Antarctic strains.

Chapter 4

Cold adaptation and stress response of *Phormidesmis priestleyi* ULC007, a cyanobacterium isolated from Lake Bruehwiler (Larsemann Hills, Antarctica)

Benoit Durieu^{1,2}, Yannick Lara³, Denis Baurain¹, Annick Wilmotte²

¹ InBioS–PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, 4000 Liège, Belgium

² BCCM/ULC, InBioS–Molecular Diversity and Ecology of Cyanobacteria, University of Liège, 4000 Liège, Belgium

³ Early Life Traces & Evolution-Astrobiology, University of Liège, 4000 Liège, Belgium

The first part of this chapter is published in **Genome Announcement** (ASM journals), 2017, Volume 5, No. 7, genomea.01546-16.

Author contribution of BD:

BD co-wrote the genome announcement, performed the comparative genomic analysis and wrote Chapter 4.

4.1. Introduction

We assembled the genome of *Phormidesmis priestleyi* ULC007 and published it as a genome announcement in 2017 (Lara et al., 2017). It was the first Antarctic draft genome assembly to be published. This chapter begins by presenting the core content of the genome announcement as an introduction, followed by the original comparative genomic analysis of this assembly on cold adaptation and stress response, carried out in 2017. The full genome announcement is available in **Section 4.8**. Although this comparative analysis of cold and stress response is now somewhat outdated, it is revisited in the general discussion (**Chapter 6**) in light of the recent literature and complemented by phylogenomic and pangenomic analyses to compare the Antarctic genomic specificities of the *Phormidesmis* cluster with those of the picocyanobacterial cluster described in **Chapter 5**.

4.2. Genome Announcement

Cyanobacteria are key organisms of the aquatic and terrestrial Antarctic food webs. They are the dominant primary producers and generally play a major role in the first steps of the colonization of deglaciated habitats (Vincent, 2002). Thin filamentous cyanobacteria of the orders Oscillatoriales and Synechococcales (Komárek et al., 2014) are abundant in Antarctica and usually form conspicuous mats and crusts in freshwater and terrestrial ecosystems. According to SSU rRNA gene phylogenies, they are clustered in several lineages, for which detailed data about their taxonomy and potential traits (e.g., nitrogen fixation, salinity tolerance) remain scarce. Moreover, no genome of Antarctic cyanobacteria is currently available.

The strain *Phormidesmis priestleyi* ULC007 (Komárek et al., 2009) was originally isolated from Lake Bruehwiler (Taton et al., 2006), also referenced as L52, a shallow freshwater lake of 1 ha (Hodgson et al., 2004). It is located in the Larsemann Hills, where air temperatures vary from -18°C during winter to 10°C during summer (Gillieson, 1990), underlining the high variability of the climatic conditions.

The purification protocol consisted of successive streaking on solidified BG11-based culture medium until obtaining an axenic culture (Rippka, 1988). After DNA extraction using GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich, St. Louis, MO), the genomic DNA of ULC007 was sequenced in two runs. The first run was performed on an Illumina HiSeq 2500 System (Illumina, Inc., San Diego, CA, USA), which generated 4,404,753 paired end reads of 100 bp. The second run,

performed using the Illumina MiSeq System (Illumina Inc, San Diego, CA, USA), generated 440,221 paired end reads of 250 bp.

Raw reads from both runs were processed and de novo assembled using Velvet v1.2.10 (Zerbino & Birney, 2008) and SPAdes v3.5 (Bankevich et al., 2012) after a thorough analysis of the kmer length distribution using kmergenie v1.6982 (Chikhi & Medvedev, 2014). The resulting sets of contigs were integrated using the CISA v1.3 software (Lin & Liao, 2013). The accuracy of the final assembly was evaluated using the REAPR v1.0.18 software (Hunt et al., 2013). Contigs were annotated using glimmer3 v3.02 (Delcher et al., 2007) and a reference database composed of 260 cyanobacterial genomes extracted from the PATRIC database. The presence of secondary metabolite biosynthesis gene clusters was tested using antiSMASH v3.0 (Weber et al., 2015).

The draft genome of ULC007 is composed of 5,684,389 bp distributed into 118 contigs, with an average G+C-content of 48.6%. Final assembly yielded an L50 value of 12 (L90 = 51) and an N50 value of 152,457 bp (N90 = 27,655 bp). A total of 5,604 protein-encoding genes (PEGs) were identified by glimmer3, of which 4,785 PEGs correspond to proteins already characterized in other cyanobacterial genomes, whereas 1,244 PEGs are apparently new. The genome contains a variety of PEGs involved in stress response (e.g., molecular chaperones DnaK and DnaJ, DNA gyrase, putative cold shock proteins of the Csp-family). The main components (KaiA, KaiB, and KaiC) of the cyanobacterial circadian clock were found, as well as PEGs related to the circadian input (CikA, LdpA, Pex) and output (CikA, SasA, RpaA, LabA) pathways (Cohen & Golden, 2015). Finally, three unknown NRPS-, NRPS/PKS-, and PKS-like gene clusters were identified, whose exploration may potentially lead to the discovery of novel secondary peptide metabolites.

This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MPPI00000000. The version described in this paper is version MPPI01000000.

4.3. Methods

For the comparative genomic analysis, the functional annotation of the *Phormidesmis priestleyi* ULC007 assembly was done using the RAST automatic server (The Rapid Annotations using Subsystems; Aziz et al., 2008). RAST uses the SEED framework for comparative genomics (Overbeek et al., 2005), a database of annotated genomes organized at different levels of functional annotations, to identify protein-encoding genes, rRNA and tRNA genes on the target genome. RAST then assigns functions and constructs a metabolic reconstruction resulting in a classification in categories, sub-categories and subsystems of gene roles (Overbeek et al., 2005).

This classification allowed us to select genes involved in the metabolic pathway of cold adaptation and stress response.

4.3.1. Obtention of good quality assemblies and statistics

We selected 72 cyanobacterial assemblies classified as “reference genomes” by NCBI GenBank. Assembly accession numbers (GCA_*) and assembly statistics (e.g., N50, numbers of contigs) were downloaded using NCBI Entrez® Direct (Kans, 2017), with the query “cyanobacteria” and the option “-db assembly”. The assembly GCA_001895925.1 (*Phormidesmis priestleyi* ULC007) and its close Arctic relative GCA_001650195.1 (*Phormidesmis priestleyi* BC1401) were then added to the dataset.

The GTBD and NCBI taxonomy of the assemblies, as well as genome completeness calculated with CheckM (Parks et al., 2015), were obtained from the GTDB database (Parks et al., 2026).

4.3.2. Cold adaptation and stress response genes

We selected genes of interest (GOIs) related to (1) cold adaptation following Barria et al. (2013) and (2) stress response following Dsouza et al. (2015) and Varin et al. (2012). Altogether, this set of genes are involved in the production of chaperones, universal stress proteins, translation initiation factors, and cell wall components, as well as in mechanisms of cold and heat-shock responses, oxidative stress, fatty acid and sugar biosynthesis, DNA replication and reparation, ribosome activity modulation, RNA processing and degradation.

For the comparative analysis between *Phormidesmis priestleyi* ULC007 and the other high-quality assemblies, we used the Network-Based SEED API (Disz et al., 2010), which provides direct access to the SEED database, to retrieve all functional roles and FIGfams (i.e., protein identification based on protein family) associated with selected GOIs across the high-quality assemblies. The data was then curated manually to avoid redundancy and misidentification. The final selected functional roles (Fun) are described in **Table 4.1**.

Table 4.1. (Part 1) Functional roles, from the SEED database, and their related genes involved in cold adaptation and stress responses, from Varin et al. (2012), Barria et al. (2013) and Dsouza et al. (2015).

Code	Functional role	Functional category	Related gene or subsystem	References
Fun01	ABC transporter component, possibly Mn transport	ABC transporters	<i>mntA</i>	Dsouza et al. 2015
Fun02	DEAD-box ATP-dependent RNA helicase CshA (EC 3.6.4.13)	ATP-dependent RNA helicases	<i>deaD</i>	Varin et al. 2012, Barria et al. 2013
Fun03	DEAD/DEAH box helicase-like protein	ATP-dependent RNA helicases	<i>deaD</i>	Varin et al. 2012, Barria et al. 2013
Fun04	Leucine dehydrogenase (EC 1.4.1.9)	Branched-Chain Amino Acid Biosynthesis	<i>bcd</i>	Dsouza et al. 2015
Fun05	Fructose-bisphosphate aldolase class I (EC 4.1.2.13)	Carbohydrate metabolism	<i>fbaA</i>	Dsouza et al. 2015
Fun06	Fructose-bisphosphate aldolase class II (EC 4.1.2.13)	Carbohydrate metabolism	<i>fbaA</i>	Dsouza et al. 2015
Fun07	Lipid A disaccharide synthetase related enzyme	Cell wall components	<i>lpxP</i>	Barria et al. 2013
Fun08	Lipid A export ATP-binding/permease protein MsbA	Cell wall components	<i>lpxP</i>	Barria et al. 2013
Fun09	L-proline glycine betaine ABC transport system permease protein ProV (TC 3.A.1.12.1)	Choline and Betaine Uptake and Betaine Biosynthesis	<i>proV</i>	Dsouza et al. 2015
Fun10	Cold shock protein	Cold-shock proteins	<i>csp</i>	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013
Fun11	RecA protein	DNA repair	<i>recA</i>	Varin et al. 2012, Barria et al. 2013
Fun12	Chromosomal replication initiator protein DnaA	DNA replication	<i>dnaA</i>	Varin et al. 2012, Barria et al. 2013
Fun13	DNA gyrase subunit A (EC 5.99.1.3)	DNA topoisomerase	<i>gyrA</i>	Varin et al. 2012, Barria et al. 2013
Fun14	DNA-binding protein HU	DNA structural proteins	<i>hupB</i>	Barria et al. 2013
Fun15	Glycosyl transferase, group 2 family protein	EPS biosynthesis	EPS	Varin et al. 2012, Dsouza et al. 2015
Fun16	Undecaprenyl-phosphate galactosephosphotransferase (EC 2.7.8.6)	EPS biosynthesis	EPS	Varin et al. 2012, Dsouza et al. 2015
Fun17	Fatty acid desaturase (EC 1.14.19.1)	Fatty acid metabolism	<i>desA</i>	Dsouza et al. 2015
Fun18	Chaperone protein DnaJ	Heat-shock proteins	<i>dnaJ</i>	Varin et al. 2012, Barria et al. 2013
Fun19	Chaperone protein DnaK	Heat-shock proteins	<i>dnaK</i>	Varin et al. 2012, Barria et al. 2013
Fun20	Gamma-glutamyltranspeptidase (EC 2.3.2.2)	Oxidative stress	Oxidative stress	Dsouza et al.
Fun21	Adenylosuccinate synthetase (EC 6.3.4.4)	Purine metabolism	<i>purA</i>	Varin et al. 2012
Fun22	Purine nucleoside phosphorylase (EC 2.4.2.1)	Pyrimidine metabolism	<i>pnp</i>	Varin et al. 2012, Barria et al. 2013
Fun23	Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase complex (EC 2.3.1.12)	Pyruvate metabolism	<i>aceE, aceF</i>	Varin et al. 2012, Barria et al. 2013
Fun24	Pyruvate dehydrogenase E1 component alpha subunit (EC 1.2.4.1)	Pyruvate metabolism	<i>aceE, aceF</i>	Varin et al. 2012, Barria et al. 2013
Fun25	Pyruvate dehydrogenase E1 component beta subunit (EC 1.2.4.1)	Pyruvate metabolism	<i>aceE, aceF</i>	Varin et al. 2012, Barria et al. 2013
Fun26	Pyruvate/2-oxoglutarate dehydrogenase complex, dihydrolipoamide dehydrogenase (E3) component, and related enzymes	Pyruvate metabolism	<i>lpd</i>	Varin et al. 2012, Barria et al. 2013
Fun27	LSU ribosomal protein L27p	Ribosome	Ribosomal proteins	Barria et al. 2013
Fun28	LSU ribosomal protein L31p	Ribosome	Ribosomal proteins	Barria et al. 2013
Fun29	LSU ribosomal protein L32p	Ribosome	Ribosomal proteins	Barria et al. 2013
Fun30	SSU ribosomal protein S15p (S13e)	Ribosome	Ribosomal proteins	Barria et al. 2013
Fun31	tRNA dihydrouridine synthase B	RNA metabolism	RNA metabolism	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013

Table 4.1. (Part 2)

Code	Functional role	Functional category	Related gene or sybsystem	References
Fun32	3'-to-5' exoribonuclease RNase R	RNA processing and degradation	<i>rnr</i>	Barria et al. 2013
Fun33	Transcription termination protein NusA	Transcription factors	<i>nusA, nusB</i>	Barria et al. 2013, Dsouza et al. 2015
Fun34	RNA polymerase sigma factor RpoE	Transcription initiation factors	RNA polymerase	Dsouza et al. 2015
Fun35	Ribosome-binding factor A	Translation initiation factors	<i>rbfA</i>	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013
Fun36	Translation initiation factor 1	Translation initiation factors	<i>infA</i>	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013
Fun37	Translation initiation factor 2	Translation initiation factors	<i>infB</i>	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013
Fun38	Translation initiation factor 3	Translation initiation factors	<i>infC</i>	Varin et al. 2012, Barria et al. 2013
Fun39	Alpha,alpha-trehalose-phosphate synthase [UDP-forming] (EC 2.4.1.15)	Trehalose Biosynthesis	<i>otsA, otsB</i>	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013
Fun40	Trehalose synthase (EC 5.4.99.16)	Trehalose Biosynthesis	<i>otsA, otsB</i>	Varin et al. 2012, Dsouza et al. 2015, Barria et al. 2013
Fun41	Universal stress protein family	Universal stress protein family	<i>tig</i>	Barria et al. 2013
Fun42	Universal stress protein UspA and related nucleotide-binding proteins	Universal stress protein family	<i>tig</i>	Barria et al. 2013

4.3.3. Clustering and statistical analysis

We then computed the number of occurrences of these functional roles (Fun) across the 72 assemblies using the R environment, with the “dcast” function from the “resphape2” package (Wickham, 2007). Based on this data, we constructed a Euclidean distance matrix of functional roles occurrence for the 72 assemblies. Hierarchical clustering was then performed on this matrix to compare patterns of cold adaptation and stress response among genome assemblies, using the “hclust” function from the “stats” package with the “McQuitty” clustering method (McQuitty, 1966). The resulting clustering was displayed alongside a heatmap of functional role occurrences. The arithmetic mean (AM) of functional occurrences across all genome assemblies was computed with standard deviation (sd), using the “mean” and “sd” functions from the “base” package (Becker et al., 1988).

4.4. Results

A total of 72 reference assemblies were retrieved from NCBI (**Table 4.2**). Except for the two polar *Phormidesmis* assemblies, the numbers of contigs ranged from 1 and 19. Completeness ranged from 95.99% and 100%, except for the assembly of Candidatus *Atelocyanobacterium thalassa* isolate ALOHA, which had a completeness of 73.92%. According to GTDB and NCBI taxonomy,

these 72 assemblies correspond to 18 or 26 families, respectively, and 8 or 14 orders, respectively (Parks et al., 2026) (**Table S4.1**).

Table 4.2. (Part 1) Assembly names, NCBI accessions, number of contigs, number of genes, and completeness (%) of the 72 assemblies used in the comparative analysis. Compl. = completeness.

Assembly	Accession	# contigs	# genes	Compl.
<i>Gloeocapsa</i> sp. PCC 7428	GCF_000317555.1	5	5532	99.78
<i>Chroococcidiopsis thermalis</i> PCC 7203	GCF_000317125.1	3	6978	99.63
<i>Microcoleus</i> sp. PCC 7113	GCF_000317515.1	9	7547	99.56
<i>Cyanobacterium stanieri</i> PCC 7202	GCA_000317655.1	1	3008	99.52
<i>Cyanobacterium aponinum</i> PCC 10605	GCF_000317675.1	2	3708	99.45
<i>Oscillatoria acuminata</i> PCC 6304	GCF_000317105.1	3	7495	99.71
<i>Trichodesmium erythraeum</i> IMS101	GCF_000014265.1	1	6500	99.71
<i>Oscillatoria nigro-viridis</i> PCC 7112	GCF_000317475.1	6	8289	99.78
<i>Arthrospira platensis</i> NIES-39	GCF_000210375.1	18	7329	99.13
<i>Cyanothece</i> sp. ATCC 51142	GCF_000017845.1	6	5159	99.96
<i>Cyanothece</i> sp. PCC 8801	GCF_000021805.1	4	4651	99.56
<i>Cyanothece</i> sp. PCC 7424	GCF_000021825.1	7	6047	99.71
<i>Cand. Atelocyanobacterium thalassa</i> is. ALOHA	GCF_000025125.1	1	1200	73.92
<i>Cyanothece</i> sp. PCC 7822	GCF_000147335.1	7	7392	99.82
<i>Microcystis aeruginosa</i> NIES-843	GCF_000010625.1	1	5777	99.89
<i>Pleurocapsa</i> sp. PCC 7327	GCF_000317025.1	1	5223	99.49
<i>Synechocystis</i> sp. PCC 6803	GCF_000009725.1	5	7862	99.78
<i>Synechococcus</i> sp. PCC 7002	GCF_000019485.1	7	3340	100
<i>Leptolyngbya</i> sp. PCC 7376	GCF_000316605.1	1	5200	99.42
<i>Calothrix</i> sp. PCC 7507	GCF_000316575.1	1	6927	99.11
<i>Calothrix</i> sp. PCC 6303	GCF_000317435.1	4	6332	99.76
<i>Nostoc</i> sp. PCC 7120	GCF_000009705.1	7	6694	99.19
<i>Nostoc azollae</i> 0708	GCF_000196515.1	3	7014	98.89
<i>Anabaena variabilis</i> ATCC 29413	GCF_000204075.1	5	6436	99.33
<i>Anabaena</i> sp. 90	GCF_000312705.1	5	5377	99.67
<i>Nostoc</i> sp. PCC 7107	GCF_000316625.1	1	5928	99.26
<i>Anabaena cylindrica</i> PCC 7122	GCF_000317695.1	7	6726	99.44
<i>Rivularia</i> sp. PCC 7116	GCF_000316665.1	3	7389	99.78
<i>Crinalium epipsammum</i> PCC 9333	GCF_000317495.1	9	5406	99.48
<i>Halotheca</i> sp. PCC 7418	GCF_000317635.1	1	4076	99.48
<i>Dactylococcopsis salina</i> PCC 8305	GCF_000317615.1	1	4168	99.55
<i>Staniera cyanosphaera</i> PCC 7437	GCF_000317575.1	6	5163	99.56
<i>Gloeobacter violaceus</i> PCC 7421	GCF_000011385.1	1	4459	99.15
<i>Phormidesmis priestleyi</i> BC1401	GCF_001650195.1	217	5869	99.29
<i>Phormidesmis priestleyi</i> ULC007	GCF_001895925.1	137	5782	95.99
<i>Leptolyngbya boryana</i> PCC 6306	GCF_000353285.1	19	7442	99.34
<i>Prochlorococcus marinus</i> str. CCMP1375	GCF_000007925.1	1	1986	100

Table 4.2. (Part 2)

Assembly	Accession	# contigs	# genes	Compl.
<i>Prochlorococcus marinus</i> str. CCMP1986	GCF_000011465.1	1	1957	99.46
<i>Prochlorococcus marinus</i> str. MIT 9313	GCF_000011485.1	1	2873	99.46
<i>Prochlorococcus marinus</i> str. NATL2A	GCF_000012465.1	1	2203	99.45
<i>Prochlorococcus marinus</i> str. MIT 9312	GCF_000012645.1	1	1985	99.73
<i>Prochlorococcus marinus</i> str. AS9601	GCF_000015645.1	1	1936	99.64
<i>Prochlorococcus marinus</i> str. MIT 9515	GCF_000015665.1	1	1957	100
<i>Prochlorococcus marinus</i> str. NATL1A	GCF_000015685.1	1	2248	99.73
<i>Prochlorococcus marinus</i> str. MIT 9303	GCF_000015705.1	1	3083	100
<i>Prochlorococcus marinus</i> str. MIT 9301	GCF_000015965.1	1	1831	99.46
<i>Prochlorococcus marinus</i> str. MIT 9215	GCF_000018065.1	1	2047	99.73
<i>Prochlorococcus marinus</i> str. MIT 9211	GCF_000018585.1	2	1961	99.73
<i>Prochlorococcus marinus</i> str. MIT 9202	GCF_000158595.1	6	2009	98.78
<i>Synechococcus</i> sp. WH 8102	GCF_000195975.1	1	2690	99.46
<i>Cyanobium gracile</i> PCC 6307	GCF_000316515.1	1	3262	99.73
<i>Synechococcus</i> sp. CC9902	GCF_000012505.1	1	2561	99.46
<i>Synechococcus</i> sp. CC9605	GCF_000012625.1	1	2858	99.73
<i>Synechococcus</i> sp. CC9311	GCF_000014585.1	1	3058	99.73
<i>Synechococcus</i> sp. WH 7803	GCA_000063505.1	1	2616	99.18
<i>Synechococcus</i> sp. RCC307	GCA_000063525.1	1	2546	99.64
<i>Synechococcus</i> sp. RS9917	GCF_000153065.1	9	2715	99.46
<i>Synechococcus</i> sp. BL107	GCF_000153805.1	6	2605	99.46
<i>Synechococcus</i> sp. RS9916	GCF_000153825.1	4	2799	99.73
<i>Synechococcus</i> sp. WH 8109	GCF_000161795.2	1	2438	99.46
<i>Geitlerinema</i> sp. PCC 7407	GCF_000317045.1	1	4313	99.87
<i>Pseudanabaena</i> sp. PCC 7367	GCF_000317065.1	2	4619	98.23
<i>Synechococcus</i> sp. PCC 7502	GCF_000317085.1	3	3804	99.76
<i>Synechococcus elongatus</i> PCC 6301	GCF_000010065.1	1	2868	99.73
<i>Synechococcus elongatus</i> PCC 7942	GCF_000012525.1	2	2915	100
<i>Synechococcus</i> sp. JA-3-3Ab	GCF_000013205.1	1	2847	100
<i>Synechococcus</i> sp. JA-2-3B'a (2-13)	GCF_000013225.1	1	2947	100
<i>Acaryochloris marina</i> MBIC11017	GCF_000018105.1	10	7710	99.53
<i>Thermosynechococcus elongatus</i> BP-1	GCF_000011345.1	1	2594	99.76
<i>Synechococcus</i> sp. PCC 6312	GCF_000316685.1	2	3839	99.29
<i>Cyanothece</i> sp. PCC 8802	GCF_000024045.1	5	4694	/
<i>Cyanothece</i> sp. PCC 7425	GCF_000022045.1	4	5440	/

4.4.1. Cold adaptation and stress response of *Phormidesmis priestleyi* ULC007

The *Phormidesmis priestleyi* ULC007 assembly contains most of the genes involved in cold adaptation and stress response, including molecular chaperones DnaK and DnaJ, DNA gyrase, and putative cold shock proteins of the Csp family. *Phormidesmis priestleyi* ULC007 clusters with the other polar representative, *Phormidesmis priestleyi* BC1401 (**Fig. 4.1**). Although phylogenetically closely related genera cluster together, such as *Prochlorococcus* and *Synechococcus*, the overall clustering pattern does not follow the species phylogeny, such as that proposed by Komárek et al. (2014).

Like most other cyanobacterial genomes, *Phormidesmis priestleyi* ULC007 lacks **Fun4** (Leucine dehydrogenase), **Fun10** (Cold shock protein), **Fun22** (Purine nucleoside phosphorylase), **Fun39** (Alpha,alpha-trehalose-phosphate synthase [UDP-forming]), and **Fun40** (Trehalose synthase) (**Fig. 4.1**). Interestingly, it contains **Fun3** (DEAD/DEAH box helicase-like protein), **Fun7** (Lipid A disaccharide synthetase related enzyme), **Fun8** (Lipid A export ATP-binding/permease protein MsbA), **Fun9** (L-proline glycine betaine ABC transport system permease protein ProV), **Fun15** (Glycosyl transferase (EPS biosynthesis)), **Fun41** (Universal stress protein family) and **Fun42** (Universal stress protein UspA and related nucleotide-binding proteins), which are recovered in a few genomes only (**Fig. 4.1** and **Fig. 4.2**).

Comparison of functional roles (Fun) copy numbers shows that **Fun06** (Fructose-bisphosphate aldolase class II, fbaA), **Fun16** (Undecaprenyl-phosphate galactosephosphotransferase, involved in EPS biosynthesis), **Fun19** (heat-shock protein DnaK), and **Fun42** (universal stress protein UspA) have a higher copy number in polar genome assemblies than in non-polar genome assemblies (**Fig. 4.2**).



Fig. 4.1. Clustering and heatmap of the 42 functional roles (Fun) involved in the cold adaptation and stress response across the 72 genome assemblies.

When the different functional roles are grouped into their respective functional category (Sub), Sub02 (Choline and Betaine Biosynthesis), Sub05 (DNA repair), Sub09 (EPS biosynthesis), and Sub16 (Heat shock DnaK genes cluster) appear to be more represented in polar genome assemblies than in non-polar genome assemblies (**Fig. 4.3**).

To account for the phylogenetic signal, we grouped the different genome assemblies by phylogenetic family. Within the Leptolyngbyaceae, the two polar *Phormidesmis* harbour more genes related to DNA repair, choline and betaine biosynthesis, heat shock response (DnaK gene cluster) and exopolysaccharide (EPS) biosynthesis (**Fig. 4.4**).

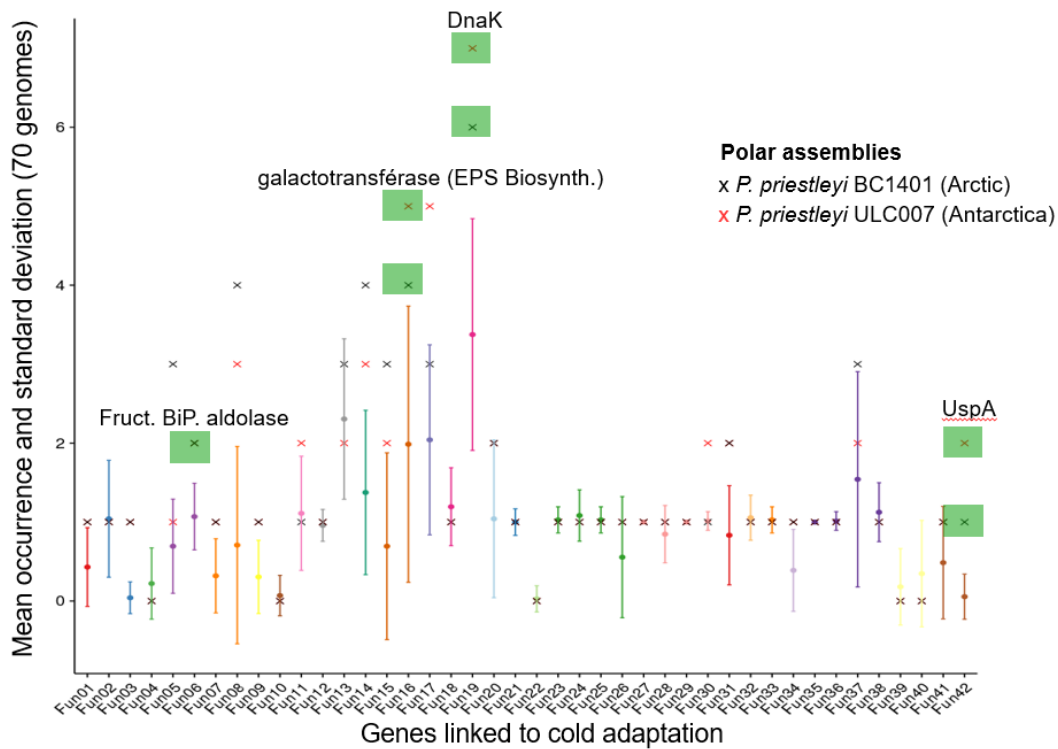


Fig. 4.2. Comparison of the mean number of occurrences of the 42 functional roles (FUN) involved in cold adaptation and stress response across the 72 assemblies, with standard deviations (vertical lines), and the corresponding numbers of occurrences in the two polar assemblies (black and red crosses).

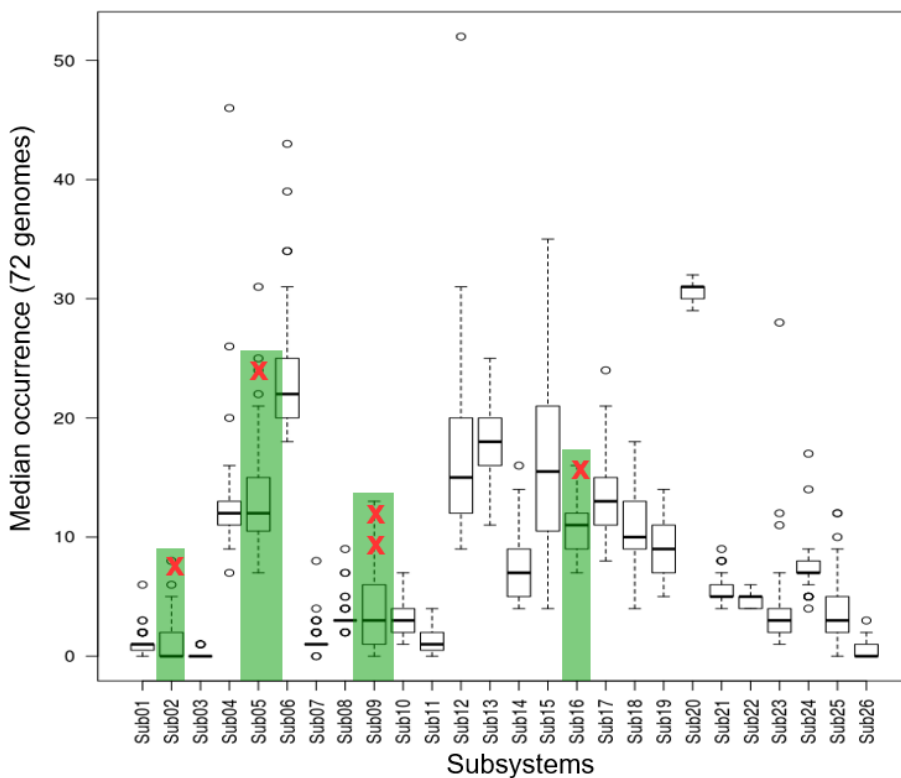


Fig. 4.3. Boxplots showing the distribution of the number of functional roles (FUN) within each functional category (Sub) across the 72 assemblies. The median is represented by the horizontal line within each box. Red crosses indicate the values for the two polar assemblies when they exceed the median.

with the other polar representative, *Phormidesmis priestleyi* BC1401, whereas the overall clustering pattern does not follow the phylogenetic signal. This result indicates that the two polar genomes share similar gene content related to cold adaptation and stress responses, which is expected from the similarity of environmental pressures between the Arctic and the Antarctica, as cold-adapted cyanobacterial lineages have repeatedly evolved across polar and alpine environments (Jungblut et al., 2010; Christmas et al., 2015).

Phormidesmis priestleyi ULC007 also contains specific genes related to cold adaptation and stress responses that are not recovered in the majority of the other genomes. Overall, as well as within the Leptolyngbyaceae family, the two polar assemblies contain higher copy numbers of genes involved in exopolysaccharide (EPS) biosynthesis, compatible solutes (e.g., glycine betaine) production, cold and stress responses, and DNA repair mechanisms. EPS biosynthesis is an adaptation that buffers the impact of freeze-thaw cycles by limiting the ice nucleation around cells (Vincent, 2007). Similarly, compatible solutes reduce the freezing point of intracellular fluids and limit cell desiccation by reducing the quantity of water required to maintain osmotic equilibrium under increasing salinity (Welsh, 2000; Zakhia et al., 2008). The molecular chaperone DnaK is induced under cold conditions and, together with DnaJ, helps maintain the native conformation of regulatory proteins such as DnaA, GyrA and RecA (DNA repair), which influence DNA topology (Varin et al., 2012). In the general discussion (**Chapter 6**), we will address again these results in light of the most recent literature.

The main limitation of this chapter is that the dataset was small, allowing us to compare only distantly related genomes. This makes it difficult to distinguish the ecological signal from the long-term evolutionary signal. For this reason, the next chapter will focus on a better-sampled cyanobacterial group, the *Synechococcus*-like picocyanobacteria subcluster 5.2, which includes widely distributed organisms ranging from oceanic to polar ecosystems.

Finally, in the general discussion (**Chapter 6**), we will compare the results of Chapters 4 and 5 by applying the same type of pangenomic analysis used in Chapter 5 to the closest relatives of *Phormidesmis priestleyi* ULC007. This approach will enable the characterization and comparison of Antarctic genomic specificities between the two clades.

4.6. References

- Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., ... Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9(1), 75. <https://doi.org/10.1186/1471-2164-9-75>
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Pribelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tester, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *Journal of Computational Biology*, 19(5), 455–477. <https://doi.org/10.1089/cmb.2012.0021>
- Barria, C., Malecki, M., & Arraiano, C. M. (2013). Bacterial adaptation to cold. *Microbiology*, 159(Pt_12), 2437–2443. <https://doi.org/10.1099/mic.0.052209-0>
- Becker, R. A., Chambers, J. M., & Wilks, A. R. (1988). *The New S Language, A Programming Environment for Data Analysis and Graphics* (Wadsworth and Brooks). (Original work published Springer)
- Chikhi, R., & Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. *Bioinformatics*, 30(1), 31–37. <https://doi.org/10.1093/bioinformatics/btt310>
- Christmas, N. A. M., Anesio, A. M., & Sánchez-Baracaldo, P. (2015). Multiple adaptations to polar and alpine environments within cyanobacteria: A phylogenomic and Bayesian approach. *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.01070>
- Cohen, S. E., & Golden, S. S. (2015). Circadian Rhythms in Cyanobacteria. *Microbiology and Molecular Biology Reviews*, 79(4), 373–385. <https://doi.org/10.1128/mmbr.00036-15>
- Delcher, A. L., Bratke, K. A., Powers, E. C., & Salzberg, S. L. (2007). Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics*, 23(6), 673–679. <https://doi.org/10.1093/bioinformatics/btm009>
- Disz, T., Akhter, S., Cuevas, D., Olson, R., Overbeek, R., Vonstein, V., Stevens, R., & Edwards, R. A. (2010). Accessing the SEED Genome Databases via Web Services API: Tools for Programmers. *BMC Bioinformatics*, 11(1), 319. <https://doi.org/10.1186/1471-2105-11-319>
- Dsouza, M., Taylor, M. W., Turner, S. J., & Aislabie, J. (2015). Genomic and phenotypic insights into the ecology of *Arthrobacter* from Antarctic soils. *BMC Genomics*, 16(1), 36. <https://doi.org/10.1186/s12864-015-1220-2>
- Gillieson, D. S. (1990). *An atlas of the lakes of the Larsemann Hills, Princess Elizabeth Land, Antarctica*. Antarctic Division, Dept. of the Arts, Sport, the Environment, Tourism and Territories. <https://nla.gov.au/nla.obj-2331557262/view>

- Hodgson, D. A., Vyverman, W., Verleyen, E., Sabbe, K., Leavitt, P. R., Taton, A., Squier, A. H., & Keely, B. J. (2004). Environmental factors influencing the pigment composition of in situ benthic microbial communities in east Antarctic lakes. *Aquatic Microbial Ecology*, 37, 247–263. <https://doi.org/10.3354/ame037247>
- Hunt, M., Kikuchi, T., Sanders, M., Newbold, C., Berriman, M., & Otto, T. D. (2013). REAPR: A universal tool for genome assembly evaluation. *Genome Biology*, 14(5), R47. <https://doi.org/10.1186/gb-2013-14-5-r47>
- Jungblut, A. D., Lovejoy, C., & Vincent, W. F. (2010). Global distribution of cyanobacterial ecotypes in the cold biosphere. *The ISME Journal*, 4(2), 191–202. <https://doi.org/10.1038/ismej.2009.113>
- Kans, J. (2017). Entrez® Direct: E-utilities on the Unix Command Line. In *Entrez® Programming Utilities Help [Internet]*. National Center for Biotechnology Information (US).
- Komárek, J., Kaštokský, J., Mareš, J., & Johansen, J. R. (2014). Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. *Preslia*, 86, 295–335.
- Komárek, J., Kaštokský, J., Ventura, S., & Turicchia, S. Š. (2009). The cyanobacterial genus *Phormidesmis*. *Algological Studies*, 41–59. <https://doi.org/10.1127/1864-1318/2009/0129-0041>
- Lara, Y., Durieu, B., Cornet, L., Verlaine, O., Rippka, R., Pessi, I. S., Misztak, A., Joris, B., Javaux, E. J., Baurain, D., & Wilmotte, A. (2017). Draft Genome Sequence of the Axenic Strain *Phormidesmis priestleyi* ULC007, a Cyanobacterium Isolated from Lake Bruehwiler (Larsemann Hills, Antarctica). *Genome Announcements*, 5(7), e01546-16. <https://doi.org/10.1128/genomeA.01546-16>
- Lin, S.-H., & Liao, Y.-C. (2013). CISA: Contig Integrator for Sequence Assembly of Bacterial Genomes. *PLOS ONE*, 8(3), e60843. <https://doi.org/10.1371/journal.pone.0060843>
- McQuitty, L. L. (1966). Similarity Analysis by Reciprocal Pairs for Discrete and Continuous Data. *Educational and Psychological Measurement*, 26(4), 825–831. <https://doi.org/10.1177/001316446602600402>
- Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H.-Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., ... Vonstein, V. (2005). The Subsystems Approach to Genome Annotation and its Use in the Project to Annotate 1000 Genomes. *Nucleic Acids Research*, 33(17), 5691–5702. <https://doi.org/10.1093/nar/gki866>
- Parks, D. H., Chaumeil, P.-A., Mussig, A. J., Rinke, C., Chuvochina, M., & Hugenholtz, P. (2026). GTDB release 10: A complete and systematic taxonomy for 715 230 bacterial and 17 245 archaeal genomes. *Nucleic Acids Research*, 54(D1), D743–D754. <https://doi.org/10.1093/nar/gkaf1040>

- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Research*, 25(7), 1043–1055. <https://doi.org/10.1101/gr.186072.114>
- Rippka, R. (1988). Isolation and purification of cyanobacteria. In *Cyanobacteria* (Vol. 167, pp. 3–27). Academic Press. [https://doi.org/10.1016/0076-6879\(88\)67004-2](https://doi.org/10.1016/0076-6879(88)67004-2)
- Taton, A., Grubisic, S., Ertz, D., Hodgson, D. A., Piccardi, R., Biondi, N., Tredici, M. R., Mainini, M., Losi, D., Marinelli, F., & Wilmotte, A. (2006). Polyphasic study of Antarctic cyanobacterial strains. *Journal of Phycology*, 42(6), 1257–1270. <https://doi.org/10.1111/j.1529-8817.2006.00278.x>
- Varin, T., Lovejoy, C., Jungblut, A. D., Vincent, W. F., & Corbeil, J. (2012). Metagenomic Analysis of Stress Genes in Microbial Mat Communities from Antarctica and the High Arctic. *Applied and Environmental Microbiology*, 78(2), 549–559. <https://doi.org/10.1128/AEM.06354-11>
- Vincent, W. F. (2002). Cyanobacterial Dominance in the Polar Regions. In B. A. Whitton & M. Potts (Eds.), *The Ecology of Cyanobacteria: Their Diversity in Time and Space* (pp. 321–340). Springer Netherlands. https://doi.org/10.1007/0-306-46855-7_12
- Vincent, W. F. (2007). Cold Tolerance in Cyanobacteria and Life in the Cryosphere. In J. Seckbach (Ed.), *Algae and Cyanobacteria in Extreme Environments* (pp. 287–301). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6112-7_15
- Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Brucoleri, R., Lee, S. Y., Fischbach, M. A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., & Medema, M. H. (2015). antiSMASH 3.0—A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Research*, 43(W1), W237–W243. <https://doi.org/10.1093/nar/gkv437>
- Welsh, D. T. (2000). Ecological significance of compatible solute accumulation by micro-organisms: From single cells to global climate. *FEMS Microbiology Reviews*, 24(3), 263–290. <https://doi.org/10.1111/j.1574-6976.2000.tb00542.x>
- Wickham, H. (2007). Reshaping Data with the reshape Package. *Journal of Statistical Software*, 21, 1–20. <https://doi.org/10.18637/jss.v021.i12>
- Zakhia, F., Jungblut, A.-D., Taton, A., Vincent, W. F., & Wilmotte, A. (2008). Cyanobacteria in Cold Ecosystems. In R. Margesin, F. Schinner, J.-C. Marx, & C. Gerday (Eds.), *Psychrophiles: From Biodiversity to Biotechnology* (pp. 121–135). Springer. https://doi.org/10.1007/978-3-540-74335-4_8
- Zerbino, D. R., & Birney, E. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research*, 18(5), 821–829. <https://doi.org/10.1101/gr.074492.107>

4.7. Supplementary Materials

Table S4.1. GTDB and NCBI taxonomy of the 72 genome assemblies.

Assembly name	GTDB_order	GTDB_family	NCBI_order	NCBI_family
<i>Gloeocapsa</i> sp. PCC 7428	Cyanobacteriales	Chroococcidiopsidaceae	Chroococcales	Chroococcaceae
<i>Chroococcidiopsis thermalis</i> PCC 7203	Cyanobacteriales	Chroococcidiopsidaceae	Chroococcidiopsidales	Chroococcidiopsidaceae
<i>Microcoleus</i> sp. PCC 7113	Cyanobacteriales	Coleofasciculaceae	Coleofasciculales	Coleofasciculaceae
<i>Cyanobacterium stanieri</i> PCC 7202	Cyanobacteriales	Cyanobacteriaceae	Chroococcales	Geminocystaceae
<i>Cyanobacterium aponinum</i> PCC 10605	Cyanobacteriales	Cyanobacteriaceae	Chroococcales	Geminocystaceae
<i>Oscillatoria acuminata</i> PCC 6304	Cyanobacteriales	Laspinemataceae	Oscillatoriales	Oscillatoriaceae
<i>Trichodesmium erythraeum</i> IMS101	Cyanobacteriales	Microcoleaceae	Oscillatoriales	Microcoleaceae
<i>Oscillatoria nigro-viridis</i> PCC 7112	Cyanobacteriales	Microcoleaceae	Oscillatoriales	Oscillatoriaceae
<i>Arthrospira platensis</i> NIES-39	Cyanobacteriales	Microcoleaceae	Oscillatoriales	Sirenicapillariaceae
<i>Cyanothece</i> sp. ATCC 51142	Cyanobacteriales	Microcystaceae	Chroococcales	Aphanothecaceae
<i>Cyanothece</i> sp. PCC 8801	Cyanobacteriales	Microcystaceae	Chroococcales	Aphanothecaceae
<i>Cyanothece</i> sp. PCC 7424	Cyanobacteriales	Microcystaceae	Chroococcales	Aphanothecaceae
Cand. <i>Atelocyanobacterium thalassa</i> is. ALOHA	Cyanobacteriales	Microcystaceae	Chroococcales	Aphanothecaceae
<i>Cyanothece</i> sp. PCC 7822	Cyanobacteriales	Microcystaceae	Chroococcales	Aphanothecaceae
<i>Microcystis aeruginosa</i> NIES-843	Cyanobacteriales	Microcystaceae	Chroococcales	Microcystaceae
<i>Pleurocapsa</i> sp. PCC 7327	Cyanobacteriales	Microcystaceae	Pleurocapsales	Hyellaceae
<i>Synechocystis</i> sp. PCC 6803	Cyanobacteriales	Microcystaceae	Synechococcales	Merismopediaceae
<i>Synechococcus</i> sp. PCC 7002	Cyanobacteriales	MRBY01	Chroococcales	Geminocystaceae
<i>Leptolyngbya</i> sp. PCC 7376	Cyanobacteriales	MRBY01	Chroococcales	Geminocystaceae
<i>Calothrix</i> sp. PCC 7507	Cyanobacteriales	Nostocaceae	Nostocales	Calotrichaceae
<i>Calothrix</i> sp. PCC 6303	Cyanobacteriales	Nostocaceae	Nostocales	Calotrichaceae
<i>Nostoc</i> sp. PCC 7120	Cyanobacteriales	Nostocaceae	Nostocales	Nostocaceae
<i>Nostoc azollae</i> 0708	Cyanobacteriales	Nostocaceae	Nostocales	Nostocaceae
<i>Anabaena variabilis</i> ATCC 29413	Cyanobacteriales	Nostocaceae	Nostocales	Nostocaceae
<i>Anabaena</i> sp. 90	Cyanobacteriales	Nostocaceae	Nostocales	Nostocaceae
<i>Nostoc</i> sp. PCC 7107	Cyanobacteriales	Nostocaceae	Nostocales	Nostocaceae
<i>Anabaena cylindrica</i> PCC 7122	Cyanobacteriales	Nostocaceae	Nostocales	Nostocaceae
<i>Rivularia</i> sp. PCC 7116	Cyanobacteriales	Nostocaceae	Nostocales	Rivulariaceae
<i>Crinalium epipsammum</i> PCC 9333	Cyanobacteriales	PCC-9333	Gomontiellales	Gomontiellaceae
<i>Halothece</i> sp. PCC 7418	Cyanobacteriales	Rubidibacteraceae	Chroococcales	Halotheceae
<i>Dactylococcopsis salina</i> PCC 8305	Cyanobacteriales	Rubidibacteraceae	Nodosilineales	Cymatolegaceae
<i>Stanieria cyanosphaera</i> PCC 7437	Cyanobacteriales	Xenococcaceae	Pleurocapsales	Dermocarpellaceae
<i>Gloeobacter violaceus</i> PCC 7421	Gloeobacterales	Gloeobacteraceae	Gloeobacterales	Gloeobacteraceae
<i>Phormidesmis priestleyi</i> BC1401	Leptolyngbyales	Leptolyngbyaceae	Leptolyngbyales	Leptolyngbyaceae
<i>Phormidesmis priestleyi</i> ULC007	Leptolyngbyales	Leptolyngbyaceae	Leptolyngbyales	Leptolyngbyaceae
<i>Leptolyngbya boryana</i> PCC 6306	Leptolyngbyales	Leptolyngbyaceae	Leptolyngbyales	Leptolyngbyaceae
<i>Prochlorococcus marinus</i> str. CCMP1375	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. CCMP1986	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9313	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. NATL2A	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae

<i>Prochlorococcus marinus</i> str. MIT 9312	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. AS9601	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9515	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. NATL1A	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9303	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9301	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9215	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9211	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Prochlorococcus marinus</i> str. MIT 9202	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Synechococcus</i> sp. WH 8102	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Cyanobium gracile</i> PCC 6307	PCC-6307	Cyanobiaceae	Synechococcales	Prochlorococcaceae
<i>Synechococcus</i> sp. CC9902	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. CC9605	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. CC9311	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. WH 7803	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. RCC307	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. RS9917	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. BL107	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. RS9916	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. WH 8109	PCC-6307	Cyanobiaceae	Synechococcales	Synechococcaceae
<i>Geitlerinema</i> sp. PCC 7407	PCC-7407	PCC-7407	Geitlerinematales	Geitlerinemataceae
<i>Pseudanabaena</i> sp. PCC 7367	Pseudanabaenales	Pseudanabaenaceae	Pseudanabaenales	Thalassoporaceae
<i>Synechococcus</i> sp. PCC 7502	Pseudanabaenales	Pseudanabaenaceae	Synechococcales	Synechococcaceae
<i>Synechococcus elongatus</i> PCC 6301	Synechococcales	Synechococcaceae	Synechococcales	Synechococcaceae
<i>Synechococcus elongatus</i> PCC 7942	Synechococcales	Synechococcaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. JA-3-3Ab	Thermotichales	Thermotichaceae	Synechococcales	Synechococcaceae
<i>Synechococcus</i> sp. JA-2-3B'a (2-13)	Thermotichales	Thermotichaceae	Synechococcales	Synechococcaceae
<i>Acaryochloris marina</i> MBIC11017	Thermosynechococcales	Thermosynechococcaceae	Acaryochloridales	Acaryochloridaceae
<i>Thermosynechococcus elongatus</i> BP-1	Thermosynechococcales	Thermosynechococcaceae	Acaryochloridales	Thermosynechococcaceae
<i>Synechococcus</i> sp. PCC 6312	Thermosynechococcales	Thermosynechococcaceae	Synechococcales	Synechococcaceae
<i>Cyanothece</i> sp. PCC 8802	/	/	/	/
<i>Cyanothece</i> sp. PCC 7425	/	/	/	/

4.8. Genome announcement

Draft Genome Sequence of the Axenic Strain *Phormidesmis priestleyi* ULC007, a Cyanobacterium Isolated from Lake Bruehwiler (Larsemann Hills, Antarctica)

Yannick Lara¹, Benoit Durieu¹, Luc Cornet^{2,c}, Olivier Verlaine¹, Rosmarie Rippka⁴, Igor S. Pessi¹, Agnieszka Misztak², Bernard Joris¹, Emmanuelle J. Javaux³, Denis Baurain², Annick Wilmotte^{1,5}

¹ InBioS—Centre for Protein Engineering, University of Liège, Liège, Belgium

² InBioS—PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, Liège, Belgium

³ UR-Geology—Palaeo-Biogeology-Botany-Palynology, University of Liège, Liège, Belgium

⁴ Institut Pasteur, Unité des Cyanobactéries, Centre National de la Recherche Scientifique (CNRS) Unité de Recherche Associée (URA) 2172, Paris, France

⁵ BCCM/ULC, University of Liège, Liège, Belgium

Abstract

Phormidesmis priestleyi ULC007 is an Antarctic freshwater cyanobacterium. Its draft genome is 5,684,389 bp long. It contains a total of 5,604 protein-encoding genes, of which 22.2% have no clear homologues in known genomes. To date, this draft genome is the first one ever determined for an axenic cyanobacterium from Antarctica.

Genome announcement

Cyanobacteria are key organisms of the aquatic and terrestrial Antarctic food webs. They are the dominant primary producers and generally play a major role in the first steps of the colonization of deglaciated habitats (Vincent, 2002). Thin filamentous cyanobacteria of the orders *Oscillatoriales* and *Synechococcales* (Komárek et al., 2014) are abundant in Antarctica and usually form conspicuous mats and crusts in freshwater and terrestrial ecosystems. According to SSU rRNA gene phylogenies, they are clustered in several lineages, for which detailed data about their taxonomy and potential traits (e.g., nitrogen fixation, salinity tolerance) remain scarce. Moreover, no genome of Antarctic cyanobacteria is currently available.

The strain *Phormidesmis priestleyi* ULC007 (Komárek et al., 2009) was originally isolated from Lake Bruehwiler (Taton et al., 2006), a shallow freshwater lake of 1 ha (Hodgson et al., 2004). It is located in the Larsemann Hills, where air temperatures vary from -18°C during winter to 10°C during summer (Gillieson, 1990), underlining the high variability of the climatic conditions.

The purification protocol consisted of successive streaking on solidified BG11-based culture medium until obtaining an axenic culture (Rippka, 1988). The genomic DNA of ULC007 was sequenced in two runs. The first run was performed on an Illumina HiSeq 2500 System (Illumina, Inc., San Diego, CA, USA), which generated 4,404,753 paired-end reads of 100 bp. The second run, performed using the Illumina MiSeq System (Illumina Inc, San Diego, CA, USA), generated 440,221 paired-end reads of 250 bp.

Raw reads from both runs were processed and de novo assembled using Velvet v1.2.10 (Zerbino & Birney, 2008) and SPAdes v3.5 (Bankevich et al., 2012) after a thorough analysis of the kmer length distribution using kmergenie v1.6982 (Chikhi & Medvedev, 2014). The resulting sets of contigs were integrated using the CISA v1.3 software (Lin & Liao, 2013). The accuracy of the final assembly was evaluated using the REAPR v1.0.18 software (Hunt et al., 2013). Contigs were annotated using glimmer3 v3.02 (Delcher et al., 2007) and a reference database composed of 260 cyanobacterial genomes extracted from the PATRIC database. The presence of secondary metabolite biosynthesis gene clusters was tested using antiSMASH v3.0 (Weber et al., 2015).

The draft genome of ULC007 is composed of 5,684,389 bp distributed into 118 contigs, with an average G+C-content of 48.6%. Final assembly yielded an L_{50} value of 12 ($L_{90} = 51$) and an N_{50} value of 152,457 bp ($N_{90} = 27,655$ bp). A total of 5,604 protein-encoding genes (PEGs) were identified by glimmer3, of which 4,785 PEGs correspond to proteins already characterized in other cyanobacterial genomes, whereas 1,244 PEGs are apparently new. The genome contains a variety of PEGs involved in stress response (e.g., molecular chaperones DnaK and DnaJ, DNA gyrase, putative cold shock proteins of the Csp-family). The main components (KaiA, KaiB, and KaiC) of the cyanobacterial circadian clock were found, as well as PEGs related to the circadian input (CikA, LdpA, Pex) and output (CikA, SasA, RpaA, LabA) pathways (Cohen & Golden, 2015). Finally, three unknown NRPS-, NRPS/PKS-, and PKS-like gene clusters were identified, whose exploration may potentially lead to the discovery of novel secondary peptide metabolites.

This genome is the first Antarctic cyanobacterial genome ever determined, and undoubtedly will contribute to our understanding of the mechanisms underpinning the survival and ecological success of cyanobacteria in high latitudes.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession MPPI00000000. The version described in this paper is version MPPI01000000.

Acknowledgements

We thank Dominic Hodgson (British Antarctic Survey) for providing the Antarctic mat sample and Arnaud Taton and Stana Grubisic for the strain isolation.

This work was supported by the BELSPO project CCAMBIO (SD/BA/03A). I.S.P. and L.C. had a PhD FRIA fellowship from the FRS–FNRS (National Fund for Scientific Research), and now L.C. has a PhD fellowship from Planet TOPERS IUAP (Inter-university attraction poles), BELSPO project (Belgian Science Policy). E.J. has funding from Planet TOPERS IUAP and ERC StG ELITE. A.M. was partially financed by the bilateral Scientific and technical cooperation between the Republic of Poland and French-speaking Community and the Walloon region of Belgium (2014–2016). A.W. is a research associate of the FRS–FNRS. This work was mostly carried out on the workstation “Brutus” funded by the Actions de Recherche Concertées (ARC) (NetRBI 2012–2017) of the Belgium French Community, but also benefited from computational resources made available on the “durandal” grid computer funded, among others, by a grant from the University of Liège, “Fonds spéciaux pour la recherche,” “Crédit de démarrage 2012” (SFRD-12/04), and by a grant from the FRS–FNRS “Crédit de recherche 2014” (CDR J.0080.15).

R.R. retired from the Pasteur Institute in 2009 and contributed to this work in the form of a collaboration with A.W. (InBioS-Centre for Protein Engineering).

References

References are included in **Section 4.6**.

Chapter 5

Out of Antarctica: new insights into Antarctic Subcluster 5.2 picocyanobacteria based on high-quality genomes

Benoit Durieu^{1,2}, Valentina Savaglia¹, Mick Van Vlierberghe², Valérian Lupo^{2,3}, Denis Baurain², Annick Wilmotte^{1*+} and Luc Cornet^{1,2+}

¹ BCCM/ULC, InBioS–Molecular Diversity and Ecology of Cyanobacteria, University of Liège, 4000 Liège, Belgium

² InBioS–PhytoSYSTEMS, Eukaryotic Phylogenomics, University of Liège, 4000 Liège, Belgium

³ BCCM/MUCL and Laboratory of Mycology, Earth and Life Institute, Catholic University of Louvain, 1348 Louvain-la-Neuve, Belgium

⁺ co-last authors

Manuscript submitted for publication in **Microbial genomics** on 26 February 2026.

Author contribution of BD:

BD performed search for relevant strains in public culture collections, strains' cultivation, DNA extractions, library preparations, bioinformatic analyses, statistical analyses and wrote chapter 5.

5.1. Abstract

Synechococcus-like cyanobacteria are cosmopolitan unicellular picocyanobacteria that have colonized diverse aquatic and terrestrial habitats and represent a large diversity of lineages. The so-called ‘subcluster 5.2’ represents a particularly diversified subgroup, including marine and freshwater organisms, that is also adapted to extreme conditions, notably polar environments. We increased the genomic representation of polar taxa in this subcluster by reconstructing new high-quality genomes from five Antarctic lacustrine strains and one Arctic freshwater isolate using a combination of small Illumina and long Nanopore reads. A phylogenomic analysis of these new assemblies combined with all publicly available good-quality assemblies of subcluster 5.2 is suggestive of a dispersal process from Antarctica. Indeed, the topology of the phylogenomic tree features one basal Antarctic lineage followed by the emergence of two lineages, one Antarctic and one non-Antarctic (Spain). This finding is further supported by an expanded 16S rRNA tree and a pangenomic analysis. Although secondary colonization of Antarctica by cyanobacteria after the cooling of the continent 34 million years ago has already been reported, this study is the first to support an ‘Out-of-Antarctica’ scenario.

5.2. Impact statement

This study provides novel insights into the evolutionary history and biogeography of *Synechococcus*-like picocyanobacteria. By expanding the genomic sampling of polar subcluster 5.2 strains, we reveal that it does not include a monophyletic Antarctic lineage but that Antarctic and non-Antarctic lineages are intermingled, suggesting an unprecedented ‘Out-of-Antarctica’ dispersal scenario. Our results, supported by 16S rRNA, phylogenomic and pangenomic analyses, highlight the role of polar regions as both a refuge and a source of microbial diversity. This work advances our understanding of cyanobacterial adaptation, evolutionary radiation and genomic evolution in extreme environments.

5.3. Data summary

Raw sequencing reads and genome assemblies from this study have been deposited in the European Nucleotide Archive (ENA) under Project accession number PRJEB103944. Raw sequencing reads are available under accession numbers ERR15933635 to ERR15933638, ERR15933641 to ERR15933644, ERR15903357, ERR15903358, ERR15905256 and ERR15905257. The corresponding assemblies can be accessed under accession numbers GCA_977927515,

GCA_977927525, GCA_977927535, GCA_977927545, GCA_977927555 and GCA_977929215. The scripts and code as well as large supplementary files generated during this study are available on Figshare and can be downloaded from <https://doi.org/10.6084/m9.figshare.30772781> (Durieu, 2025). Supplementary data files are also provided and contain supplementary figures (Fig. S5.1-S5.4) and tables (Tables S5.1-S5.6).

5.4. Introduction

Synechococcus-like cyanobacteria are a group of unicellular picocyanobacteria (solitary cells smaller than 2–3 μm in diameter), including notably the *Synechococcus* and *Cyanobium* genera. Their taxonomy remains under debate because their simple and resembling morphology has led to classification into polyphyletic groups (Callieri et al., 2012; Komárek et al., 2020; Strunecký et al., 2023). They have colonized a wide range of aquatic and terrestrial habitats and exhibit a cosmopolitan distribution from equatorial to polar regions (Callieri et al., 2013; Huber et al., 2017; Scanlan et al., 2009). Among these lineages, subcluster 5.2 is a particularly diversified subgroup, with a low internal 16S rRNA identity (circa 97.3 - 98.4%) (Doré et al., 2020). Originally described as a marine group (Dufresne et al., 2008; Huang et al., 2012), subsequent studies have shown that subcluster 5.2 also includes non-marine microorganisms thriving in variable and extreme conditions, encompassing a wide range of salinity and temperature (Callieri et al., 2013; Castenholz et al., 2001; Chen et al., 2022; Doré et al., 2020; Xia et al., 2020).

Earlier phylogenetic studies based on 16S rRNA data revealed the presence of one group of three Antarctic cyanobacterial strains within subcluster 5.2 (Callieri et al., 2013; Everroad & Wood, 2012; Powell et al., 2005; Vincent et al., 2000), among which one strain, isolated from the marine-derived Ace Lake (Vestfold Hills), has been studied by genomic analyses (Panwar et al., 2022; Tang et al., 2019). This strain plays a major ecological role in the lake by forming dense blooms in a seasonal dynamic influenced by the polar light cycle (Panwar et al., 2020).

Cosmopolitan and endemic cyanobacteria are present in Antarctica (Durieu et al., 2025). Many lineages across the cyanobacterial phylogeny show phylogenetic radiations where Antarctic strains diverge from cold-tolerant (e.g., Arctic, high altitude) or more temperate ancestors. These patterns can be interpreted as adaptative radiations after colonization of the Antarctic continent from elsewhere (Christmas et al., 2015). Nevertheless, no events of dispersal from Antarctica towards another location, i.e., an 'Out-of-Antarctica' scenario, have been reported to date. In this study, we increased the genomic representation of subcluster 5.2 with polar strains by reconstructing six high-quality genome assemblies from five Antarctic and one Arctic isolates.

These new genome assemblies allow us to present a phylogenomic and pangenomic analysis demonstrating that Antarctic strains in subcluster 5.2 do not form a monophyletic group but instead are intermingled with non-Antarctic lineages. Interestingly, those secondarily non-Antarctic cyanobacteria contain a pseudogeneized version of a gene that is likely functional in their Antarctic relatives. These findings have important evolutionary history implications as, in addition to multiple documented colonization events of Antarctica, our results open the possibility of a reverse, 'Out-of-Antarctica', scenario in this case.

5.5. Methods

5.5.1. Sampling

We reconstructed the genome of four polar strains of *Cyanobium* sp. from the BCCM/ULC Cyanobacteria Collection (**Table 5.1**). Three strains were isolated by W. Vincent in Antarctic lakes from Livingston Island (South Shetland Islands, Antarctic Peninsula): ULC081 (Limnopol Lake), ULC082 (Chester Cone) and ULC084 (Laguna Chica). Strain ULC065 was also isolated by W. Vincent in Bylot Island (Arctic, Canada) and deposited on BCCM/ULC Collection under the original name "Cyanobium sp. O-154". Additionally, two Antarctic *Synechococcus* sp. strains (CS-602 and CS-603) isolated from Pendant Lake and Lake Abraxas (Vestfold Hills, Eastern Antarctica), respectively, were ordered from the Australian National Algae Culture Collection, ANACC (**Table 5.1**). Strains were grown in BG11 medium at a temperature of 12 °C.

Strain	Culture collection	Isolation source	Deposit year	ENA accession
ULC065	BCCM/ULC	Bylot Island, Canada, Arctic	2011	GCA_977927515.1
ULC081	BCCM/ULC	Limnopol Lake, Antarctica	2006	GCA_977927525.1
ULC082	BCCM/ULC	Chester Cone, Antarctica	2011	GCA_977927535.1
ULC084	BCCM/ULC	Laguna Chica, Antarctica	2011	GCA_977927545.1
CS-602	ANACC	Pendant Lake, Antarctica	1993	GCA_977927555.1
CS-603	ANACC	Lake Abraxas, Antarctica	1993	GCA_977929215.1

Table 5.1. Culture collections, isolation sources and ENA accessions of the strains sequenced in the present study.

5.5.2. DNA extraction and de novo sequencing

For each strain, 6 ml of liquid culture were centrifuged at 12,000 g to remove water and obtain biomass pellets for DNA extraction using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich, St. Louis, MO). DNA amount and quality were checked by gel electrophoresis (0.8% agarose, 80 V, 100 mins) and with the Quant-iT™ PicoGreen™ dsDNA Assay Kit (Thermo Fisher Scientific, Waltham, MA).

Short paired-end read sequencing was performed for the six strains on an Illumina MiSeq 600 cy V3 sequencer (2 x 300 bp) and the Nextera XT DNA Library preparation Kit (Illumina, San Diego, CA, USA). Long read sequencing was performed on a MinION Flow Cell R9.4.1 (Oxford Nanopore Technologies, UK) after barcoding and ligation sequencing preparation using Native Barcoding Kit 24 (Oxford Nanopore Technologies, UK). A sequencing depth of 100 x was expected for the two sequencing technologies, based on a hypothesized average genome size of 3 Mbp.

5.5.3. Genome assembly and quality assessment

Raw reads from Illumina sequencing were first filtered with PRINSEQ (Schmieder & Edwards, 2011) with custom parameters (-min_len: 80, -min_qual_mean: 15, -trim_qual_right: 25, -qual_noscale) and then with Trimmomatic (Bolger et al., 2014) to remove Illumina adaptors (leading: 20, trailing: 20, slidingwindow: 10:20, crop: 270, minlen: 80). Base-calling of the MinION raw reads was performed using the Guppy basecaller (Oxford Nanopore Technologies, UK) and a filter step was applied with Nanofilt (De Coster et al., 2018) to retain only sequences with Q-score ≥ 10 . Hybrid de novo assembly of filtered reads was then performed with Unicycler (--no_correct, --mode normal; Wick et al., 2017). We identified cyanobacterial contigs in the hybrid assemblies using Kraken2 (Wood et al., 2019) default parameters and databases, and those cyanobacterial contigs were merged after checking for redundancies using mummer and mummerplot (Marçais et al., 2018). CheckM (Parks et al., 2015) was then used with the "Taxonomic and Lineage-specific" workflow to determine completeness and contamination. Information about length, taxonomy and mummer redundancies of the contigs is provided in **Table S5.1**.

5.5.4. Multilocus maximum likelihood phylogeny (phylogenomic analysis)

The 2336 genomes assemblies of the order Synechococcales available as of 7 June 2024 in the NCBI Assembly database were downloaded using the Genome-downloader.nf workflow

implemented in the GEN-ERA toolkit (Cornet et al., 2023). These assemblies were compared to our six polar assemblies by Average Nucleotide Identity (ANI) using the ANI.nf workflow (also from GEN-ERA), with default parameters (Cornet et al., 2023). We conserved 291 assemblies sharing at least 77% of ANI with one or more of our polar assemblies, slightly below the value of 80-84% used by Dufresne *et al.* (Dufresne et al., 2008) to distinguish subclusters. Quality criteria (contamination $\leq$ 5%, completeness $\geq$ 95%) were computed using the GENcontams.nf workflow (GEN-ERA), with default parameters (Cornet et al., 2023). When multiple assemblies were available for a single cyanobacterial strain, only the highest-quality assembly was retained for further analyses based on completeness and contamination values, resulting in the selection of 48 of the 291 public assemblies for phylogenomic analysis. We then added *Synechocystis* PCC 7509 (GCF_000332075.2) as an outgroup (**Table S5.2**). A phylogenetic inference was performed on these 55 assemblies using the Phylogeny.nf workflow from the GEN-ERA toolbox, with default parameters [26]: 476 Orthologous groups (OGs), corresponding to unicopy genes shared by all genomes, were retrieved (Cornet et al., 2023). A multi locus maximum likelihood phylogeny using RAxML (Stamatakis, 2014) with 1000 rapid bootstraps under the PROTGAMMALGF model implemented in the Phylogeny.nf workflow (GEN-ERA) (Cornet et al., 2023) was performed on these 476 OGs. The phylogenomic tree was further arranged with iTOL v7 (Letunic & Bork, 2024) to delineate the different polar groups.

5.5.5. 16S rRNA phylogeny

16S rRNA genes were extracted from assemblies with RNAmmer v1.2 (Lagesen et al., 2007). We then retrieved all available sequences from the NCBI “core_nt” database (carried out on 10 March 2025) at a 97% similarity threshold, to encompass the diversity of the phylogenomic “group D”. The retained 16S rRNA gene sequences were aligned using MAFFT v8.2.12 (Katoh & Standley, 2013) using default parameters and a phylogenetic tree was inferred using RAxML (Stamatakis, 2014) with 100 rapid bootstraps under the GTRGAMMA model. Geographic locations for the whole dataset were retrieved in batch from GenBank using EFetch (Entrez Fetch) utility (Schuler et al., 1996). The final tree was visualized with iTOL v7 (Letunic & Bork, 2024).

5.5.6. Genome annotations and pangenomic analysis

Among the 55 assemblies included in the phylogenomic analysis, 44 genomes were retained for the pangenomic analysis. Selection was based on assemblies containing no more than 100 contigs and showing less than 50% strain heterogeneity, according to CheckM results.

Pangenomic analysis was performed using the 'anvi-pan-genome' program in Anvi'o, after genome formatting with 'anvi-gen-contigs-database' and annotation using 'anvi-run-ncbi-cogs', 'anvi-run-kegg-kofams', and 'anvi-run-pfams' [34–40] (Figshare identifier: 10.6084/m9.figshare.30772781). The workflow used BLASTP to compute the amino acid identity level between all predicted proteins and remove matches below a bitscore threshold (minbit heuristic = 0.5; Altschul et al., 1990). Homologous gene clusters (GCs) were then determined using the Markov Clustering Algorithm (MCL inflation = 8; (Eren et al., 2015; Utter et al., 2020)) (Figshare identifier: 10.6084/m9.figshare.30772781).

Functional annotations significantly enriched (q-value <0.05) in Antarctic genomes compared to non-Antarctic genomes were identified with the program 'anvi-compute-functional-enrichment-in-pan'. A false discovery rate correction was applied to p-values generated using the Rao score test for equality of proportions to account for multiple testing. To strengthen the robustness of our functional annotation analysis, we used the COG and KOfam/Pfam databases and retained only enriched functions for which the GC annotations corresponded across both databases.

5.5.7. In-depth analysis of specific Antarctic gene clusters and functional enrichment analysis

The homologous gene clusters (GCs) inferred from the pangenomic analysis were examined with respect to their Antarctic or non-Antarctic origin. GCs specific to Antarctic (i.e., present only in Antarctic assemblies or in all Antarctic assemblies and not in their closest relatives) were counted using a custom Perl script (Figshare identifier: 10.6084/m9.figshare.30772781). From these GCs, those with functional annotations significantly overrepresented in Antarctic assemblies (i.e., genes *insA* and *higB*) were selected for further analysis (Figshare: 10.6084/m9.figshare.30772781). As *higB* was sometimes unexpectedly absent from assemblies despite its supposedly tight functional association with *higA*, being in the same operon, we applied a more sensitive detection method using Forty-Two (Simion et al., 2017), as in Lupo et al. (Lupo et al., 2025). To this end, HigB protein sequences were aligned with Clustal Omega (Sievers et al., 2011), manually curated, and a PhyML tree (Guindon et al., 2009) was inferred under the

GTRGAMMA model to separate the protein sequences annotated as HigB in homologous groups (Guindon et al., 2010). Each homologous group of HigB protein sequences was then aligned with MAFFT (Kato & Standley, 2013) and used as TBLASTN queries to mine the 44 assemblies with Forty-Two. The protein sequences found by Forty-Two were then classified as (a) HigB (97% similarity, green), (b) potential HigB (85% similarity, grey), (c) potentially not HigB (<74% similarity, black), (d) partial sequences (turquoise), (e) truncated sequences (brown) and (f) undetermined HigB (presence of “X”, red). The exact genomic positions of the HigB protein sequences detected with Forty-Two were then added to the GFF file generated by Anvi’o for each assembly using custom Python scripts (Figshare: 10.6084/m9.figshare.30772781). Finally, the resulting GFF files were used in GeneSpy (Garcia et al., 2019) to visualize the genomic context of *higB* gene sequences across all assemblies.

5.6. Results

5.6.1. General characteristics of the assemblies

For our six new polar assemblies, the number of cyanobacterial contigs identified by Kraken2 ranges from 1 (ULC065) to 13 (CS-602) for a minimum completeness of 99.46% (UC081 and CS-602) and a maximal contamination level of 0.54% (ULC081 and ULC065). Statistics for these new assemblies and the 49 others retrieved from the NCBI Assembly database are summarized in **Table S5.2**. Briefly, the maximum number of contigs is 224 (*Synechococcus lacustris* GCF_003011125.1), while the minimum completeness is 97.69% (*Cyanobium* sp. A1C AMD GCF_024346495.1) and the maximum contamination level is 1.72% (*Synechococcus lacustris* GCF_003011125.1). Genome size ranges from 2,164,299 bp (*Ircinia ramosa* symbiont *Synechococcus* sp. SB0662 GCA_009840445.1) to 4,074,513 bp (freshwater *Cyanobium* sp. FGCU6 GCA_024958655.1), GC content from 40.62% (*Synechococcus* sp. PCC 7502 GCF_000317085.1) to 69.59% (*Cyanobium* sp. LC18 GCF_009834675.1) and the number of genes from 2140 (*Synechococcus* sp. SB0662 GCA_009840445.1) to 4280 (*Cyanobium* sp. FGCU6 GCA_024958655.1).

5.6.2. Antarctic genomes are paraphyletic in the phylogenomic analysis of subcluster 5.2

The phylogenomic analysis reveals that Antarctic and Arctic genomes are not monophyletic and are found in two and three different groups, respectively, within subcluster 5.2 (**Fig. 5.1**). These

groups contain genomes from microorganisms inhabiting glacial or temperate lakes, as well as marine environments. **Group A** encompasses freshwater lacustrine *Cyanobium/Synechococcus* strains sampled from three temperate locations (Mexico, Spain, USA), one strain from the post-glacial freshwater lake Constance (Switzerland), and strain CS-1332, which is from the Arctic (Lake C3, Nunavut, Ellesmere Island, Canada). **Group B** is composed of four freshwater lacustrine strains from temperate locations (Mexico, Spain) and two from the Arctic: one, marine, from Northwest Territories (Eastern Beaufort Sea) and the other from the freshwater Lake A (Ellesmere Island, Nunavut, Canada). **Group C** is a mixture of lacustrine (New Zealand, Spain, Argentina) and marine (Black Sea, Bulgaria, and Menai Strait, UK) strains, while 11F2 is reported as an Antarctic strain from Edmonson Point (Victoria Land, Antarctica). We ordered and resequenced this strain because it was nearly identical to other strains of the Group C, both in terms of ANI (100%; **Table S5.3**), number of genes and genome size (**Table S5.2**). However, we found 99.99% ANI between the assembly used in this study and the re-sequenced strain. This highlights a potential traceability issue that might only be resolved through re-analysis of the original environmental sample. **Group E**, is characterized by taxa from very diverse lacustrine sources, such as the Arctic (Lake A, Nunavut, Bylot Island, Canada), freshwater post-glacial high-altitude lakes (Lake Lugano, Switzerland, and Lake Mondsee, Austria), the alkaline lake Atexcac (Mexico), an artificial freshwater lake in Czech Republic (Mostecké Jezero), a freshwater pond in Montréal (Canada), the freshwater Loriguilla Reservoir (Chulilla, Spain) and an undetermined source of water found in the digestive system of a honeybee from Beijing (China).

Regarding **group D** which includes our new polar genomes, it encompasses freshwater and saline Antarctic strains, respectively from Livingston Island (near the Antarctic Peninsula) and from the Vestfold Hills (Eastern Antarctica) and two freshwater strains originating from lakes situated in the mountains of the Serrania de Cuenca (Central-Eastern Spain). Interestingly, the Antarctic strains from group D are divided in two monophyletic lineages separated by the Spanish assemblies (**Fig. 5.1**). The first lineage contains the brackish Antarctic genomes CS-601, CS-602 and CS-603 from different lakes of the Vestfold Hills, which share 99.2 to 99.6% ANI (subgroup D1). The second one includes the four freshwater Antarctic genomes ULC081, ULC082, ULC084 and CS-1327, sampled from different lakes in Byers Peninsula in Livingston Island and sharing 98.1 to 99.9% ANI (subgroup D2) (**Table S5.3**). The subgroup D2 is more genetically distant from the group D1 (81.4 to 81.6% ANI) than from the Spanish group (subgroup D3), comprising the nearly identical strains Cruz_CV_v_12 and Tobar12_5_mg (84.6 to 84.7% ANI).

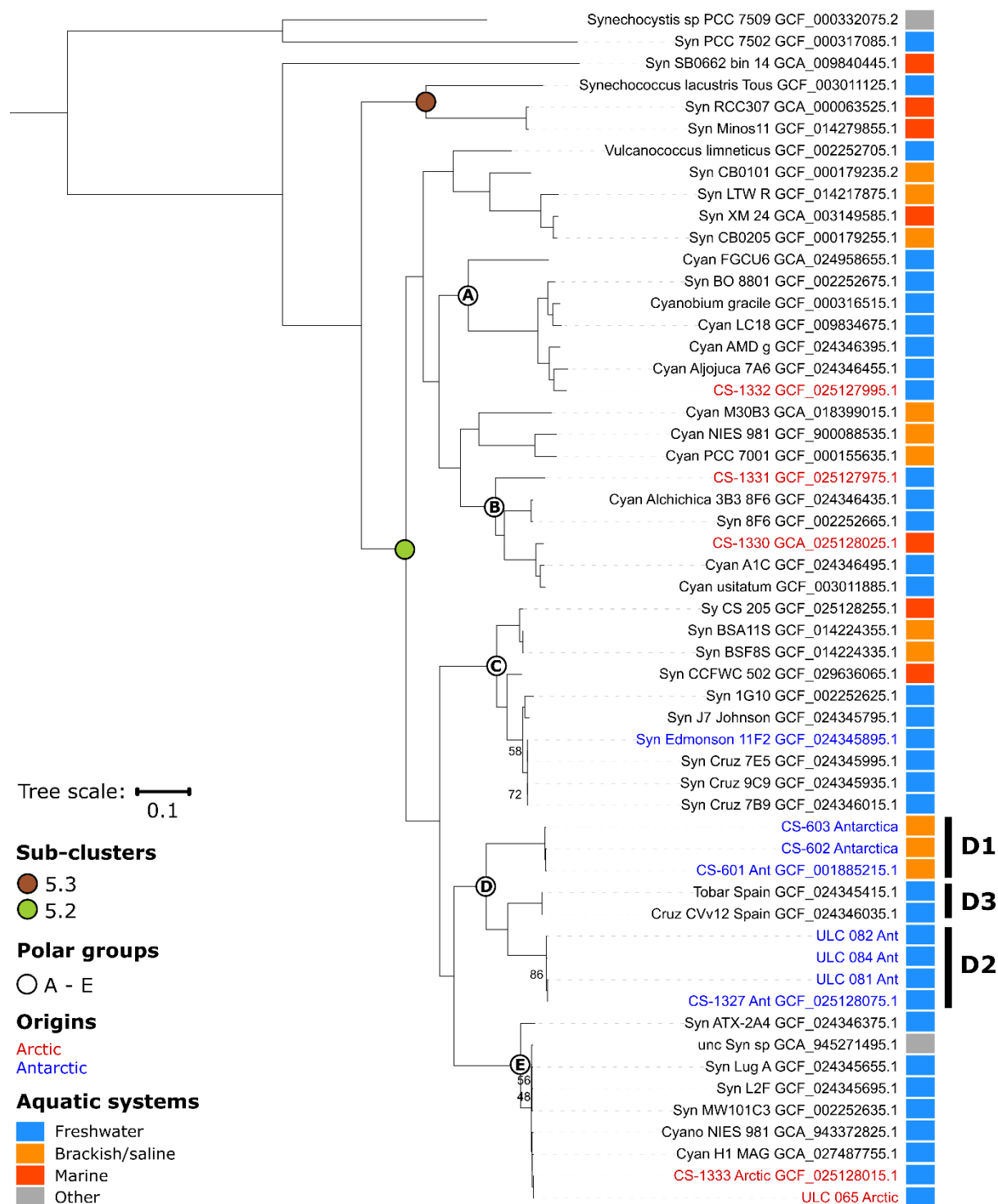


Fig. 5.1. Multilocus maximum likelihood phylogeny with 1000 rapid bootstraps under the PROTGAMMALGF model of 55 cyanobacterial assemblies inferred from a supermatrix of 476 genes totalizing 69,007 unambiguously aligned amino acid positions and 0.40% of missing character states. Bootstrap values <100 are indicated in front of the corresponding nodes.

5.6.3. Densification of group D with environmental 16S rRNA sequences confirms the presence of a non-Antarctic lineage between the two Antarctic ones

The paraphyletic nature of Antarctic strains in group D was confirmed by the phylogenetic analysis of their 16S rRNA sequences. With a bootstrap above 90%, this analysis confirms that group D contains non-Antarctic taxa (subgroup D3) between the two Antarctic subgroups D1 and D2 (**Fig. 5.2**). We could not recover any 16S rRNA sequence from GCF_024345415.1 (*Synechococcus* sp. Tobar12_5mg) and GCF_024346035.1 (*Synechococcus* sp. Cruz_CVv12) assemblies, but other 16S rRNA sequences retrieved from the NCBI “core_nt” database did fall in subgroup D3. These non-Antarctic sequences were from polar fjord (Greenland), high-altitude lake (Pyrenees), temperate terrestrial (Portugal, Acores) and coastal marine (Russia) locations.

We also note that subcluster 5.3 is split into two groups in the 16S rRNA phylogeny (**Fig. 5.2**), while it is monophyletic and well supported in the phylogenomic tree (**Fig. 5.1**). This kind of incongruence has already been reported in a previous comparison between 16S rRNA-based phylogeny and phylogenomic analysis of picocyanobacteria (Mareš, 2018) and is due to the lower phylogenetic resolution of the 16S rRNA (Dessimoz & Gil, 2010; Gontcharov et al., 2004; Lunter et al., 2008; Mareš, 2018; Wong et al., 2008). Finally, as in the phylogenomic analysis, group E is the closest relative to group D.

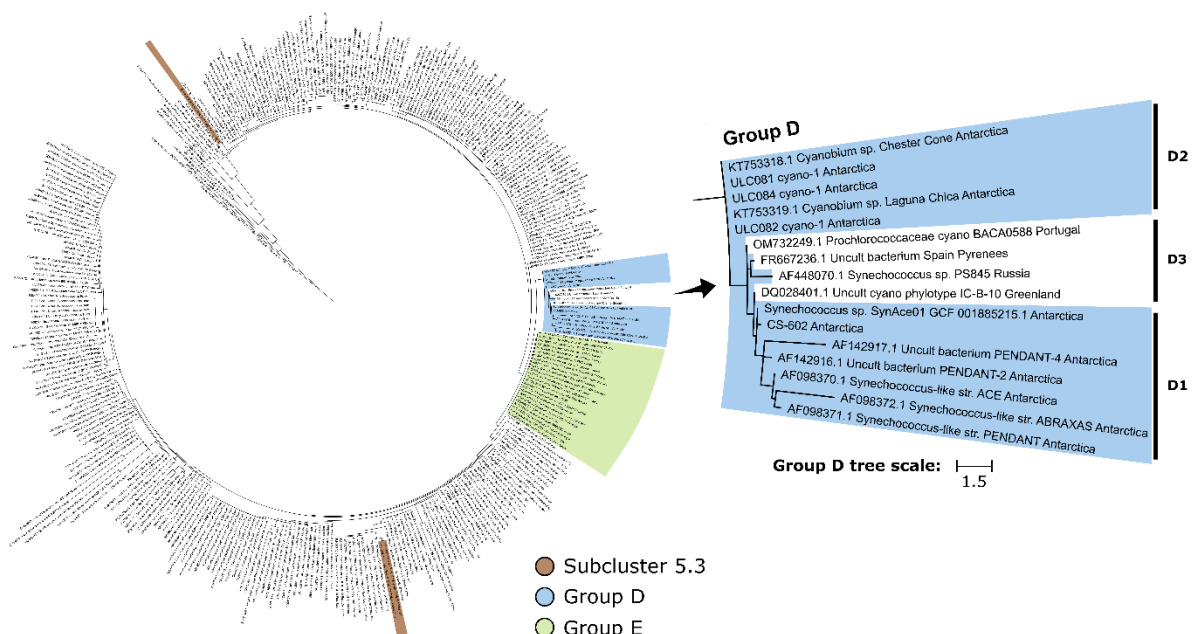


Fig. 5.2. Maximum likelihood phylogeny of the 405 16S rRNA sequences contextualizing the Antarctic Group D, with 100 rapid bootstraps under the GTRGAMMA model for an alignment of 746 unambiguously aligned nucleotide positions and 6.63% missing character states. Subgroup D3 corresponds to all non-Antarctic 16S rRNA sequences. D1 and D2 positions are inverted compared to the phylogenomic tree (**Fig. 5.1**). The scalebar (in substitutions per site) only refers to the enlarged subtree. The figure in better

resolution is available on Figshare and can be downloaded from: <https://doi.org/10.6084/m9.figshare.33200250>.

5.6.4. Group D Antarctic genomes are also grouped together in the pangenomic analysis and contain unique GCs

A total of 137,390 proteins were predicted across the 44 assemblies selected for the pangenomic analysis, which clusters these sequences into 22,886 gene clusters (GCs) (**Fig. S5.1**). Group D has 33 unique GCs, while subgroup D2/D3 assemblies share 72 unique GCs (**Fig. S5.1**). At the group level, D1, D2 and D3 assemblies have 302, 141 and 417 unique GCs, respectively, corresponding to their accessory genomes (**Fig. S5.1**).

To better characterize group D and highlight the potential specificities of Antarctic taxa, we determined the GCs that are unique to the subgroups D1 and D2 and absent from the subgroup D3 (**Fig. 5.1**, **Fig. 5.2** and **Fig. S5.1**). Hence, the genome assemblies of Antarctic groups D1 and D2 specifically share 16 GCs not found in the non-Antarctic assemblies of subgroup D3, thereby defining strictly Antarctic GCs in group D. Moreover, 84 GCs are present in all D1/D2 assemblies and in some assemblies outside of group D but are totally absent from subgroup D3. The descriptions of these 16 + 84 GCs found in all D1/D2 assemblies and missing from D3 assemblies are available in **Table S5.4**.

5.6.5. Only two Antarctic gene clusters (GCs) display functional enrichment in the Antarctic genomes

Focusing on the 16 + 84 GCs present in subgroup D1 and D2 Antarctic assemblies but not in the two Spanish assemblies (subgroup D3), only two of them are enriched in the Antarctic genomes compared to the non-Antarctic ones, in terms of functional annotation (COG and KOfam/Pfam) with an adjusted q-value of 0.069, close to the significance threshold (GC_00000016 and GC_00002029; **Table S5.4**). Here, “functional enrichment” refers to functions that are overrepresented in a specific group of genomes. For the other 98 GCs, annotations are identical to those of GCs from assemblies outside group D; their functional annotations are therefore no longer specifically enriched in Antarctic subgroups D1/D2.

Hence, GC_00000016 are genes that code for a transposase (*insA*) and GC_00002029 are genes that code for a toxin-antitoxin (TA) system killer protein (*higB*). *InsA* is present with a higher number of copies in the D1/D2 Antarctic assemblies than in other assemblies (**Fig. S5.2**). The diversity of

the *InsA* protein sequences is important (19 different GCs). *higB* has less copies per genome than *insA* but still shows significant enrichment. It is also less diversified (2 different GCs). Copy numbers of *insA* and *higBA* per assembly are detailed in **Fig. S5.2**. Considering that HigB is a toxin protein that is part of a toxin-antitoxin (TA) system (Fraikin et al., 2020), we checked the presence of the companion antitoxin HigA and its position in the assembly. Yet, in contrast to HigB, HigA is not enriched in Antarctic genomes compared to non-Antarctic ones (**Fig. S5.2**).

5.6.6. In-depth analysis of *higBA* gene clusters reveals potential pseudogenes in the most-closely related non-Antarctic genomes within group D

To better characterize the gene integrity of the GCs of interest, an in-depth search for the genes encoding HigB and HigA was carried out with Forty-Two (**Fig. 5.3**) and their genomic contexts were visualized with GeneSpy (**Fig. S5.3**). This analysis showed that HigB, when present, is almost always (37 of 38 occurrences) concomitant with HigA (**Table S5.5**). The alignment of the new sequences of HigB predicted by Forty-Two with those of Prodigal shows that the HigB gene sequences found in *Synechococcus* sp. Cruz CV_v_12 and Tobar_12_5_mg were truncated when compared to their close relatives from group D (**Fig. S5.4**). The genomic context of *higA* and *higB* (upstream *higB* ORFs and downstream *higA* ORFs) is conserved between Cruz CV_v_12 and Tobar12mg, but not between those and the genomes of group D2 or group D1. Instead, GeneSpy analysis revealed that *higB* and *higA* genes in group D1 and D2 are almost always bordered by, e.g., genes *hicB* (in 5 of the 7 assemblies), *hinT* (in 5 of the 7 assemblies), *Mj0435* (in 4 of the 7 assemblies) and *HEPN* (in 4 of the 7 assemblies), while *higB* and *higA* in “Cruz” and “Tobar” genomes are bordered by *wcaA* and *bcasA* genes (**Fig. S5.3**). Looking at the assemblies outside of group D, Forty-Two also predicted an additional *higB* sequence for *Synechococcus* sp. L2F and *Cyanobium* sp. FGCU (**Fig. 5.3**).

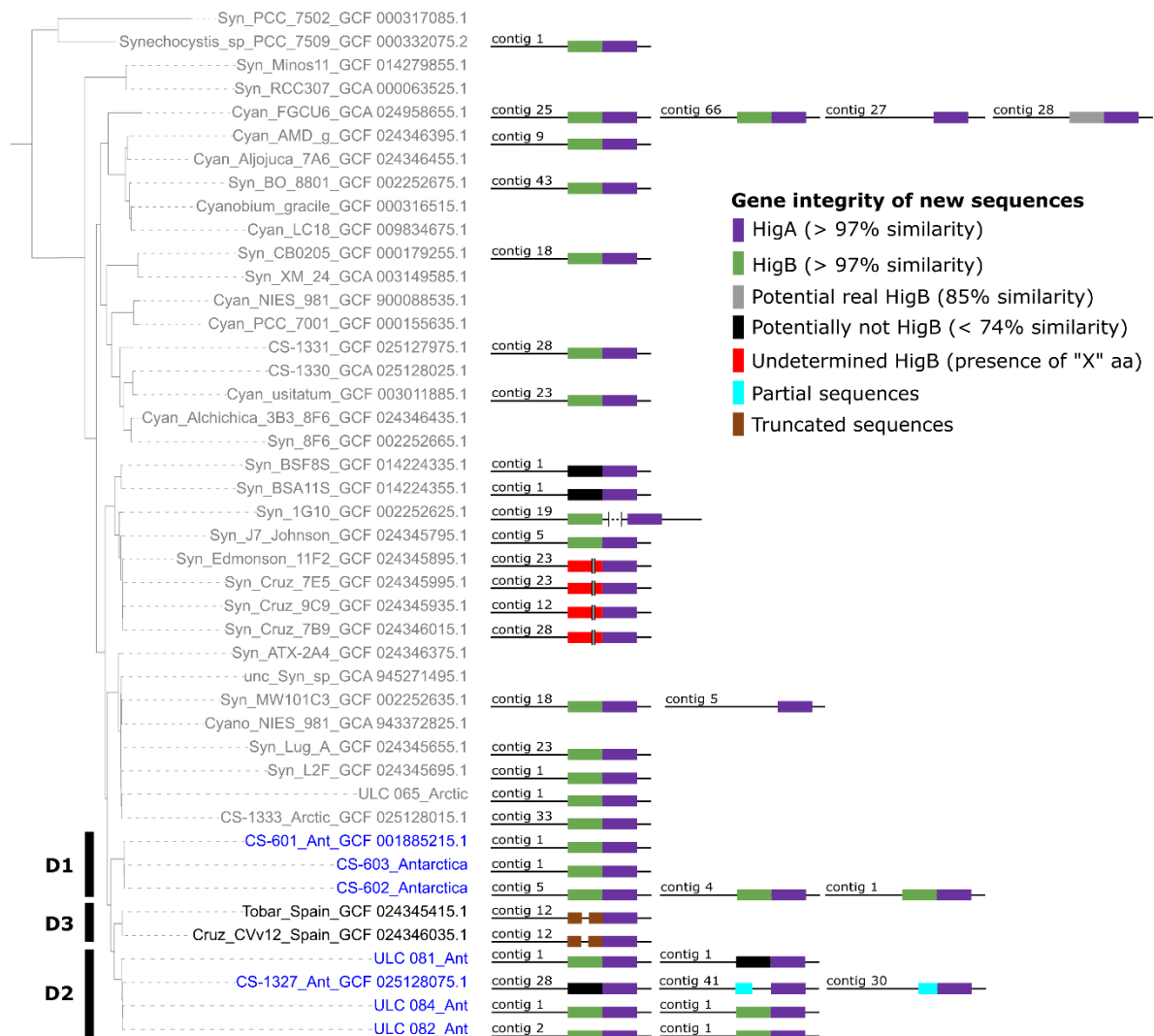


Fig. 5.3. Details of the number and gene integrity of higBA clusters in the 44 assemblies of the pangenomic analysis, based on Forty-Two analysis.

5.6.7. Functional annotations enriched in Antarctic genomes of the group D

Beyond the GC level, enrichment was also examined at the functional annotation level, across multiple GCs. In total, 34 and 37 functional annotations, respectively from COG and KOfam/Pfam databases, were significantly enriched (q -value < 0.05) in the D1 and D2 Antarctic genomes. The intersection of these two annotation sources reveals that 19 equivalent functions (where GC annotations match in both COG and KOfam/Pfam databases) are significantly enriched (**Fig. 5.4** and **Table S5.6**). Out of these, 21% are not predicted in detail (COG category R “General function prediction only”) and 21% are unknown (S “Function unknown”); 32% of them are indirectly linked to stress or cold adaptation (**Fig. 5.4**).

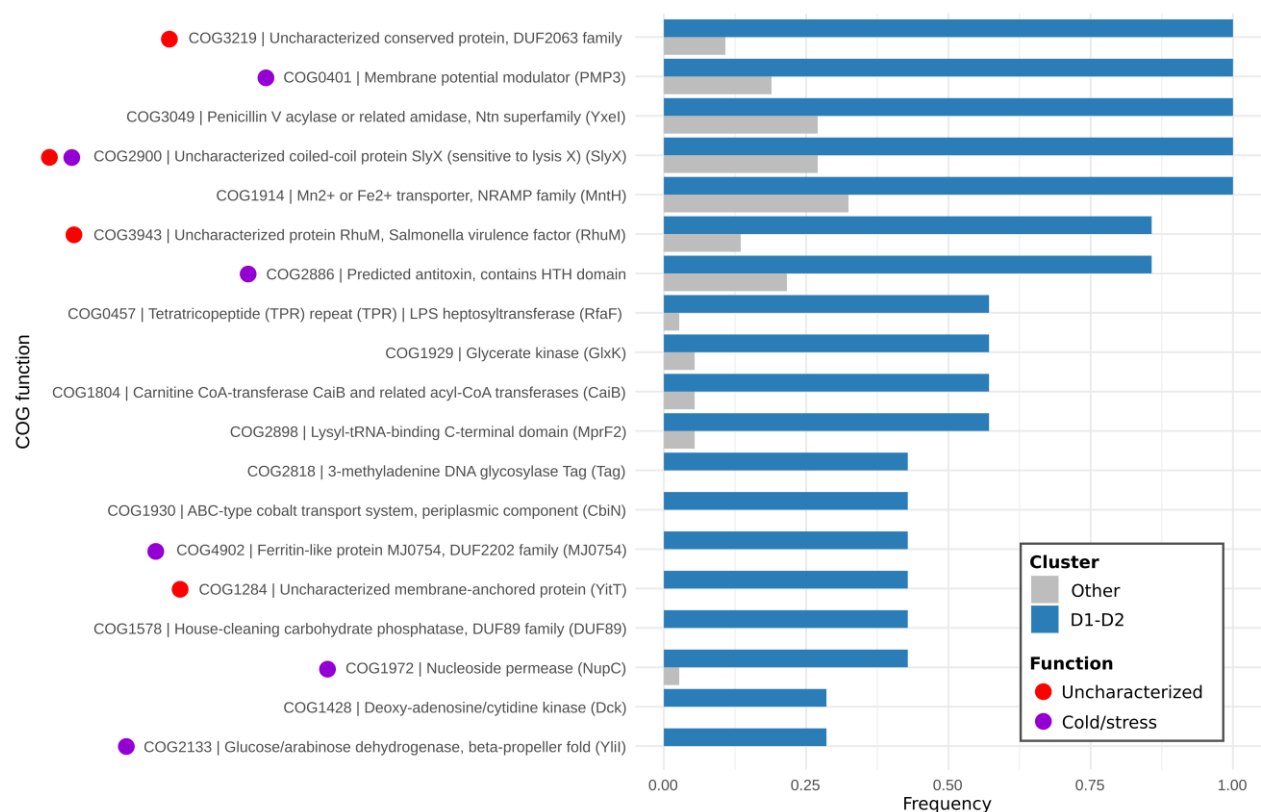


Fig. 5.4. Frequency of the 19 significantly enriched COG functional annotations for the Antarctic D1-D2 groups compared with the other studied genomes. The corresponding KOfam/Pfam functional annotations are described in **Table S5.6**.

5.7. Discussion

5.7.1. Group D: an Antarctic lineage as an origin for dispersal in non-Antarctic ecological niches?

The subcluster 5.2 represents a particularly diversified subgroup, including marine and freshwater organisms. Among this subcluster, Antarctic genomes (e.g., CS-601, ULC084) were previously described as belonging to a monophyletic lineage in phylogenomic studies (Panwar et al., 2022; Tang et al., 2019) and 16S rRNA phylogenies (Callieri et al., 2013; Everroad et al., 2013; Tang et al., 2019), with as closest relatives two strains from the Spanish lakes La Cruz and Tobar (Cabello-Yeves et al., 2022). Building on newly generated and publicly available polar genomes, our phylogenomic analysis shows that those Antarctic genomes form a group with other Antarctic taxa (CS-602, CS-603, CS-1327, ULC081, ULC082) and, interestingly, non-Antarctic ones (Spanish strains Tobar and Cruz_CV12) (**Fig. 5.1**). We designated this group with a majority of Antarctic strains as group D. Our 16S rRNA phylogenetic analysis, which specifically enriched the closest relatives of group D, confirms the presence of non-Antarctic strains between the two

Antarctic subgroups of group D, even broadening their geographic distribution with the Azores, French Pyrenees, Russia and Greenland (Dessimoz & Gil, 2010; Dufresne et al., 2008; Gontcharov et al., 2004; Lunter et al., 2008; Mareš, 2018; Wong et al., 2008). Overall, our trees suggest that the ancestors of group D within subcluster 5.2 likely lived in Antarctic. This inference leads to the hypothesis that, in addition to being colonized several times by cyanobacterial taxa (Christmas et al., 2015), Antarctica has also been the origin of at least one lineage that subsequently dispersed outside the continent in other ecological niches. To the best of our knowledge, apart from the perspective proposed for *Nodularia* in Christmas et al. (2015), no study on cyanobacteria has yet reported such an "Out-of-Antarctica" scenario (Christmas et al., 2015). However, we cannot determine the precise timing of these events because the sampled lakes are only 8000 years old (Fulford-Smith & Sikes, 1996; Toro et al., 2013), whereas the Antarctic lineages studied here diverged dozens million years ago, based on the classical 16S rRNA substitution rate (Ochman & Wilson, 1987).

5.7.2. The specificities of genomes of the group D: no clear "Antarctic signal" and a pseudogene that could be a signature of ancestry

Our pangenomic analysis of the 44 genomes determined 16 Anvi'o gene clusters (GCs) unique to the Antarctic genomes of group D (subgroups D1 and D2), and 84 GCs not unique to those Antarctic genomes but strictly absent from the non-Antarctic genomes of group D (subgroup D3). Due to our larger genome dataset, this number is inevitably smaller than the 901 reported in Tang et al. (2019), who compared only one Antarctic genome with three non-Antarctic ones (Tang et al., 2019). Moreover, Tang et al. (2019) determined the unique genes of one specific genome, whereas we determined the genes shared across a collection of genomes from the same environment.

Among these 16 + 84 GCs, we did not detect any genetic signature related to stress or cold adaptation, which could have been expected in Antarctic environments, with two exceptions concerning the two GCs enriched in Antarctic genomes. It is worth noting that these GCs, a transposase (InsA) and a toxin-antitoxin (TA) system killer protein (HigB), can be related to stress adaptation. Indeed, TA systems have been described as important for adaptation to stress and metabolism regulation (e.g., biofilm formation, carbon metabolism, ion uptake; (Song et al., 2021)), with an important variation from one species to another (Loris & Garcia-Pino, 2014; Pavelich et al., 2025). Because HigB is part of a TA system, we searched for the presence of the *higA* gene. The HigB and HigA protein pair were not encoded in all genomes, but *higA* was present alone, or along with an altered version of *higB*, in eight genomes (**Fig. 5.3**). This is in line with other

studies, in which the *higA* gene was also found alone, a fact never reported for *higB* (Budde et al., 2007). When *higB* and *higA* genes are both present, we observed that they are encoded consecutively on the chromosome as *higBA*. This was also reported in other studies (Deter et al., 2017; Fraikin et al., 2020). The deeper analysis of *higB* genes reveals the presence of truncated sequences in the two non-Antarctic (i.e., Spanish) genomes of subgroup D3, which could be the consequence of pseudogenization. Pseudogenization in bacteria is regarded as a process in which certain genes become disposable to the fitness of the organism due to a shift of ecological niche, and hence loss of selective pressure (Goodhead & Darby, 2015; Kuo & Ochman, 2010). Consequently, the truncated *higB* sequences found in the non-Antarctic subgroup D3 genomes provide an additional clue for supporting an ‘Out-of-Antarctica’ scenario.

Aside from *insA* and *higB*, we identified 19 functions, beyond mere sequences homology (Delmont & Eren, 2018), that are significantly enriched in D1 and D2 Antarctic genomes compared to non-Antarctic genomes (**Fig. 5.4**). Among these enriched functional annotations, 32% are indirectly linked to stress or cold adaptation, which may provide an ecological advantage in Antarctic conditions. COG0401 (membrane potential modulator PMP3/YqaE) is part of the Esi3/RCI2/PMP3 gene family that is induced in response to biotic and abiotic stresses in plants, including cold (Brunetti et al., 2018). COG1972 (nucleoside permease NupC) is part of nucleoside transporters that are involved in the cAMP/CRP-regulated incorporation of extracellular ADP-glucose enhanced under cell envelope stress conditions in *E. coli* (Almagro et al., 2018). Nucleosides are essential to repair DNA damage, which is particularly important in Antarctica due to high UV radiation and oxidative stress. COG4902 (Ferritin-like protein MJ0754) belongs to proteins that are involved in cold stress response, as reported in *Listeria monocytogenes* (Miladi et al., 2012). COG2931 (CA²⁺ binding protein, RTX toxin-related) belongs to RTX toxins that use CA²⁺ binding to modulate immune system and promote virulence (Coote, 1992). CA²⁺ binding to an antifreeze protein in the Antarctic bacterium *Marinomonas primoryensis* may serve as an adhesin to transiently bind the bacteria to ice (Guo et al., 2012). Interestingly, another component related to TA systems, different from the couple HigBA described above, was found to be enriched: COG2886 (predicted antitoxin, contains HTH domain). This result reinforces the hypothesis of a potential role for TA systems in the Antarctic environment. It is possible that the absence of the antitoxin gene might lead to, e.g., the modification of the transcription of genes (Fraikin et al., 2020; Liu et al., 2020; Qin et al., 2021) or the death of the bacterium (Hurley & Woychik, 2009; Singh et al., 2021; Song et al., 2021). Moreover, another type of transposase gene than *insA* appears to be enriched: COG2900 (Transposase IS66 family). An enrichment of transposases corroborates the observation of Tang et al. (2019), who showed a high content in genes encoding

diverse transposases in the Antarctic SynAce01 genome (corresponding to strain CS-601 studied here) (Tang et al., 2019). Transposition plays an important role in genomic rearrangements, and it influences gene regulation and adaptation processes that can underline microevolution in cyanobacteria (Mikheeva et al., 2013), and thus possibly in Antarctica too.

The stress and cold adaptation signals in the Antarctic genomes of group D in subcluster 5.2 might be minimized by the great diversity of locations and environmental parameters experienced by the members of the subcluster, which may dilute or mask lineage-specific genomic adaptations. Indeed, many strains of the subcluster 5.2 have been reported in stressful or variable environments, like estuaries (Fucich et al., 2019; Xia et al., 2020; Zhu et al., 2015), oligotrophic lakes (Cabello-Yeves et al., 2018), Black Sea (Di Cesare et al., 2020), sub-alpine lakes (Tang et al., 2019), polar lakes (*Synechococcus* sp. CS-1330, CS-1331, CS-1332 and CS-1333), glacial lakes (*Synechococcus* sp. BO 8801, J7 Johnson, 1G10, and Lug A), and volcanic lakes (*Cyanobium* sp. Aljojuca 7A6, *Synechococcus* sp. 8F6, ATX 2A4 and *Vulcanococcus limneticus* LL). Yet, previous studies have shown that cold tolerance genes are present in most cyanobacteria, not specifically those thriving in cold environments (Christmas et al., 2016; Lumian et al., 2024). This observation is consistent with the widespread observation that cyanobacteria are psychrotolerant rather than psychrophile (Tang et al., 1997; Taton et al., 2006; Velázquez et al., 2011; Vincent, 2007).

5.8. Conclusion

Taken together, our phylogenomic, phylogenetic and pangenomic analyses converge towards a scenario in which the Antarctic group D within subcluster 5.2 represents not only a diversification into polar environments, but also a potential source of dispersal back toward non-polar ecological niches. Truncated *higB* gene sequences in non-Antarctic lineage members, combined with the topology of the phylogenies, provide evidence for an ancestral Antarctic origin followed by subsequent expansion outside the continent. The patchy distribution of stress-related functions suggests that cold adaptation in these lineages may reflect pre-existing psychrotolerant capacities rather than specialized Antarctic innovations, even though transposases and TA systems might ultimately be shown to be important in Antarctica.

5.9. Author statements

5.9.1. Author contributions

BD performed all analyses and prepared the Figures. VS aided with pangenomic analysis and preparation of figures, MVV aided with the gene cluster analysis, and VL contributed to the Forty-Two analysis. LC, DB and AW supervised the study. BD, VS, LC, DB and AW wrote the initial manuscript. All authors read and approved the final manuscript.

5.9.2. Funding information

This work was supported by research grants (PDR T.0018.24 OR-OX-PHOT-IN-CYN and CDR J.0195.23) from the Belgian National Fund for Scientific Research (F.R.S.-FNRS) to DB and AW respectively. BD was supported by a F.R.I.A fellowship from the Belgian National Fund for Scientific Research (F.R.S.-FNRS) and by BelSPO (BCCM/ULC). LC is supported by a mandate of “Collaborateur scientifique” from the Belgian National Fund for Scientific Research (F.R.S.-FNRS). AW is Senior Research Associate of the FRS-FNRS . Computational resources were provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the F.R.S.-Le Fonds National de la Recherche Scientifique (FNRS) (2.5020.11).

5.9.3. Acknowledgements

This paper is a contribution to the Scientific Committee on Antarctic Research (SCAR) Scientific Research Program 'Integrated science to inform Antarctic and Southern Ocean Conservation' (Ant-ICON). We thank Prof. Warwick Vincent (Laval University, Quebec, Canada) for depositing the studied strains in the BCCM/ULC collection of cyanobacteria, as the technical staff of the BCCM/ULC collection, especially Kim Beets, for their help with the cultivation of the strains used in this study.

OpenAI (ChatGPT-5) was used as a support in the design of script 1, script 2 and script 3 (Figshare identifier: 10.6084/m9.figshare.30772781).

5.10. References

- Almagro, G., Viale, A. M., Montero, M., Muñoz, F. J., Baroja-Fernández, E., Mori, H., & Pozueta-Romero, J. (2018). A cAMP/CRP-controlled mechanism for the incorporation of extracellular ADP-glucose in *Escherichia coli* involving NupC and NupG nucleoside transporters. *Scientific Reports*, 8(1), 15509. <https://doi.org/10.1038/s41598-018-33647-w>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/https://doi.org/10.1016/S0022-2836(05)80360-2)
- 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>
- Brunetti, S. C., Arseneault, M. K. M., & Gulick, P. J. (2018). Characterization of the Esi3/RCI2/PMP3 gene family in the Triticeae. *BMC Genomics*, 19(1), 898. <https://doi.org/10.1186/s12864-018-5311-8>
- Budde, P. P., Davis, B. M., Yuan, J., & Waldor, M. K. (2007). Characterization of a higBA toxin-antitoxin locus in *Vibrio cholerae*. *Journal of Bacteriology*, 189(2), 491–500. <https://doi.org/10.1128/jb.00909-06>
- Cabello-Yeves, P. J., Callieri, C., Picazo, A., Schallenberg, L., Huber, P., Roda-Garcia, J. J., Bartosiewicz, M., Belykh, O. I., Tikhonova, I. V., Torcello-Requena, A., De Prado, P. M., Puxty, R. J., Millard, A. D., Camacho, A., Rodriguez-Valera, F., & Scanlan, D. J. (2022). Elucidating the picocyanobacteria salinity divide through ecogenomics of new freshwater isolates. *BMC Biology*, 20(1), 1–24. <https://doi.org/10.1186/s12915-022-01379-z>
- Cabello-Yeves, P. J., Picazo, A., Camacho, A., Callieri, C., Rosselli, R., Roda-Garcia, J. J., Coutinho, F. H., & Rodriguez-Valera, F. (2018). Ecological and genomic features of two widespread freshwater picocyanobacteria. *Environmental Microbiology*, 20(10), 3757–3771. <https://doi.org/10.1111/1462-2920.14377>
- Callieri, C., Coci, M., Corno, G., Macek, M., Modenutti, B., Balseiro, E., & Bertoni, R. (2013). Phylogenetic diversity of nonmarine picocyanobacteria. *FEMS Microbiology Ecology*, 85(2), 293–301. <https://doi.org/10.1111/1574-6941.12118>
- Callieri, C., Cronberg, G., & Stockner, J. G. (2012). Freshwater Picocyanobacteria: Single Cells, Microcolonies and Colonial Forms. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (pp. 229–269). Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_8
- Castenholz, R. W., Wilmotte, A., Herdman, M., Rippka, R., Waterbury, J. B., Itean, I., & Hoffmann, L. (2001). Phylum BX. Cyanobacteria. In D. R. Boone, R. W. Castenholz, & G. M. Garrity (Eds.), *Bergey's Manual® of Systematic Bacteriology* (pp. 473–599). Springer New York. https://doi.org/10.1007/978-0-387-21609-6_27

Chen, J., Li, Y., Jing, H., Zhang, X., Xu, Z., Xu, J., & Liu, H. (2022). Genomic and transcriptomic evidence for the diverse adaptations of *Synechococcus* subclusters 5.2 and 5.3 to mesoscale eddies. *New Phytologist*, 233(4), 1828–1842. <https://doi.org/10.1111/nph.17903>

Christmas, N. a. M., Anesio, A. M., & Sánchez-Baracaldo, P. (2015). Multiple adaptations to polar and alpine environments within cyanobacteria: a phylogenomic and Bayesian approach. *Frontiers in Microbiology*, 6(October), 1–10. <https://doi.org/10.3389/fmicb.2015.01070>

Christmas, N. A. M., Barker, G., Anesio, A. M., & Sánchez-baracaldo, P. (2016). Genomic mechanisms for cold tolerance and production of exopolysaccharides in the Arctic cyanobacterium *Phormidesmis priestleyi* BC1401. *BMC Genomics*, 1–14. <https://doi.org/10.1186/s12864-016-2846-4>

Coote, J. G. (1992). Structural and functional relationships among the RTX toxin determinants of Gram-negative bacteria. *FEMS Microbiology Letters*, 88(2), 137–162. <https://doi.org/10.1111/j.1574-6968.1992.tb04961.x>

Cornet, L., Durieu, B., Baert, F., D’hooge, E., Colignon, D., Meunier, L., Lupo, V., Cleenwerck, I., Daniel, H. M., Rigouts, L., Sirjacobs, D., Declerck, S., Vandamme, P., Wilmotte, A., Baurain, D., & Becker, P. (2023). The GEN-ERA toolbox: unified and reproducible workflows for research in microbial genomics. *GigaScience*, 12, 1–10. <https://doi.org/10.1093/gigascience/giad022>

De Coster, W., D’Hert, S., Schultz, D. T., Cruts, M., & Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics*, 34(15), 2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>

Delmont, T. O., & Eren, E. M. (2018). Linking pangenomes and metagenomes: The *Prochlorococcus* metapangenome. *PeerJ*, 2018(1), 1–23. <https://doi.org/10.7717/peerj.4320>

Dessimoz, C., & Gil, M. (2010). Phylogenetic assessment of alignments reveals neglected tree signal in gaps. *Genome Biology*, 11(4), R37. <https://doi.org/10.1186/gb-2010-11-4-r37>

Deter, H. S., Jensen, R. V., Mather, W. H., & Butzin, N. C. (2017). Mechanisms for differential protein production in toxin–antitoxin systems. *Toxins*, 9(7), 1–13. <https://doi.org/10.3390/toxins9070211>

Di Cesare, A., Dzhenbekova, N., Cabello-Yeves, P. J., Eckert, E. M., Slabakova, V., Slabakova, N., Peneva, E., Bertoni, R., Corno, G., Salcher, M. M., Kamburska, L., Bertoni, F., Rodriguez-Valera, F., Moncheva, S., & Callieri, C. (2020). Genomic Comparison and Spatial Distribution of Different *Synechococcus* Phylotypes in the Black Sea. *Frontiers in Microbiology*, 11, 1979. <https://doi.org/10.3389/fmicb.2020.01979>

Doré, H., Farrant, G. K., Guyet, U., Haguit, J., Humily, F., Ratin, M., Pitt, F. D., Ostrowski, M., Six, C., Brillet-Guéguen, L., Hoebeke, M., Bisch, A., Le Corguillé, G., Corre, E., Labadie, K., Aury, J.-M., Wincker, P., Choi, D. H., Noh, J. H., ... Garczarek, L. (2020). Evolutionary Mechanisms of Long-Term Genome Diversification Associated With Niche Partitioning in Marine Picocyanobacteria. *Frontiers in Microbiology*, 11, 567431. <https://doi.org/10.3389/fmicb.2020.567431>

Dufresne, A., Ostrowski, M., Scanlan, D. J., Garczarek, L., Mazard, S., Palenik, B. P., Paulsen, I. T., de Marsac, N. T., Wincker, P., Dossat, C., Ferriera, S., Johnson, J., Post, A. F., Hess, W. R., & Partensky, F. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biology*, 9(5), R90. <https://doi.org/10.1186/gb-2008-9-5-r90>

Durieu, B. (2025). Picocyanobacteria Antarctica. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.30772781>.

Durieu, B., Savaglia, V., Lara, Y., Lambion, A., Pessi, I. S., Vyverman, W., Verleyen, E., & Wilmotte, A. (2025). (Sub-)Antarctic endemic cyanobacteria from benthic mats are rare and have restricted geographic distributions. *Ecography*, (1), 1–12. <https://doi.org/10.1111/ecog.07489>

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 platform for 'omics data. *PeerJ*, 2015(10), 1–29. <https://doi.org/10.7717/peerj.1319>

Everroad, R. C., Woebken, D., Singer, S. W., Burow, L. C., Kyrpides, N., Woyke, T., Goodwin, L., Detweiler, A., Prufert-Bebout, L., & Pett-Ridge, J. (2013). Draft Genome Sequence of an Oscillatorian Cyanobacterium, Strain ESFC-1. *Genome Announcements*, 1(4), 13–14. <https://doi.org/10.1128/genomeA.00527-13>

Everroad, R. C., & Wood, A. M. (2012). Phycoerythrin evolution and diversification of spectral phenotype in marine *Synechococcus* and related picocyanobacteria. *Molecular Phylogenetics and Evolution*, 64(3), 381–392. <https://doi.org/10.1016/j.ympev.2012.04.013>

Fraikin, N., Goormaghtigh, F., & Van Melderen, L. (2020). Type II Toxin-Antitoxin Systems: Evolution and Revolutions. *Journal of Bacteriology*, 202(7), 1–14. <https://doi.org/10.1128/JB.00763-19>

Fucich, D., Marsan, D., Sosa, A., & Chen, F. (2019). Complete Genome Sequence of Subcluster 5.2 *Synechococcus* sp. Strain CB0101, Isolated from the Chesapeake Bay. *Microbiology Resource Announcements*, 8(35), e00484–19. <https://doi.org/10.1128/MRA.00484-19>

Fulford-Smith, S. P., & Sikes, E. L. (1996). The evolution of Ace Lake, Antarctica, determined from sedimentary diatom assemblages. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 124(1), 73–86. [https://doi.org/https://doi.org/10.1016/0031-0182\(96\)01657-4](https://doi.org/https://doi.org/10.1016/0031-0182(96)01657-4)

Garcia, P. S., Jauffrit, F., Grangeasse, C., & Brochier-Armanet, C. (2019). GeneSpy, a user-friendly and flexible genomic context visualizer. *Bioinformatics*, 35(2), 329–331. <https://doi.org/10.1093/bioinformatics/bty459>

Gontcharov, A. A., Marin, B., & Melkonian, M. (2004). Are Combined Analyses Better Than Single Gene Phylogenies? A Case Study Using SSU rDNA and rbcL Sequence Comparisons in the Zygnematophyceae (Streptophyta). *Molecular Biology and Evolution*, 21(3), 612–624. <https://doi.org/10.1093/molbev/msh052>

Goodhead, I., & Darby, A. C. (2015). Taking the pseudo out of pseudogenes. *Current Opinion in Microbiology*, 23, 102–109. <https://doi.org/https://doi.org/10.1016/j.mib.2014.11.012>

- Guindon, S., Delsuc, F., Dufayard, J.-F., & Gascuel, O. (2009). Estimating maximum likelihood phylogenies with PhyML. *Methods in Molecular Biology* (Clifton, N.J.), 537, 113–137. https://doi.org/10.1007/978-1-59745-251-9_6
- Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Systematic Biology*, 59(3), 307–321. <https://doi.org/10.1093/sysbio/syq010>
- Guo, S., Garnham, C. P., Whitney, J. C., Graham, L. A., & Davies, P. L. (2012). Re-Evaluation of a Bacterial Antifreeze Protein as an Adhesin with Ice-Binding Activity. *PLoS ONE*, 7(11), e48805. <https://doi.org/10.1371/journal.pone.0048805>
- Huang, S., Wilhelm, S. W., Harvey, H. R., Taylor, K., Jiao, N., & Chen, F. (2012). Novel lineages of *Prochlorococcus* and *Synechococcus* in the global oceans. *The ISME Journal*, 6(2), 285–297. <https://doi.org/10.1038/ismej.2011.106>
- Huber, P., Diovisalvi, N., Ferraro, M., Metz, S., Lagomarsino, L., Llames, M. E., Royo-Llonch, M., Bustingorry, J., Escaray, R., Acinas, S. G., Gasol, J. M., & Unrein, F. (2017). Phenotypic plasticity in freshwater picocyanobacteria. *Environmental Microbiology*, 19(3), 1120–1133. <https://doi.org/10.1111/1462-2920.13638>
- Hurley, J. M., & Woychik, N. A. (2009). Bacterial Toxin HigB Associates with Ribosomes and Mediates Translation-dependent mRNA Cleavage at A-rich Sites*. *Journal of Biological Chemistry*, 284(28), 18605–18613. <https://doi.org/https://doi.org/10.1074/jbc.M109.008763>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>
- Komárek, J., Johansen, J. R., Smarda, J., & Strunecky, O. (2020). Phylogeny and taxonomy of *Synechococcus*-like cyanobacteria. *Fottea*, 20(2), 171–191. <https://doi.org/10.5507/fot.2020.006>
- Kuo, C. H., & Ochman, H. (2010). The extinction dynamics of bacterial pseudogenes. *PLoS Genetics*, 6(8), 1–7. <https://doi.org/10.1371/journal.pgen.1001050>
- Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H. H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Research*, 35(9), 3100–3108. <https://doi.org/10.1093/nar/gkm160>
- Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, 52(W1), W78–W82. <https://doi.org/10.1093/nar/gkae268>
- Liu, Y., Gao, Z., Liu, G., Geng, Z., Dong, Y., & Zhang, H. (2020). Structural Insights Into the Transcriptional Regulation of HigBA Toxin–Antitoxin System by Antitoxin HigA in *Pseudomonas aeruginosa*. *Frontiers in Microbiology*, 10(January), 1–13. <https://doi.org/10.3389/fmicb.2019.03158>

Loris, R., & Garcia-Pino, A. (2014). Disorder- and dynamics-based regulatory mechanisms in toxin-antitoxin modules. In *Chemical Reviews* (Vol. 114, Number 13, pp. 6933–6947). American Chemical Society. <https://doi.org/10.1021/cr400656f>

Lumian, J., Sumner, D. Y., Grettenberger, C. L., Jungblut, A. D., Irber, L., Pierce-Ward, N. T., & Brown, C. T. (2024). Biogeographic distribution of five Antarctic cyanobacteria using large-scale k-mer searching with sourmash branchwater. *Frontiers in Microbiology*, 15(February), 1–10. <https://doi.org/10.3389/fmicb.2024.1328083>

Lunter, G., Rocco, A., Mimouni, N., Heger, A., Caldeira, A., & Hein, J. (2008). Uncertainty in homology inferences: Assessing and improving genomic sequence alignment. *Genome Research*, 18(2), 298–309. <https://doi.org/10.1101/gr.6725608>

Lupo, V., Roomans, C., Royen, E., Ongena, L., Jacquemin, O., Mullender, C., Kerff, F., & Baurain, D. (2025). Identification and characterization of archaeal pseudomurein biosynthesis genes through pangenomics. *MSystems*, 10(3), 1–24.

Marçais, G., Delcher, A. L., Phillippy, A. M., Coston, R., Salzberg, S. L., & Zimin, A. (2018). MUMmer4: A fast and versatile genome alignment system. *PLOS Computational Biology*, 14(1), e1005944. <https://doi.org/10.1371/journal.pcbi.1005944>

Mareš, J. (2018). Multilocus and SSU rRNA gene phylogenetic analyses of available cyanobacterial genomes, and their relation to the current taxonomic system. *Hydrobiologia*, 811(1), 19–34. <https://doi.org/10.1007/s10750-017-3373-2>

Mikheeva, L. E., Karbysheva, E. A., & Shestakov, S. V. (2013). The role of mobile genetic elements in the evolution of cyanobacteria. *Russian Journal of Genetics: Applied Research*, 3(2), 91–101. <https://doi.org/10.1134/S2079059713020032>

Miladi, H., Soukri, A., Bakhrouf, A., & Ammar, E. (2012). Expression of ferritin-like protein in *Listeria monocytogenes* after cold and freezing stress. *Folia Microbiologica*, 57(6), 551–556. <https://doi.org/10.1007/s12223-012-0172-z>

Ochman, H., & Wilson, A. C. (1987). Evolution in bacteria: Evidence for a universal substitution rate in cellular genomes. *Journal of Molecular Evolution*, 26(1), 74–86. <https://doi.org/10.1007/BF02111283>

Panwar, P., Allen, M. A., Williams, T. J., Hancock, A. M., Brazendale, S., Bevington, J., Roux, S., Páez-espino, D., Nayfach, S., Berg, M., Schulz, F., Chen, I. A., Huntemann, M., Shapiro, N., Kyrpides, N. C., Woyke, T., Elie-fadrosh, E. A., & Cavicchioli, R. (2020). Influence of the polar light cycle on seasonal dynamics of an Antarctic lake microbial community. 8(116), 1–24. <https://doi.org/10.1186/s40168-020-00889-8>

Panwar, P., Williams, T. J., Allen, M. A., & Cavicchioli, R. (2022). Population structure of an Antarctic aquatic cyanobacterium. *Microbiome*, 10(1), 1–22. <https://doi.org/10.1186/s40168-022-01404-x>

- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and meta-genomes. *Genome Research*, 25(7), 1043–1055. <https://doi.org/10.1101/gr.186072.114>
- Pavelich, I. J., Schureck, M. A., Srinivas, P., Blackburn, T. M., Wang, D., Hoffer, E. D., Boamah, M., Zaldana, K., Onuoha, N., Miles, S. J., Grabowicz, M., Okafor, C. D., & Dunham, C. M. (2025). Antitoxin control of optimal transcriptional repression in the atypical HigB-HigA toxin-antitoxin system from *Proteus vulgaris*. *Nucleic Acids Research*, 53(13), 1–15. <https://doi.org/10.1093/nar/gkaf610>
- Powell, L. M., Bowman, J. P., Skerratt, J. H., Franzmann, P. D., & Burton, H. R. (2005). Ecology of a novel *Synechococcus* clade occurring in dense populations in saline Antarctic lakes. *Marine Ecology Progress Series*, 291, 65–80. <https://doi.org/10.3354/meps291065>
- Qin, Q., Muhammad, K., & R., I. J. (2021). The higBA-Type Toxin-Antitoxin System in IncC Plasmids Is a Mobilizable Ciprofloxacin-Inducible System. *MSphere*, 6(3), 10.1128/msphere.00424-21. <https://doi.org/10.1128/msphere.00424-21>
- Scanlan, D. J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W. R., Post, A. F., Hagemann, M., Paulsen, I., & Partensky, F. (2009). Ecological Genomics of Marine Picocyanobacteria. *Microbiology and Molecular Biology Reviews*, 73(2), 249–299. <https://doi.org/10.1128/MMBR.00035-08>
- Schmieder, R., & Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics*, 27(6), 863–864. <https://doi.org/10.1093/bioinformatics/btr026>
- Schuler, G. D., Epstein, J. A., Ohkawa, H., & Kans, J. A. B. T.-M. in E. (1996). Entrez: Molecular biology database and retrieval system. In *Computer Methods for Macromolecular Sequence Analysis* (Vol. 266, pp. 141–162). Academic Press. [https://doi.org/https://doi.org/10.1016/S0076-6879\(96\)66012-1](https://doi.org/10.1016/S0076-6879(96)66012-1)
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539. <https://doi.org/10.1038/msb.2011.75>
- Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D. J., Di Franco, A., Roure, B., Satoh, N., Quéinnec, É., Ereskovsky, A., Lapébie, P., Corre, E., Delsuc, F., King, N., Wörheide, G., & Manuel, M. (2017). A Large and Consistent Phylogenomic Dataset Supports Sponges as the Sister Group to All Other Animals - Forty-two Mat-Meth. *Current Biology*, 27(7), 958–967. <https://doi.org/10.1016/j.cub.2017.02.031>
- Singh, G., Yadav, M., Ghosh, C., & Rathore, J. S. (2021). Bacterial toxin-antitoxin modules: classification, functions, and association with persistence. *Current Research in Microbial Sciences*, 2, 100047. <https://doi.org/10.1016/j.crmicr.2021.100047>

- Song, Y., Luo, G., Zhu, Y., Li, T., Li, C., He, L., Zhao, N., Zhao, C., Yang, J., Huang, Q., Mu, X., Tang, X., Kang, M., Wu, S., He, Y., & Bao, R. (2021). *Pseudomonas aeruginosa* antitoxin HigA functions as a diverse regulatory factor by recognizing specific pseudopalindromic DNA motifs. *Environmental Microbiology*, 23(3), 1541–1558. <https://doi.org/10.1111/1462-2920.15365>
- Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9), 1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>
- Strunecký, O., Ivanova, A. P., & Mareš, J. (2023). An updated classification of cyanobacterial orders and families based on phylogenomic and polyphasic analysis. *Journal of Phycology*, 59(1), 12–51. <https://doi.org/10.1111/jpy.13304>
- Tang, E. P. Y., Tremblay, R., & Vincent, W. F. (1997). Cyanobacterial dominance of polar freshwater ecosystems: are high-latitude mat-formers adapted to low temperature? *Journal of Phycology*, 33(2), 171–181. <https://doi.org/10.1111/j.0022-3646.1997.00171.x>
- Tang, J., Du, L.-M., Liang, Y.-M., & Daroch, M. (2019). Complete Genome Sequence and Comparative Analysis of *Synechococcus* sp. CS-601 (SynAce01), a Cold-Adapted Cyanobacterium from an Oligotrophic Antarctic Habitat. *International Journal of Molecular Sciences*, 20(1), 152. <https://doi.org/10.3390/ijms20010152>
- Taton, A., Grubisic, S., Ertz, D., Hodgson, D. A., Piccardi, R., Biondi, N., Tredici, M. R., Mainini, M., Losi, D., Marinelli, F., & Wilmotte, A. (2006). Polyphasic study of antarctic cyanobacterial strains. *Journal of Phycology*, 42(6), 1257–1270. <https://doi.org/10.1111/j.1529-8817.2006.00278.x>
- Toro, M., Granados, I., Pla, S., Giralt, S., Antoniadis, D., Galán, L., Cortizas, A. M., Lim, H. S., & Appleby, P. G. (2013). Chronostratigraphy of the sedimentary record of Limnopolar Lake, Byers Peninsula, Livingston Island, Antarctica. *Antarctic Science*, 25(2), 198–212. <https://doi.org/DOI:10.1017/S0954102012000788>
- Utter, D. R., Borisy, G. G., Eren, A. M., Cavanaugh, C. M., & Mark Welch, J. L. (2020). Metapangenomics of the oral microbiome provides insights into habitat adaptation and cultivar diversity. *Genome Biology*, 21(1), 1–25. <https://doi.org/10.1186/s13059-020-02200-2>
- Velázquez, D., Rochera, C., Camacho, A., & Quesada, A. (2011). Temperature effects on carbon and nitrogen metabolism in some Maritime Antarctic freshwater phototrophic communities. *Polar Biology*, 34(7), 1045–1055. <https://doi.org/10.1007/s00300-011-0964-7>
- Vincent, W. F. (2007). Cold Tolerance in Cyanobacteria and Life in the Cryosphere. In J. Seckbach (Ed.), *Algae and Cyanobacteria in Extreme Environments* (pp. 287–301). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6112-7_15
- Vincent, W. F., Bowman, J. P., Rankin, L. M., & Mcmeekin, T. A. (2000). Phylogenetic diversity of picocyanobacteria in Arctic and Antarctic ecosystems. *Microbial Biosystems: New Frontiers. Proceedings of the 8th International Symposium on Microbial Ecology*, 1, 317–322.

- Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLOS Computational Biology*, 13(6), e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Wong, K. M., Suchard, M. A., & Huelsenbeck, J. P. (2008). Alignment Uncertainty and Genomic Analysis. *Science*, 319(5862), 473–476. <https://doi.org/10.1126/science.1151532>
- Wood, D. E., Lu, J., & Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. *Genome Biology*, 20(257), 1–13. <https://doi.org/10.1186/s13059-019-1891-0>
- Xia, X., Lee, P., Cheung, S., Lu, Y., & Liu, H. (2020). Discovery of Euryhaline Phycoerythrobilin-Containing *Synechococcus* and Its Mechanisms for Adaptation to Estuarine Environments. *MSystems*, 5(6), e00842–20. <https://doi.org/10.1128/mSystems.00842-20>
- Zhu, X., Li, Q., Yin, C., Fang, X., & Xu, X. (2015). Role of spermidine in overwintering of cyanobacteria. *Journal of Bacteriology*, 197(April), 2325–2334. <https://doi.org/10.1128/JB.00153-15>

5.11. Supplementary information

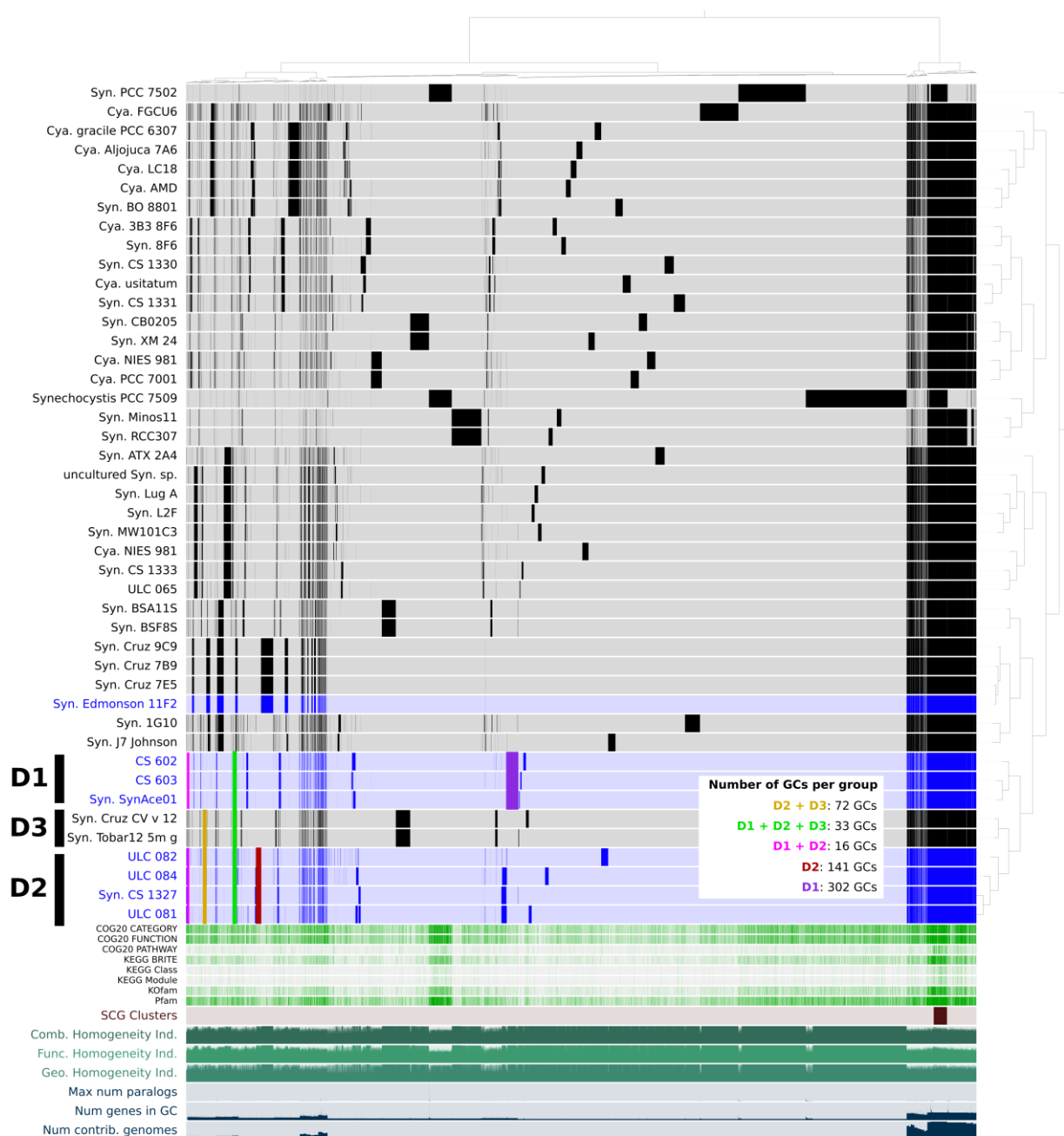


Fig. S5.1. Pangenomic analysis of the 44 cyanobacterial assemblies (minbit = 0.5, mcl inflation = 8) revealing 5 groups of GCs linked to Antarctic genomes of cluster D (groups 'D2 + D3', 'D1 + D2 + D3', 'D1 + D2', 'D2' and 'D1').

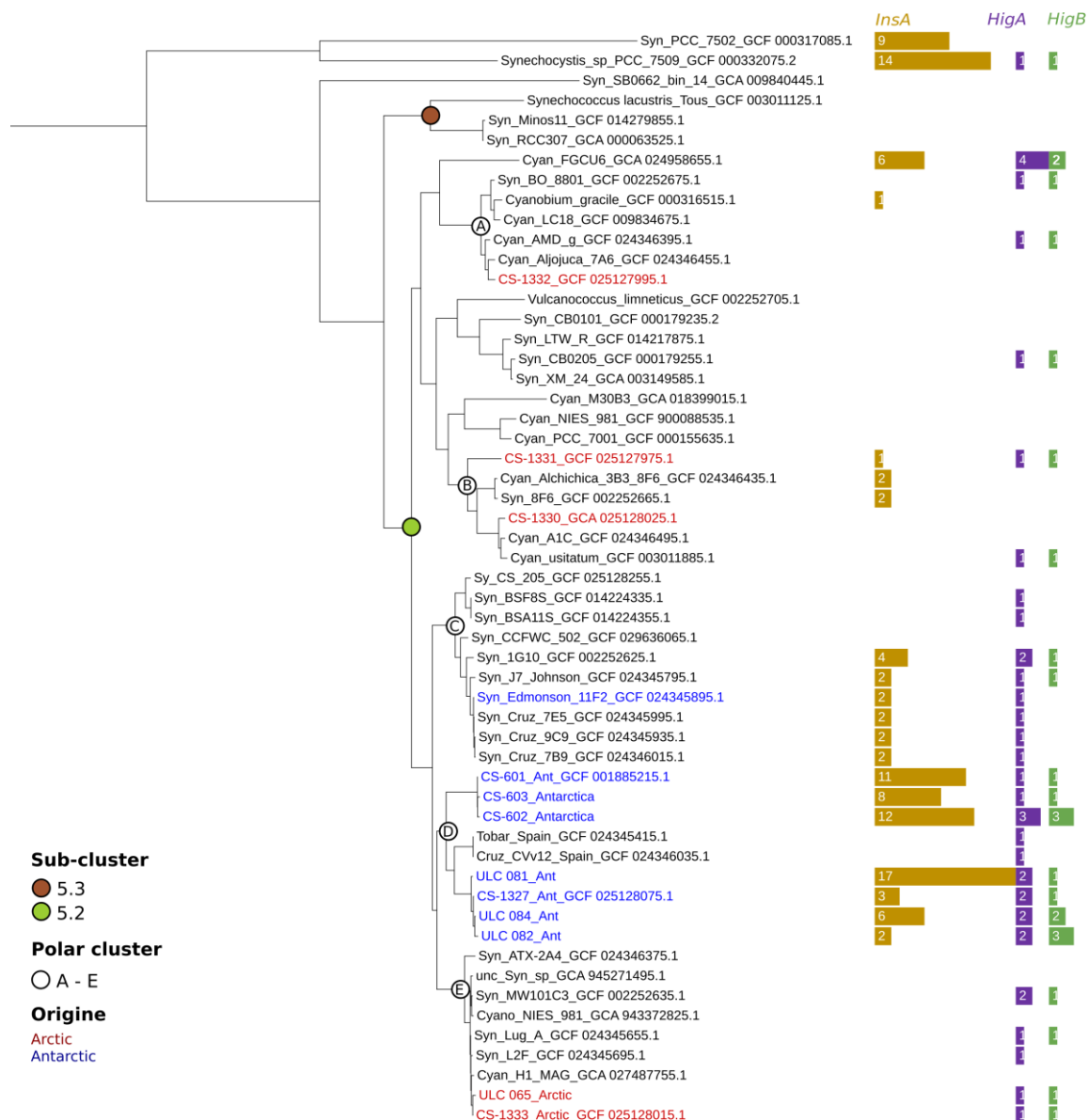


Fig. S5.2. Multilocus maximum likelihood phylogeny with 1000 rapid bootstraps under the PROTGAMMALGF model of 55 cyanobacterial assemblies inferred from a supermatrix of 476 genes totaling 86,369 unambiguously aligned amino acid positions and 0.21% of missing character states. Number of copies of *insA* (gold), *higA* (purple) and *higB* (green) genes predicted by Anvi'o workflow are represented by horizontal barplots.



Fig. S5.3. Genomic contexts of HigA and HigB protein sequences in the 44 assemblies.

The figure in better resolution is available on Figshare and can be downloaded from:

<https://doi.org/10.6084/m9.figshare.33200250>

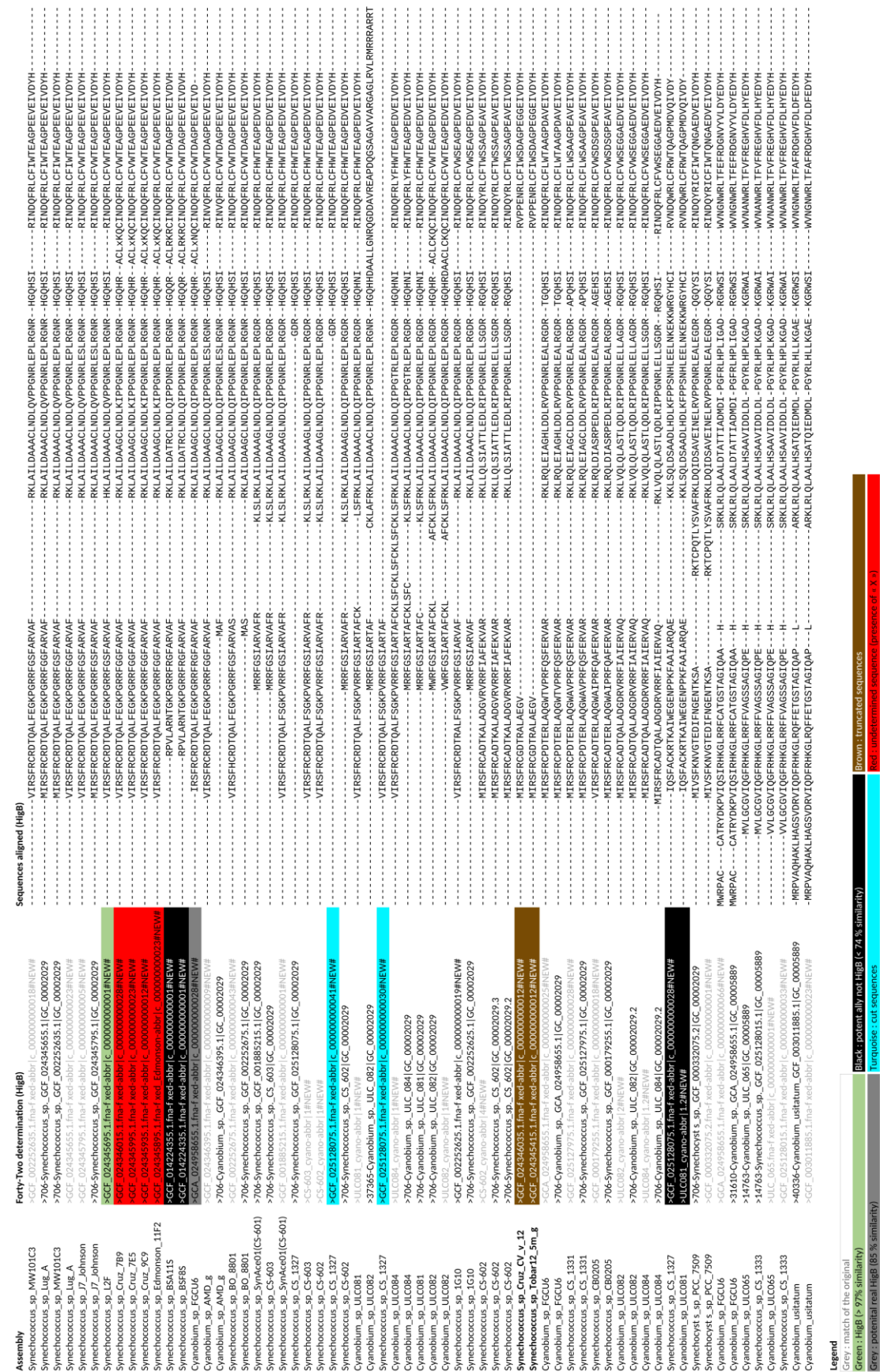


Table S5.1. Contig code, length (bp), taxonomy (determine with kraken2) and redundancy (determine with mummer) of the contigs of the new assemblies. Green: kept in the final assembly. Red: removed from the final assembly because (1) redundant or (2) too small.

# contigs	Contig size (bp)	Taxonomy (kraken2)	Redundancy
ULC084			
1	3073311	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
2	21675	<i>Synechococcus</i> sp. CBW1107 (taxid 2789857)	not redundant
ULC082			
1	1744470	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
2	1140949	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
3	54114	<i>Synechococcus</i> sp. CBW1107 (taxid 2789857)	not redundant
ULC081			
1	3135004	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
4	62081	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
27	1358	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
CS-602			
5	2743189	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
64	84999	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
95	41963	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
107	32814	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
124	26448	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
214	7383	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
243	5530	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	redundant
248	5309	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
263	4622	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
291	3658	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
310	2981	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	redundant
327	2652	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
382	1812	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	redundant
408	1597	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
457	1307	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
508	1170	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	redundant
567	1031	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
CS-603			
4	2671853	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
21	49466	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	/
22	42711	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
33	24602	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	/
57	3484	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	redundant
61	2829	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	not redundant
70	1788	<i>Synechococcus</i> sp. SynAce01 (taxid 1916956)	redundant

72	1522	Synechococcus sp. SynAce01 (taxid 1916956)	not redundant
73	1522	Synechococcus sp. SynAce01 (taxid 1916956)	redundant
74	1151	Synechococcus sp. SynAce01 (taxid 1916956)	/
75	1144	Synechococcus sp. SynAce01 (taxid 1916956)	redundant
76	1103	Synechococcus sp. SynAce01 (taxid 1916956)	redundant
77	899	Synechococcus sp. SynAce01 (taxid 1916956)	
78	484	Synechococcus sp. SynAce01 (taxid 1916956)	
79	382	Synechococcus sp. SynAce01 (taxid 1916956)	
80	382	Synechococcus sp. SynAce01 (taxid 1916956)	
81	379	Synechococcus sp. SynAce01 (taxid 1916956)	
82	294	Synechococcus sp. SynAce01 (taxid 1916956)	
83	288	Synechococcus sp. SynAce01 (taxid 1916956)	
84	260	Synechococcus sp. SynAce01 (taxid 1916956)	
85	255	Synechococcus sp. SynAce01 (taxid 1916956)	
87	200	Synechococcus sp. SynAce01 (taxid 1916956)	
88	183	Synechococcus sp. SynAce01 (taxid 1916956)	

Table S5.2. (Part 1) metadata of the selected publicly available assemblies and the new assemblies of this study. Clu. g.: cluster in the **Fig.5.1**, Co. #: number of contigs, Gen. size (bp): genome size (in bp), Ge. #: number of genes, Av. Gene length: average gene length, Compl.: completeness, Conta.: contamination, Het.: heterogeneity.

Clu. g.	Name_assembly	RefSeq/GenBank / ENA	BioProject	Co. #	Gen. size (bp)	Ge. #	Av. Gene length	GC content	Compl.	Conta.	Het.
A	Cyanobium_sp_Aljojuca_7A6_GCF_024346455_1	GCF_024346455.1	PRJNA718564	34	3131121	3198	896,5178236	0,673054475	99.73	0.82	33.33
A	Cyanobium_sp_AMD_g_GCF_024346395_1	GCF_024346395.1	PRJNA718564	13	3121205	3126	912,768714	0,679395938	99.73	0.27	0.00
A	Cyanobium_gracile_PCC_6307_GCF_000316515_1	GCF_000316515.1	PRJNA158695	1	3342364	3340	901,0526946	0,687071785	99.73	0.00	0.00
A	Synechococcus_sp_BO_8801_GCF_002252675_1	GCF_002252675.1	PRJNA396543	47	3270344	3334	902,6625675	0,691591257	98.91	0.27	0.00
A	Synechococcus_sp_CS_1332_GCF_025127995_1	GCF_025127995.1	PRJNA736758	132	2973784	3028	896,8652576	0,67349747	97.28	0.27	0.00
A	Cyanobium_sp_Copco_Reservoir_LC18_GCF_009834675_1	GCF_009834675.1	PRJNA198871	28	3210832	3205	920,0096724	0,695986897	100.00	0.00	0.00
B	Cyanobium_usitatum_GCF_003011885_1	GCF_003011885.1	PRJNA435996	25	2515215	2701	854,5035172	0,625836757	99.73	0.27	0.00
B	Synechococcus_sp_8F6_GCF_002252665_1	GCF_002252665.1	PRJNA396543	84	2498494	2760	829,7271739	0,658465666	99.46	0.54	0.00
B	Synechococcus_sp_CS_1331_GCF_025127975_1	GCF_025127975.1	PRJNA736758	47	2509836	2800	820,4496429	0,646136242	99.46	0.27	0.00
B	Cyanobium_sp_Alchichica_3B3_8F6_GCF_024346435_1	GCF_024346435.1	PRJNA718564	34	2480576	2734	830,020117	0,659230356	99.46	0.54	0.00
B	Synechococcus_sp_CS_1330_GCA_025128025_1	GCA_025128025.1	PRJNA736758	72	2513790	2834	809,694072	0,626482721	98.91	0.27	0.00
B	Cyanobium_sp_A1C_AMD_GCF_024346495_1	GCF_024346495.1	PRJNA718564	132	2464954	2759	821,5770207	0,632123764	97.69	0.95	25.00
C	Synechococcus_sp_J7_Johnson_GCF_024345795_1	GCF_024345795.1	PRJNA718564	68	2995359	3286	831,2017651	0,652604913	99.73	0.00	0.00
C	Synechococcus_sp_BSF8S_GCF_014224335_1	GCF_014224335.1	PRJNA556564	44	2884859	3114	841,2447013	0,657586385	99.73	0.27	0.00
C	Synechococcus_sp_BSA11S_GCF_014224355_1	GCF_014224355.1	PRJNA556564	42	2883234	3111	842,0819672	0,657742313	99.73	0.27	0.00
C	Synechococcus_sp_Cruz_7B9_GCF_024346015_1	GCF_024346015.1	PRJNA718564	32	3155662	3342	860,5152603	0,652240322	99.73	0.00	0.00
C	Synechococcus_sp_Cruz_7E5_GCF_024345995_1	GCF_024345995.1	PRJNA718564	31	3155391	3340	861,0170659	0,652243415	99.73	0.00	0.00
C	Synechococcus_sp_Edmonson_11F2_GCF_024345895_1	GCF_024345895.1	PRJNA718564	31	3150148	3339	861,1518419	0,652442361	99.73	0.00	0.00
C	Synechococcus_sp_CCFWC_502_GCF_029636065_1	GCF_029636065.1	PRJNA880272	1	2860008	3074	854,3314899	0,660741185	99.73	0.14	100.00
C	Synechococcus_sp_1G10_GCF_002252625_1	GCF_002252625.1	PRJNA396543	84	3330538	3684	815	0,645711293	99.46	0.27	0.00
C	Synechococcus_sp_Cruz_9C9_GCF_024345935_1	GCF_024345935.1	PRJNA718564	33	3152229	3334	861,1817636	0,652211181	99.46	0.00	0.00
C	Synechococcus_sp_CS_205_GCF_025128255_1	GCF_025128255.1	PRJNA736758	53	2360644	2551	853,0784006	0,668900944	97.28	0.14	100.00
D	Synechococcus_sp_SynAce01_GCF_001885215_1	GCF_001885215.1	PRJNA352894	1	2750634	2933	850,4872145	0,639157009	99.73	0.00	0.00
D	ULC_084	GCA_977927545.1	PRJEB103944	2	3094986	3214	878,6228998	0,64932507	99.73	0.00	0.00
D	ULC_082	GCA_977927535.1	PRJEB103944	3	2939533	3087	868,5519922	0,6507748	99.73	0.27	50.00
D	Synechococcus_sp_Cruz_CV_v_12_GCF_024346035_1	GCF_024346035.1	PRJNA718564	52	3022307	3172	860,1469105	0,658529064	99.46	0.72	0.00
D	Synechococcus_sp_Tobar12_5m_g_GCF_024345415_1	GCF_024345415.1	PRJNA718564	47	2951029	3104	859,9416881	0,657971169	99.46	0.27	0.00
D	Synechococcus_sp_CS_1327_GCF_025128075_1	GCF_025128075.1	PRJNA736758	41	2806576	2917	879,8063079	0,652066433	99.46	0.54	0.00

D	ULC_081	GCA_977927525.1	PRJEB103944	3	3198443	3320	871,0346386	0,644653664	99.46	0.54	0.00
D	CS_602	GCA_977927555.1	PRJEB103944	13	2954813	3188	837,3717064	0,636436214	99.46	0.00	0.00
D	CS_603	GCA_977929215.1	PRJEB103944	4	2718915	2890	855,0020761	0,638798197	100.00	0.00	0.00
E	Synechococcus_sp_ATX_2A4_GCF_024346375_1	GCF_024346375.1	PRJNA718564	42	2917741	2867	900,3547262	0,661151213	99.46	0.54	0.00
E	Cyanobium_sp_H1_MAG_00018_GCA_027487755_1	GCA_027487755.1	PRJNA767443	159	2633236	2753	859,9211769	0,673308431	98.91	1.39	66.67
E	Synechococcus_sp_CS_1333_GCF_025128015_1	GCF_025128015.1	PRJNA736758	34	2931480	2907	891,9793602	0,662642761	98.64	0.27	0.00
E	Synechococcus_sp_Lug_A_GCF_024345655_1	GCF_024345655.1	PRJNA718564	33	3010822	3037	875,7688508	0,663080049	100.00	0.54	0.00
E	Cyanobium_sp_NIES_981_GCA_943372825_1	GCA_943372825.1	PRJEB49198	24	3098747	3046	893,1116218	0,660277202	100.00	0.27	0.00
E	uncultured_Synechococcus_sp_GCA_945271495_1	GCA_945271495.1	PRJNA599289	22	3156465	3059	899,1722785	0,661737735	100.00	0.27	0.00
E	Synechococcus_sp_MW101C3_GCF_002252635_1	GCF_002252635.1	PRJNA396543	18	3028460	2996	890,3841789	0,661341408	100.00	0.41	0.00
E	Synechococcus_sp_L2F_GCF_024345695_1	GCF_024345695.1	PRJNA718564	10	3037934	2948	906,5881954	0,662283973	100.00	0.54	0.00
E	ULC_065	GCA_977927515.1	PRJEB103944	1	3174230	2987	911,6494811	0,65908299	100.00	0.54	0.00
/	Cyanobium_sp_M30B3_GCA_018399015_1	GCA_018399015.1	PRJNA721826	1	3180296	3155	896,79271	0,677647615	99.73	0.54	50.00
/	Synechococcus_sp_PCC_7502_GCF_000317085_1	GCF_000317085.1	PRJNA159509	3	3583735	3701	854,5185085	0,406211397	99.76	0.00	0.00
/	Cyanobium_sp_NIES_981_GCF_900088535_1	GCF_900088535.1	PRJEB14035	1	3021545	3021	898,1996028	0,686163205	99.73	0.00	0.00
/	Synechocystis_sp_PCC_7509_GCF_000332075_2	GCF_000332075.2	PRJNA159501	4	4908825	4842	881,1982652	0,41672872	99.67	0.00	0.00
/	Synechococcus_sp_Minor11_GCF_014279855_1	GCF_014279855.1	PRJNA596899	1	2284343	2560	832,5867188	0,604613668	99.64	0.00	0.00
/	Synechococcus_sp_RCC307_GCA_000063525_1	GCA_000063525.1	PRJNA13654	1	2224914	2549	814,5641428	0,608435026	99.64	0.00	0.00
/	Cyanobium_sp_FGCU6_GCA_024958655_1	GCA_024958655.1	PRJNA813570	69	4074513	4280	851,7028037	0,668496885	99.59	0.82	0.00
/	Cyanobium_sp_PCC_7001_GCF_000155635_1	GCF_000155635.1	PRJNA19301	2	2834252	2799	911,7845659	0,687376024	99.46	0.27	0.00
/	Synechococcus_sp_CB0205_GCF_000179255_1	GCF_000179255.1	PRJNA46503	78	2427308	2766	808,0661605	0,629675344	99.18	0.27	0.00
/	Synechococcus_sp_SB0662_bin_14_GCA_009840445_1	GCA_009840445.1	PRJNA555144	85	2164299	2140	919,9233645	0,620727497	97.01	0.82	75.00
/	Vulcanococcus_limneticus_GCF_002252705_1	GCF_002252705.1	PRJNA396543	160	3540491	3668	863,4348419	0,683744586	100.00	1.09	40.00
/	Synechococcus_sp_XM_24_GCA_003149585_1	GCA_003149585.1	PRJNA445564	31	2376968	2618	846,8009931	0,632653868	100.00	0.27	0.00
/	Synechococcus_sp_LTW_R_GCF_014217875_1	GCF_014217875.1	PRJNA645008	1	2415533	2632	831,2028875	0,625626725	100.00	0.27	100.00
/	Synechococcus lacustris GCF_003011125_1	GCF_003011125.1	PRJNA435996	224	2606100	2896	838,2493094	0,517592686	98.55	1.72	75.00
/	Synechococcus sp. CB0101_GCF_000179235_2	GCF_000179235.2	PRJNA46501	1	2606100	3121	817,2399872	0,64059565	100.00	0.54	0.00

Table S5.2. (Part 2) is available on Figshare and can be downloaded from <https://doi.org/10.6084/m9.figshare.33200250>

[illegible]

Table S5.4, Table S5.5 and **Table S5.6** are available on Figshare and can be downloaded from <https://doi.org/10.6084/m9.figshare.33200250>

Chapter 6

General Discussion

Although numerous studies have investigated cyanobacteria in Antarctica, current knowledge of their diversity and distribution remains concentrated on specific geographic areas. Consequently, our understanding of Antarctic cyanobacterial biogeography is so far based on a collection of local-scale studies. This prevents us from robustly testing whether the Antarctic Conservation Biogeographic Regions (ACBRs), which were mainly designed using macroorganisms as a framework, also apply to cyanobacteria. Previous studies on microorganisms have yielded mixed results, with some confirming the relevance of ACBRs (e.g., diatoms; Verleyen et al., 2021) and others showing that ACBRs do not fully reflect the actual distribution of soil bacteria (Varliero et al., 2024). Moreover, Antarctic biogeographic patterns are contrasted for cyanobacteria, with available evidence supporting either high levels of Antarctic endemism (e.g., Pessi et al., 2018; Taton et al., 2006) or the prevalence of cold-adapted ecotypes across both polar regions (e.g. Jungblut et al., 2010; Kleinteich et al., 2017).

The vast majority of microbial biogeographic studies investigate patterns, such as endemism, by sampling entire microbial communities at broad taxonomic resolution using conserved marker genes (Chase & Martiny, 2018). For example, endemism can be determined for each taxon within a microbial community by comparing its marker-gene sequence with all available data in public databases (e.g., NCBI GenBank; Pessi et al., 2018). The genetic region used as a marker, and particularly its genetic resolution, can influence the biogeographic patterns observed (Cho & Tiedje, 2000). A key limitation is that, although these conserved regions can capture a broad range of microbial diversity, they are limited in their ability to detect fine-scale genetic variation (Chase & Martiny 2018). Recent advances in genomics have shown that, despite the existence of a broadly shared core genome among species or strains within the same genus, each taxon also possesses a set of specific genes, some of which may be linked to ecological niches, i.e., the ecogenome (e.g. Doré et al., 2020; Cabello-Yeves et al., 2022). These specific genes might therefore provide complementary information for characterizing Antarctic endemism.

This Ph.D. thesis aimed to improve our understanding of the biogeography of Antarctic cyanobacteria and to identify and characterize their genomic specificities. The research was structured in two main parts. In the **first part**, we used cyanobacterial 16S rRNA amplicon sequences from (1) 84 lakes representing a broad sampling across eight ACBRs in Antarctica and two sub-Antarctic islands (**Chapter 2**), and (2) 100 terrestrial samples representing a more local

and denser sampling effort in the Sør Rondane Mountains, Dronning Maud Land (**Chapter 3**). With the lacustrine dataset of Chapter 2, we tested the effects of both environmental parameters (pH and salinity) and geographic distance on the structure of cyanobacterial communities. We also determined and compared the proportions of endemic taxa in Antarctica and the sub-Antarctic islands. For the terrestrial dataset presented in Chapter 3, the availability of higher-resolution sampling and more extensive metadata allowed to describe cyanobacterial communities more precisely in relation to their microhabitats and to perform a finer-scale analysis of the effects of soil characteristics on community structure.

In the **second part** of this Ph.D. thesis, we used genomics to search for genetic specificities associated with the Antarctic environment, with the aim of providing complementary information for characterizing Antarctic endemism. To this end, we developed two case studies focusing on genera with cosmopolitan records in Antarctica, according to the criteria applied in Chapters 2 and 3: the first focused on filamentous strains (*Phormidesmis*) and the second on unicellular strains (*Synechococcus*-like picocyanobacteria). Although these genera are distributed across broad geographic scales, genomic analyses can reveal unique features of their Antarctic ecotypes. To address this hypothesis, we first used an *a priori* approach based on genes of interest (GOIs) identified from the literature to characterize the genome of *Phormidesmis priestleyi* ULC007. This study was conducted in 2017 and, although it has not yet been published, remains relevant to the comparative framework developed in this thesis and was therefore included as **Chapter 4**. Then, we used an *a posteriori*, or discovery-based, approach to identify Antarctic-specific genes in Antarctic picocyanobacteria of subcluster 5.2 (**Chapter 5**).

Together, these results help us to refine the concept of endemism. Because endemism is a central criterion in biodiversity conservation, our results provide additional data to support conservation policies for Antarctic ecosystems.

6.1. Biogeographic patterns of Antarctic cyanobacterial communities

6.1.1. From continental to local scales

In **Chapter 2**, benthic microbial mats were sampled from 84 lakes across eight ACBRs and two sub-Antarctic islands. By analyzing 16S rRNA gene sequences from these environmental samples, we observed that, at continental geographic scale, the structure of cyanobacterial lacustrine communities is significantly different between Antarctica and sub-Antarctic islands,

but not among ACBRs. This confirms that the Antarctic continent is geographically isolated for microorganisms, and that microbial communities do not reflect ACBR subdivisions (Pessi et al., 2018; Varliero et al., 2024). Moreover, both Antarctic and sub-Antarctic lakes contain a high number of potentially endemic taxa. This is in line with previous studies on Antarctic microbiota (Taton et al., 2006; De Wever et al., 2009; Vyverman et al., 2010; Pessi et al., 2018). Interestingly, the proportion of endemic taxa is higher among the low-abundance taxa (taxa encompassing <1% of the total taxa reads in a sample) than among the high-abundance taxa (taxa encompassing >90% of the total taxa reads) in both Antarctic and sub-Antarctic lakes. Also interestingly, rare (low-abundance) taxa are restricted to a few areas within Antarctica, whereas high-abundant taxa are more widely distributed across the continent. This is consistent with previous observations that the vast majority of rare bacterial taxa have restricted geographic distributions and occur at low relative abundances (Pessi et al., 2018; Bickel & Or, 2021).

At a broad geographical scale (Antarctica and sub-Antarctic islands), distance is almost three times more structuring than environmental parameters (pH and conductivity) ([E]=4.9%, [S]=14.8%; **Table 6.1**). This difference in proportion between distance and environmental parameters seems to decrease at continental scale ([E]=6.5%, [S]=9.2%; **Table 6.1**). To test this pattern on a smaller geographical scale, a higher resolution sampling is mandatory. Yet this result should be taken with caution because our variation partitioning analysis only explains 18.7% (Antarctica and sub-Antarctic islands) or 14.6% (Antarctica only) of cyanobacterial community structures. In this analysis, the relative importance of distance is overestimated because only two environmental variables were recorded. More environmental variables should be considered to increase the explained fraction of the variation. In practice, a complete quantification of all selective variables is nearly impossible to achieve, so the distance effect will probably be almost always overestimated (Hanson et al., 2012).

Response variables (model type)	Env. Variables [E]	Spatial filters (PCNMs) [S]	Fraction of variation as partial adj. R ²				
			[E]	[S]	[E S]	[S E]	[E+S]
Antarctica + Sub-Ant. Islands (71 samples)	Cond + pH	1, 2, 3, 4, 6, 9	4.9	14.8	4.0	13.8	18.7
Antarctica (57 samples)	Cond + pH	1, 2, 3	6.5	9.2	5.4	8.2	14.6

Table 6.1. Variation partitioning showing the relative contributions of spatial (PCNMs) and environmental (pH and conductivity) variables to cyanobacterial community structure at different spatial scales. The calculation of PCNMs is detailed in the methodology of Chapter 2. This table is a modification of **Table S2.4** of Chapter 2, with the addition of variation partitioning for Antarctica samples only (second row).

In **Chapter 3**, by analyzing 16S rRNA gene sequences from 100 environmental samples collected from terrestrial habitats at a local scale (Sør Rondane Mountains, in Dronning Maud Land), we observed that different rock substrates host distinct cyanobacterial communities. Hence, granitic soils host the highest diversity, with numerous filamentous cyanobacterial families. Communities on other types of substrate were less diversified and generally less abundant (in terms of biomass), with the extreme case of the moraine samples of Dry Valley, which harbour very low abundances of cyanobacterial reads and show the lowest ASV richness among all substrate types. Consequently, sites with different substrate types host distinct cyanobacterial communities. At this local scale (Sør Rondane Mountain), the impact of environmental parameters is important (mostly pH, TN and N-NO_3^- and conductivity), almost comparable to the effect of distance. This confirms previous studies showing that pH (Rousk et al., 2010; Griffiths et al., 2011; Bottos et al., 2020; Santini et al., 2022), conductivity (Lozupone & Knight, 2007; X. Wang et al., 2022) and nutrient content (Andrew et al., 2012; Cruz-Martínez et al., 2012) are among the most important abiotic factors structuring terrestrial bacterial communities.

Interestingly, the relative contribution of environmental parameters and geographic distance to explain cyanobacterial community structures is lower in the lacustrine dataset (**Chapter 2**) than in the terrestrial dataset (**Chapter 3**). This could be explained by the higher homogeneity of habitats in the lacustrine compared to the terrestrial dataset. Indeed, lacustrine samples are mostly freshwater and show some variation in pH and conductivity, but overall exhibit limited environmental contrasts, except for a few samples (**Table S2.1, Chapter 2**). These lakes might therefore act as “replicate refugia” located in different parts of the continent, where cyanobacteria encounter comparable environmental conditions. Conversely, terrestrial samples encompass a variety of substrates and microniches, each differently driving cyanobacterial community composition, but at the same time geographically structured, so that distance is correlated with variations of environmental parameters (**Table S3.7, Chapter 3**). This correlation between environmental characteristics and geographic distance, together with the stronger contrasts in environmental characteristics among terrestrial samples compared with the lacustrine dataset, probably increases the effects of both environmental parameters and geographic distance on cyanobacterial community structures. As the lacustrine and terrestrial datasets differ in both sampling resolution and geographic spatial scale, this comparison should nonetheless be interpreted with caution and requires a more robust analysis (more details in **Section 6.3.1 and 6.3.4**).

6.1.2. Endemism comparison between aquatic and terrestrial communities

Is the proportion of endemism different between lacustrine and terrestrial Antarctic habitats? To answer this question, we can examine the case of Dronning Maud land, where both lacustrine and terrestrial datasets are available. Because taxa were delineated using two different approaches - 99% OTUs in Chapter 2 and ASVs in Chapter 3 - alpha diversity could not be directly compared. Indeed, these two approaches differ in taxonomic resolution, as OTU-based datasets artificially reduce community richness compared with ASV-based datasets. However, as a first approximation, the proportion of endemic taxa can be compared. As a reminder, the proportion of endemic taxa in a sample or group of samples is calculated by first determining the **geographic distribution** of each taxon. To do this, the genetic sequence of each taxon is compared with all similar sequences ($\geq 99\%$ gene sequence similarity) available in the NCBI GenBank database. Each taxon can then be assigned to one of the following categories: endemic, polar, alpine, or cosmopolitan (see **Section 2.3.4** for more details). Once the geographic distribution of each taxon within a sample is known, the relative proportion of each category (e.g., endemic) can be calculated by pooling taxa by category. This approach is based on relative abundance, thereby canceling the influence of differences in taxon counts between OTU-based and ASV-based datasets. The only requirement is to use the same database to assign geographic distributions to taxa across both datasets.

In practice, we first pooled the lacustrine samples from Chapter 2 by ACBR, retaining only ACBRs with a minimum of five samples to ensure statistical robustness. We then determined the geographic distribution of taxa (endemic, polar, alpine, or cosmopolitan) from these pools, as well as the geographic distribution of taxa of Chapter 3. The lacustrine samples from Dronning Maud Land were then compared with the terrestrial samples from the same region, as well as with lacustrine samples from the other ACBRs (**Fig. 6.1**).

Interestingly, the differences in the proportion of endemics between lacustrine and terrestrial sites do not appear to be significant. Terrestrial communities seem to harbour slightly more endemic taxa than lacustrine communities, particularly within the abundant fraction of the communities. Stronger environmental filtering in terrestrial habitats compared with lacustrine habitats may explain this trend. Although, to our knowledge, no study has directly shown that Antarctic terrestrial cyanobacterial communities contain more endemic taxa than lacustrine communities, this pattern is consistent with several observations from literature. In Antarctic lakes, some cyanobacterial phylotypes are widely distributed across distinct ice-free regions, whereas potentially endemic phylotypes tend to be rare and restricted to few lakes (Pessi et al., 2018; Chapter 2). By contrast, terrestrial Antarctic microenvironments appear to be strongly

structured by environmental constraints, particularly water availability and endolithic habitats (Rego et al., 2019). More generally, environmental harshness and geographic isolation have been proposed as important drivers of microbial regionalization and endemism in Antarctica (Vyverman et al., 2010). Therefore, the slightly higher proportion of endemic taxa observed in terrestrial communities could reflect stronger environmental filtering and lower habitat buffering compared with lacustrine environments, although this hypothesis requires further testing with better suited sampling designs. Dispersal, however, is more difficult to compare between our terrestrial and aquatic studies. Indeed, in Antarctica, different soil substrates should act as island refuges (Lebre et al., 2021), whereas aquatic habitats such as perennially ice-covered lakes can provide stable, low-disturbance environments with persistent benthic microbial mats, where cyanobacteria can colonize, adapt, and maintain complex communities over time (Zhang et al., 2015; Greco et al., 2020).

Previous comparisons of terrestrial and aquatic habitats revealed that there is a general lack of differentiation for prokaryotes according to habitat type, but that filamentous cyanobacteria in aquatic habitats are significantly different from those in terrestrial habitats (Obbels et al., 2016). In our results, although we do not have enough data to assess the significance of this difference, it appears not important (**Fig. 6.1**), which may be due to physical connection between aquatic and terrestrial habitats. Indeed, it has been shown, in Arctic for example, that freshwater bacterial taxa can originate from surrounding soil communities (Crump et al., 2012). Conversely, some Antarctic cyanobacteria may possess an “escape mechanism” whereby benthic communities form so-called “lift-off” mats that rise from the lake bottom to the surface (Dudeja et al., 2010). This process may facilitate their subsequent dispersal into other aquatic and terrestrial environments (Wharton Jr et al., 1983; Tanabe & Kudoh, 2012). In Antarctica, habitats are also connected like, e.g., in the McMurdo Dry Valley where glaciers and soils are connected to lakes via streams (Takacs-Vesbach et al., 2010).

Our analysis also suggests that differences between terrestrial and aquatic communities are smaller than between aquatic community assemblages from different areas (e.g., between Dronning Maud Land and Victoria Land). However, this observation should be taken with caution and requires better sampling, because it is based only on one terrestrial site (more details in **Section 6.3**).

Polar taxa (i.e., present only in Antarctica and Arctic) were recorded in both aquatic and terrestrial habitats (**Fig. 6.1**). This might suggest a biogeographic connection and a regular exchange between the two polar regions (Kleinteich et al., 2017). However, in our terrestrial dataset as well as in the lacustrine dataset, the polar fraction is the least represented category, compared with

the alpine, endemic, and cosmopolitan fractions. Pessi et al. (2018) also observed this trend in lacustrine habitats.

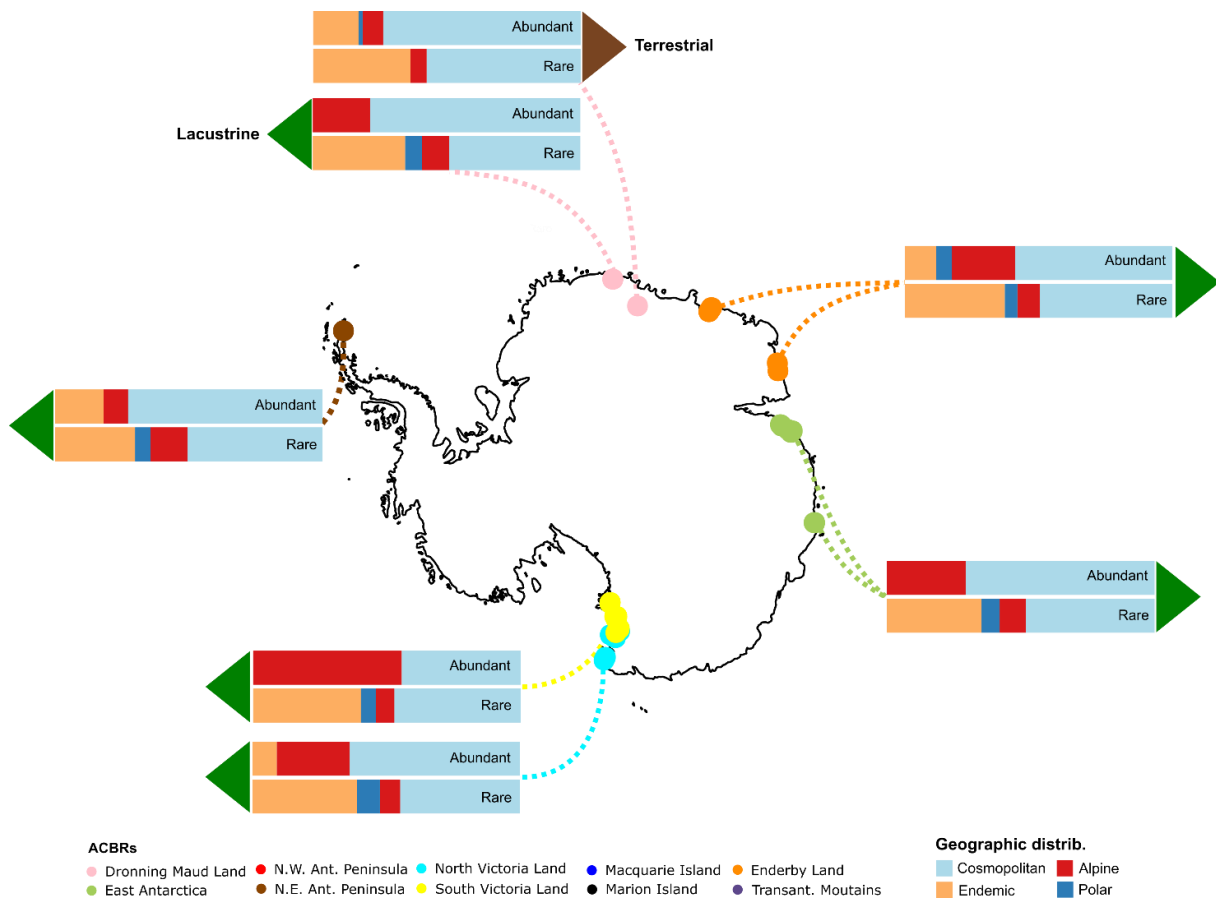


Fig. 6.1. Comparison of the geographic distribution of taxa (OTUs or ASVs) among the different ACBRs. In Dronning Maud Land, lacustrine (**Chapter 2**) and terrestrial (**Chapter 3**) communities are compared. The relative proportion of high-abundance taxa (i.e., taxa encompassing >90% of the total taxa reads in a sample) and the low-abundance taxa (i.e., those encompassing <1%), classified as ‘endemic’, ‘polar’, ‘alpine’ and ‘cosmopolitan’ (see **Chapter 2**) are expressed as number of taxa, based on presence/absence data).

We should note that, within our two datasets (**Chapter 2** and **3**), almost all genera comprise both endemic and cosmopolite OTUs/ASVs. This is because our criterion for defining endemism - matching only Antarctic hits at a 99% 16S rRNA similarity threshold - is finer than the accepted threshold for genus delineation, which is 94.5% (Yarza et al., 2014).

Our comparison between aquatic and terrestrial habitats has several limitations and should therefore be regarded as preliminary. It remains an open question that requires further investigation, including optimized sampling, more comprehensive metadata collection, and harmonized methodologies for taxon delineation (more details in **Section 6.3**).

In the next section, we will focus on cyanobacterial genomics to determine whether it could help to refine the concept of endemism. For example, *Cyanobium* sp. ULC084 is considered cosmopolitan according to the criteria used in **Chapters 2 and 3**, but our pangenomic analysis in **Chapter 5** actually identified Antarctic-specific genes in its genome.

6.2. Characterization of the specificities of Antarctic genomes

6.2.1. *A priori* (GOIs) and *a posteriori* (discovery-based) approaches

In **Chapter 4**, we performed a comparative genomic analysis of one Antarctic genome assembly (*Phormidesmis priestleyi* ULC007), one Arctic one (*Phormidesmis priestleyi* BC1401) and 70 non-polar genome assemblies, using an *a priori* approach based on a set of genes of interest (GOIs) related to stress response and cold adaptation. We observed that some genes had higher copy numbers in polar genomes than in non-polar genomes, including *fbaA* (involved in carbohydrate degradation), genes involved in exopolysaccharides (EPS) biosynthesis, *dnaK* (involved in DNA replication and stress response), and gene coding for universal stress protein UspA (involved in stress response). At a broader metabolic level (functional categories), polar assemblies contained more genes related to choline and betaine biosynthesis (compatible solutes), DNA repair mechanisms, (EPS) biosynthesis and *dnaK* genes cluster. These results demonstrate the adaptation of *Phormidesmis priestleyi* ULC007 to the cold temperatures and stress conditions of Antarctica: EPS buffers the impact of freeze-thaw cycles by limiting ice nucleation around cells (Vincent, 2007), compatible solutes reduce the freezing point of intracellular fluids and limit cell desiccation by reducing the quantity of water required to maintain osmotic equilibrium under increasing salinity (Welsh, 2000; Zakhia et al., 2008), DnaK is induced under cold conditions and, together with DnaJ, helps maintain the native conformation of regulatory proteins such as DnaA, GyrA and RecA (DNA repair), which influence DNA topology (Varin et al., 2012).

This chapter based on analyses conducted in 2017 had several limitations, including the scarcity of polar samples and the lack of control for phylogenetic relationships (see **Section 6.3.1 and 6.3.3** for more details). Therefore, the genes identified as enriched in the two polar *Phormidesmis* sp. genomes should be interpreted with caution.

More robustly, Effendi et al. (2022) compared the genome of the Antarctic *Nostoc* sp. strain SO-36 with closely related available genomes to identify Antarctic-specific features. They showed that genes involved in EPS biosynthesis are conserved within the Nostocaceae, with the Antarctic *Nostoc* SO-36 harbouring simpler EPS-biosynthetic mechanisms than the other related assemblies. Similarly, Marcello (2025) selected a more complete set of GOIs linked to cold

adaptation, together with an updated set of cyanobacterial assemblies. He used 21 cold-tolerance genes based on Christmas et al. (2016) and Barria et al. (2013), 7 genes related to membrane fluidity, and 10 photosynthesis genes encoding membrane-bound PS I and II protein complexes. He selected 183 assemblies from temperate environments and 43 from cold-freezing environments. He found no significant differences in gene counts between cyanobacteria from cold-freezing and temperate environments: both the gene-counts heatmap and PCA analysis clustered genomes according to phylogenetic relationships rather than environmental conditions (Marcello 2025). That was the case for the whole dataset, as well as for *Synechococcus* and *Microcystis* analyzed independently. The statistical analysis restricted to filamentous cyanobacteria remained inconclusive because of the limited number of available assemblies (Marcello 2025). Marcello (2025) also showed that amino acid substitution patterns in cold-freezing cyanobacteria are highly variable. Some patterns align with expectations for cold tolerant microorganisms, such as a lower Arg/Lys ratio, and a higher proportion of residues that promote flexibility in coils, loops, or solvent-exposed regions compared with non-cold tolerant microorganisms, whereas many others deviate from these predictions (Marcello, 2025). However, he also shows that GC content distributions differ significantly between temperate and cold environments within specific phylogenetic groups (i.e., *Synechococcus*, *Microcystis*), but not at the broader cyanobacterial phylogenetic level (Marcello, 2025).

An alternative way to find Antarctic-specific genes is to move from an *a priori* approach to an *a posteriori* (or discovery-based) approach. In a discovery-based approach, we do not select a set of GOs in advance; rather, we compare orthologous genes clusters (GCs) among assemblies in order to identify, *a posteriori*, the GCs that are unique to the group of assemblies of interest. In **Chapter 5**, we applied this approach to genomes newly sequenced in that study, complemented by high-quality publicly available genomes from subcluster 5.2 picocyanobacteria, as defined by Dufresne et al. (2008), to identify genes present in all Antarctic assemblies but absent from non-Antarctic ones. Subcluster 5.2 comprises assemblies originating from a wide range of habitats (i.e., marine, temperate freshwater, high-altitude, polar). It is important to recognize that the “Antarctic habitat” represents a multifactorial variable that shares several characteristics with Arctic, high-altitude, and desert habitats, such as extreme cold, dryness, nutrient limitation, and low water availability. These shared environmental constraints may promote convergent genomic adaptations (Cao et al., 2020; Contreras et al., 2023). By accounting for the phylogenetic structure of the dataset, we identified 16 GCs that are present in all Antarctic genomes but absent from non-Antarctic ones, thus corresponding to the Antarctic ecogenome. Interestingly, although the Antarctic genomes belonging to subcluster 5.2 are considered cosmopolitan according to the

criteria used in **Chapters 2 and 3**, this discovery-based genomic approach revealed the presence of Antarctic-specific genes. We discuss how this finding may provide additional insight into the concept of endemism in **Section 6.4**.

6.2.2. Antarctic-specific genes: comparison between picocyanobacteria and *Phormidesmis* groups

Do Antarctic filamentous cyanobacteria of the genus *Phormidesmis* have more Antarctic-specific genes than Antarctic picocyanobacteria? To answer this question, we should update the analysis of *Phormidesmis priestleyi* ULC007 using a discovery-based approach, applying the same methodology as used for the picocyanobacteria in **Chapter 5**. To do so, we retrieved all Phormidesmidaceae and Leptolyngbyaceae assemblies from GTDB (version r232, 23 April 2026) and retained only the high-quality assemblies (completeness >90%, contamination <5%), in order to place *P. priestleyi* ULC007 in its phylogenomic context. Since the analysis presented in **Chapter 4**, which dates from 2017, the number of available genomes has increased considerably and now includes close relatives of *P. priestleyi* ULC007. This assembly grouped within a cluster we named cluster A (>70% ANI with *P. priestleyi* ULC007), that encompasses cyanobacteria from diverse habitats, including the Arctic genomes already used in **Chapter 4**, and another Antarctic *Phormidesmis* sp. available since April 2024 (**Fig. 6.2**). To confirm the topology, we selected only the 39 assemblies from cluster A and two outgroups, allowing us to increase the size of the gene supermatrix used for the phylogenomic analysis (**Fig. 6.3A**). We then performed pangenomic analyses on cluster A, as well as on a series of smaller clusters corresponding to increasing ANI thresholds with ULC007 (B: >72%, C: >75% and D: 82%). These analyses aimed to evaluate the number of genes specific to Antarctic assemblies (**Fig. 6.3B**). Hence, when considering cluster A, *P. priestleyi* ULC007 (named “*Phormidesmis A priestleyi A*” in GTDB) and the other Antarctic strain (*Phormidesmis A sp038227635*, %ANI with ULC007 = 81.8%) shared 37 specific genes (Antarctic-specific, i.e., present in the two Antarctic genomes but absent from all other genomes of the cluster).

Interestingly, compared with the 37 Antarctic-specific genes, only 4 polar-specific genes were identified (**Fig. 6.3B**). More precisely, the Arctic assembly and ULC007 share only 5 GCs, even though they are phylogenetically closer to each other (82.6% ANI) than the two Antarctic strains are to one another (81.8% ANI). Similarly, the other Antarctic assembly and the Arctic assembly, despite their higher phylogenetic similarity (87.3% ANI), share only 21 GCs, fewer than the two more distant Antarctic assemblies (81.8% ANI, 37 GCs). These results indicate that higher

phylogenetic similarity (vertical inheritance) does not necessarily imply a higher number of shared specific genes. Instead, the opposite trend is observed: the two Antarctic assemblies share more specific genes with each other than either does with the Arctic assembly, independently of the phylogenetic signal. It should be noted that the Antarctic and Arctic genomes have comparable levels of completeness and contamination, suggesting that this pattern is unlikely to result from artificial biases related to genome quality. Different, non-mutually exclusive hypotheses could explain this pattern that suggests a decoupling between phylogenetic relatedness and specific gene content. Antarctic genomes may share more specific genes with each other than Arctic and Antarctic genomes either because these genes were newly acquired, e.g., through horizontal gene transfer within similar ecological niches sharing the same microbial communities, or because older genes were selectively retained in Antarctic lineages, while being lost or pseudogenized in non-Antarctic relatives. Such processes are consistent with the dynamic nature of prokaryotic pangenomes, in which accessory gene repertoires, which may reflect ecological adaptation, can represent a substantial proportion of the genome compared to the core gene repertoire (McInerney et al., 2017; Maistrenko et al., 2020).

Decreasing the phylogenetic depth increased the number of Antarctic-specific genes: 38 GCs for cluster B (>72% ANI with ULC007), 98 for cluster C (>75% ANI with ULC007) and 105 for cluster D (82% ANI with ULC007). This suggests that broad comparisons may conflate ancient phylogenetic differences with ecological specificities, whereas comparisons among closer relatives can better reveal gene repertoires associated with Antarctic environmental filtering.

Compared with the patterns found for picocyanobacteria in **Chapter 5**, the number of Antarctic-specific genes is more than twice as high in Antarctic *Phormidesmis* (37 GCs) as in picocyanobacteria (16 GCs), even though the *Phormidesmis* analysis was conducted in a broader phylogenetic context (>75% ANI) than the picocyanobacterial analysis (>77% ANI).

Further analyses are required to explain why the *Phormidesmis* group contains a higher number of Antarctic-specific genes (37) than the picocyanobacteria (16). Indeed, several characteristics of the two datasets could have opposing effects on this comparison, including differences in phylogenetic depth, sample size and density, genome size, morphological complexity, the Antarctic habitats represented, and the potential effects of artificial selection following strain isolation and maintenance in culture collections (see **Section 6.3** for more details).

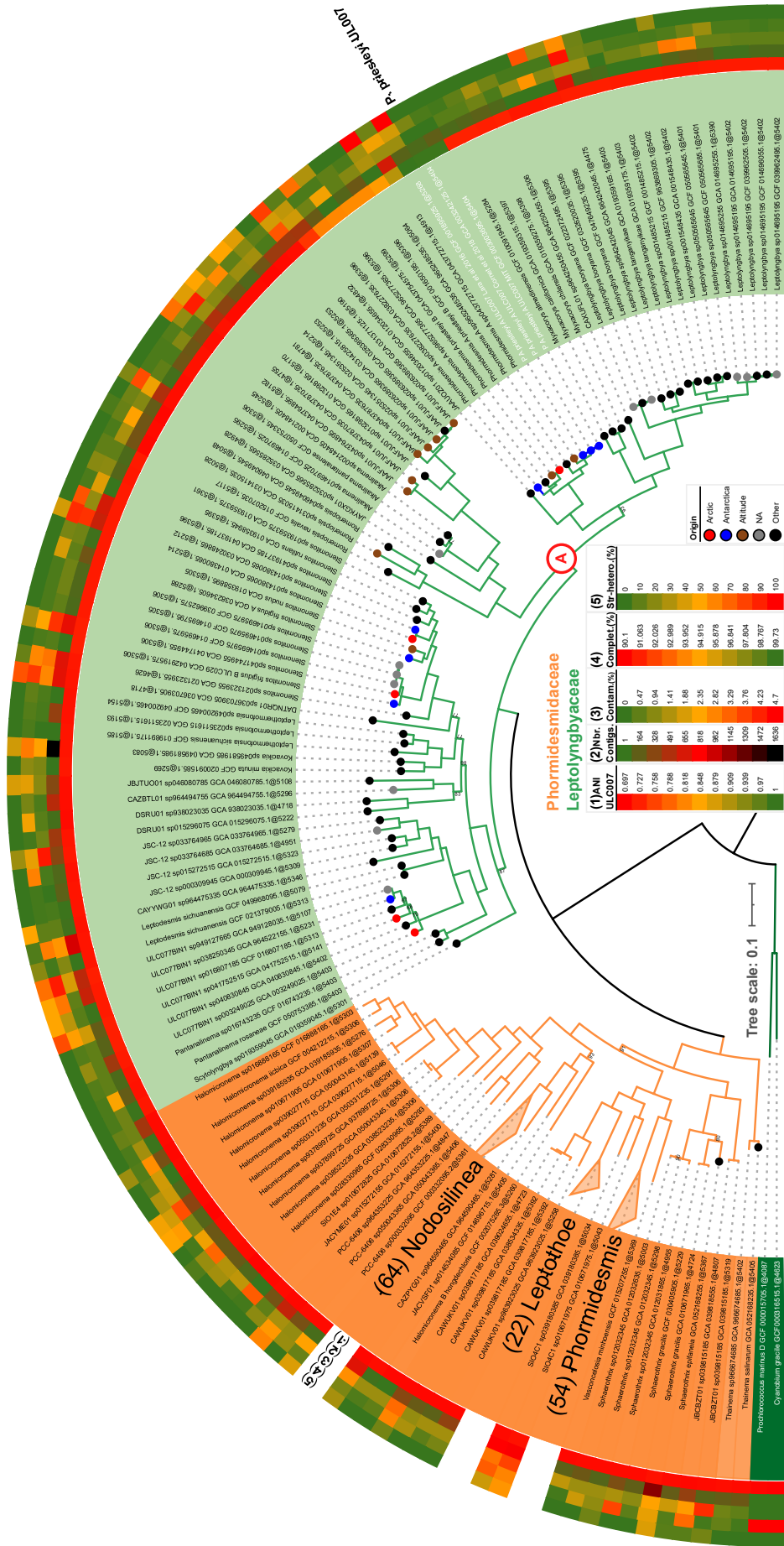


Fig. 6.2. Phylogenomic context of the strain *Phormidismis priestleyi* ULC007 showing the cluster A (in red) within the Leptolyngbyaceae. All public assemblies of Leptolyngbyaceae and Phormidaceae families available on GTDB were retrieved, then the good quality ones (>90% completeness and <5% contamination) were retained. Multilocus maximum likelihood phylogeny with 1000 rapid bootstraps under the PROTAMMALGF model of the 262 assemblies inferred with RAXML (Stamatakis, 2006) from a supermatrix of 47 genes totaling 5406 unambiguously aligned amino acid positions and 3.08% of missing character states.

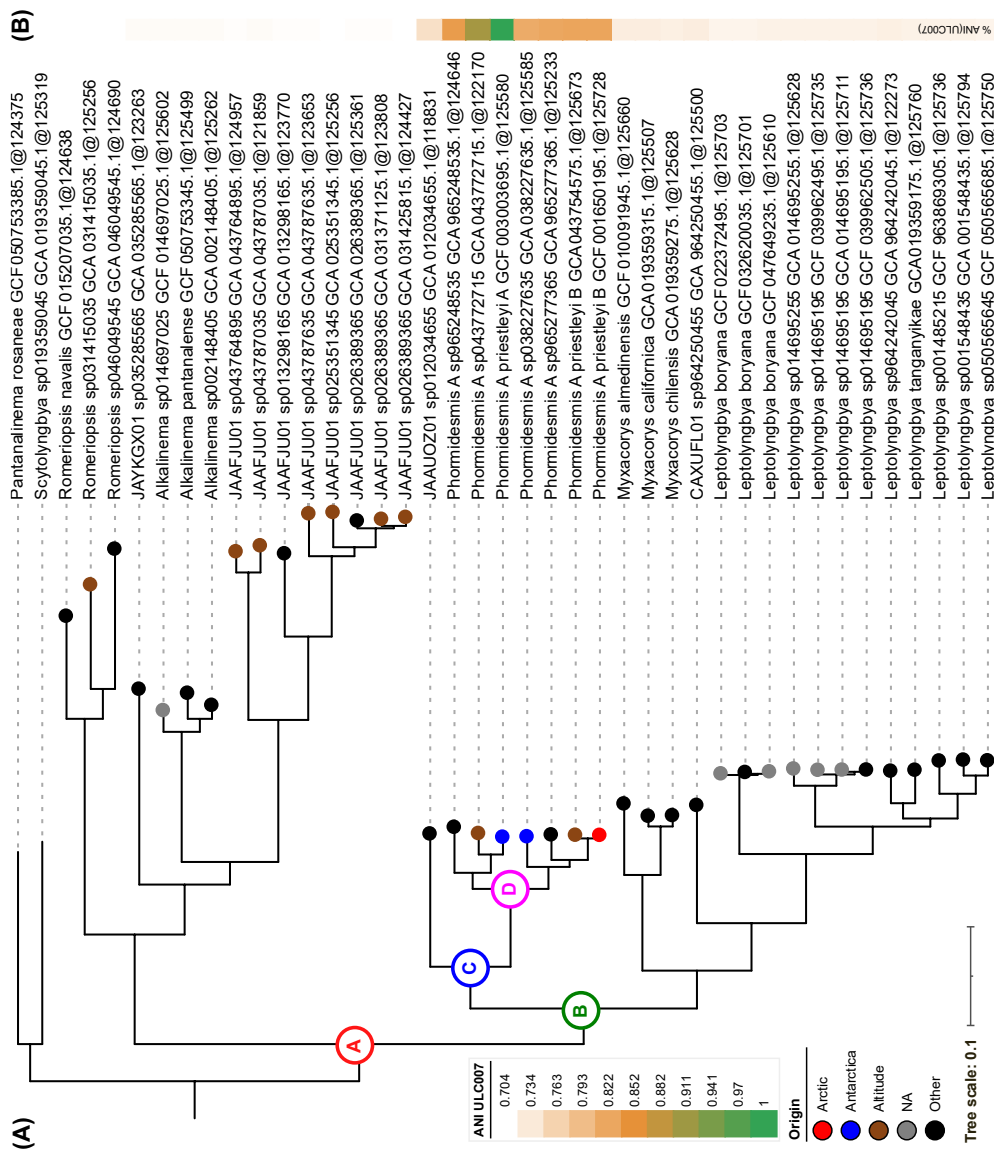


Fig 6.3. (A) Multilocus maximum likelihood phylogeny with 1000 rapid bootstraps under the PROTAMMALGF model of the 39 assemblies of cluster A and two outgroups inferred with RaxML (Stamatakis, 2006) from a supermatrix of 512 genes totaling 125823 unambiguously aligned amino acid positions and 0.71% of missing character states. **(B)** Number of unique genes for the different categories (Antarctic, Altitude *Phormidesmis*, Other *Phormidesmis*, Polar, Third pole *Phormidesmis*, *Phormidesmis*), at different ANI thresholds (70, 72, 75 and 82% to ULC007) corresponding to nested phylogenetic clusters. % ANI drawn on the plot were calculated with pyANI. The values were slightly different when calculated with fastANI (used in Chapter 5): cluster A = 75.3% ANI, cluster B = 75.7% ANI, cluster C = 76.9% and cluster D = 81.9% ANI with ULC007.

6.2.3. Determination of the origin of the Antarctic-specific genes

To investigate the origin of the Antarctic-specific genes in both Antarctic *Phormidesmis* and picocyanobacteria, and to track potential horizontal gene transfers (HGTs), profile Hidden Markov Models (pHMMs) were built from the 37 and 16 Antarctic-specific genes, respectively, and searched against GTDB database r232, that contain 189,801 representative bacterial species, using HMMER v.3.4 (Finn et al., 2011), adapting the protocol of Hambücker et al. (2026) (**Table 6.2**). Eight of the 37 *Phormidesmis* GCs and two of the 16 picocyanobacterial GCs contained a maximum of 50% cyanobacterial taxa among the first 200 hits, making them candidates for potential HGTs. Inspection of the ten corresponding gene trees was then carried out, complemented by functional annotations and 3D protein structure modeling, as illustrated in **Fig. 6.4**. The two picocyanobacterial GCs are predicted to be potential HGTs. For *Phormidesmis*, two GCs are potential HGTs, another one is a potential HGT, but not in Antarctica, and one GC is a potential innovation of the Antarctic *Phormidesmis* (**Table 6.2**). The detailed results for each GCs are available in **Table S6.1** online (Figshare: <https://doi.org/10.6084/m9.figshare.33200250>).

As mentioned above, a general limitation of our analyses is that the identification of Antarctic-specific genes is sensitive to sampling (see **Section 6.3.1** for more details). Future work should therefore aim to investigate the origin of these Antarctic-specific genes more robustly, while also exploring other genomic patterns, such as gene enrichment (see **Section 6.4** for more details).

Antarctic-specific GCs	GCs candidates ¹	Potential HGT ²	Potential innovation ³
<i>Phormidesmis</i>			
37	8	3	1
Picocyanobacteria			
16	2	2	0

Table 6.2. Identification of potential HGT events in the Antarctic-specific GCs, performed in collaboration with Mick Van Vlierberghe and Louise Hambücker. ¹To target HGT, only the GCs containing a maximum of 50% of cyanobacterial taxa among the first 200 hits were retained for subsequent analysis OmpaPa (<https://metacpan.org/dist/Bio-MUST-Apps-OmpaPa>; D. Baurain). ²The candidates were then further filtered based on their e-value and sequence length using. The filtered sequences were then aligned using MAFFT v.7 (Katoh & Standley, 2013). Alignment sites were conserved if they had at most 50% gaps and the sequences with a minimum of 50% length (with respect to the longest unaligned sequence) were exported using ali2phylipl.pl (<https://metacpan.org/release/Bio-MUST-Core>; D. Baurain). A gene tree was then inferred from the filtered alignment using IQ-TREE v.3 (Wong et al., 2026) using the LG4X substitution model and 1000 ultrafast bootstrap replicates. For each GC, functional protein domains were annotated with InterProScan v.5 (Jones et al., 2014) and the 3D protein structures were modeled using the AlphaFold3 server (Abramson et al., 2024). The resulting structures were visualized and aligned using ChimeraX v.1.11 (Meng et al., 2023). ³When the analysis did not identify any homologous sequence in GTDB, the GC was considered to be a potential innovation.

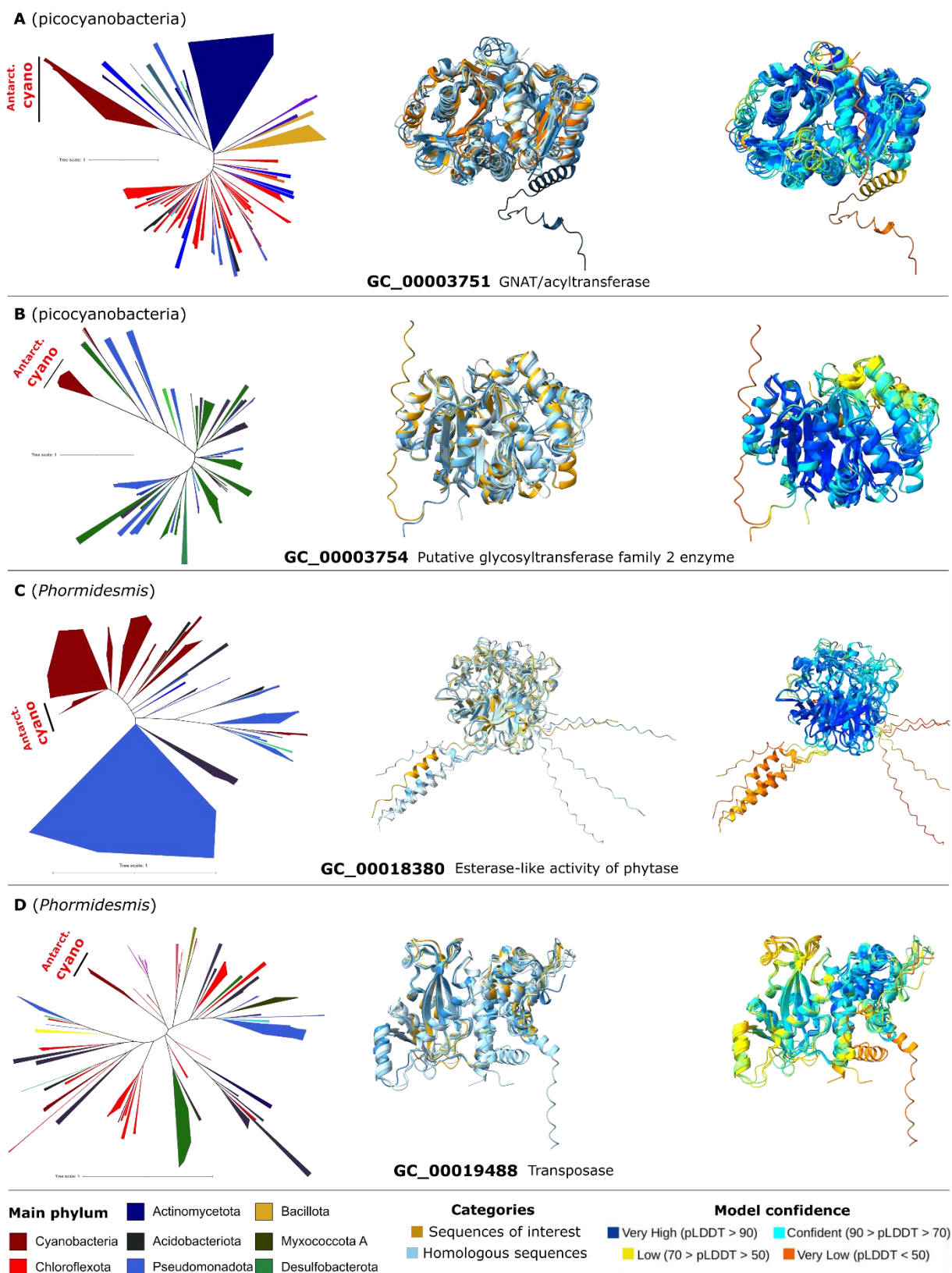


Fig. 6.4. Examples of Antarctic-specific genes obtained through potential HGT. **Left:** gene trees showing the placement of the Antarctic cyanobacterial GC sequences of interest (in red, with brackets) and their homologous sequences identified through the pHMM analysis (see **Table 6.2** for methodological details). **Middle:** 3D structures of the Antarctic sequences (orange) and their homologues (blue). **Right:** model confidence for the 3D structures. Inspection of the trees suggests potential HGT events (**A**, **B** and **D**) from

non-cyanobacterial taxa to the Antarctic genomes (Antarct. cyano). Panel **C** shows that GC_00018380 originated within cyanobacterial lineage. If it was acquired through HGT, the donor was likely another cyanobacterial taxon. 3D protein structures are conserved among the different GCs, as shown by the structural similarity between the Antarctic sequences of interest and their homologous.

6.3. Methodological considerations and limitations

6.3.1. Sampling issues

One of the main limitations of this thesis concerns sampling. Indeed, both within and across the different chapters, the **number** and **types** of samples may strongly influence the results. This limitation applies to both diversity analyses (**Chapters 2 and 3**) and comparative genomic analyses (**Chapters 4 and 5**).

The impact of sample **size** is illustrated by the comparative analysis presented in **Chapter 4** (*Phormidesmis* analysis), which included only one Antarctic and one Arctic assembly. Consequently, it was not possible to identify Antarctic-specific genes, i.e., the Antarctic ecogenome, but only strain-specific genes (singletons). We addressed this limitation for picocyanobacteria in **Chapter 5** (Durieu et al., 2025). Indeed, while Tang et al. (2019) identified 901 genes unique to a single Antarctic genome, corresponding to strain-specific genes, we found that this number decreased to only 16 gene clusters (GCs) when closely related Antarctic genomes were included in the analysis, representing the Antarctic ecogenome. This example clearly illustrates the strong influence of sampling effort on comparative genomic analyses. Indeed, a gene may appear unique simply because Antarctic genomes are under-sampled, or because available assemblies incompletely represent the true accessory gene pool (Méric et al., 2014; Ceres et al., 2022).

Along the same lines, we included more picocyanobacterial Antarctic genomes (seven) in **Chapter 5** than in the *Phormidesmis* analysis (two) of **Chapter 4**. However, there is no phylogenetic continuum among all Antarctic picocyanobacterial genomes: instead, they form two groups of very closely related genomes (~99% ANI), subgroups D1 and D2, separated by a genetic distance almost similar to that measured between the two Antarctic *Phormidesmis* genomes (~81% ANI). In spite of this comparable distance, the higher number of genomes in the picocyanobacterial dataset reduces the number of potentially shared genes.

The impact of sample **type** is illustrated by the comparison between **Chapters 4 and 5**. Indeed, *Phormidesmis* genomes are substantially larger (~5.7 Mb) than picocyanobacterial genomes (~3 Mb). This biases the comparison because larger genomes are likely to have broader adaptive

potential than smaller genomes (Larsson et al., 2011), which might increase the number of Antarctic-specific genes detected in larger genomes. A normalization by genome size should therefore be implemented to enable a more robust comparison (Dewar et al., 2025). Related to genome size, here, we compare multicellular filamentous organisms (*Phormidesmis*) with unicellular organisms (picocyanobacteria). As multicellularity involves more complex biological mechanisms, such as intracellular communication, cell differentiation, and molecular transfers between cells, than unicellularity (Herrero et al., 2016), their physiology and ecology differ, which may interfere with the comparison of Antarctic-specific genes.

6.3.2. Lack of (meta)data

The availability of **data** on Antarctic cyanobacteria progressively decreases from morphological observations to genetic data (e.g., 16S rRNA gene sequences), metagenomes, and, ultimately, complete genomes. This limitation affects estimates of endemism: because current databases represent only a subsample of the true diversity, sequence similarity-based approaches likely overestimate the proportion of endemic taxa. Indeed, sequences are compared only against the limited diversity currently available in databases, and the absence of a match does not necessarily indicate real endemism. For comparative genomic analyses, the accessibility and identification of suitable genomes remain major limitations. Based on the literature, we retrieved 42 Antarctic genome assemblies that had previously been investigated in scientific publications. However, a direct search of public databases such as NCBI GenBank yielded a larger number of Antarctic assemblies (yet many without an associated publication). Thus, a substantial pool of potentially informative Antarctic genomic data remains underutilized in public databases and has yet to be incorporated into comparative analyses.

The availability and quantity of **metadata** also represent a major limitation across the different analyses presented in this thesis. More metadata is required to enable a robust comparison between the lacustrine cyanobacterial communities examined in **Chapter 2** and the terrestrial communities studied in **Chapter 3**. Indeed, “aquatic” and “terrestrial” are broad ecological variables that encompass a wide range of environmental parameters (e.g., nutrients, pH, salinity, water availability), many of which may covary. Such metadata would make it possible to refine these comparisons by identifying the environmental parameters driving variation within each habitat separately, as in **Chapter 3**, as well as between habitat types. Similarly, in the comparative genomic analyses presented in **Chapter 4** and **5**, the lack of environmental metadata limits our

ability to explain the genomic differences observed between Antarctic and non-Antarctic assemblies.

It should be kept in mind that the scarcity of samples and associated metadata is an inherent limitation of Antarctic research, as sample collection and in situ measurements are logistically more challenging and costly than in more accessible regions of the world.

6.3.3. Phylogenetic sampling depth and density, and pangenomic issues

In this thesis, **phylogenetic depth** and **density** is another dimension of our sampling limitations. In **Chapter 4**, we compared the Antarctic genome with a closely related Arctic genome and more distantly related non-Antarctic genomes, without robustly controlling for phylogenetic relationships. This prevented us from distinguishing the ecological signal, linked to adaptation to Antarctica, from the long-term evolutionary signal (Dewar et al., 2025). Therefore, the genes identified as enriched in the two polar *Phormidesmis* genomes should be interpreted with caution, as additional closely related genomes are required to confirm (or infirm) this pattern.

Furthermore, the datasets used in **Chapters 4** and **5** differ in their phylogenetic depth. Although the difference is small, the *Phormidesmis* dataset includes more distantly related assemblies (>75% ANI) than the picocyanobacterial dataset (>77% ANI). Because the *Phormidesmis* dataset encompasses a broader diversity of non-Antarctic genomes, genes present in Antarctic genomes are mechanically more likely to have homologues in at least one non-Antarctic genome. Consequently, such genes would not be classified as Antarctic-specific, potentially reducing the number of putative Antarctic-specific gene clusters (GCs). Fewer Antarctic-specific GCs would therefore be expected in the *Phormidesmis* analysis than in the picocyanobacterial analysis. However, we observed the opposite trend.

In addition to the potential biological explanations discussed above (**Section 6.3.2**), one way to account for this surprising result is that, in the *Phormidesmis* analysis (**Fig. 6.3A**), the phylogenetic distance between cluster D and its closest relatives (JAAUOZ01 sp012034655 GCA_012034655.1) is greater than the distance separating the Antarctic-containing polar group D from its closest relative, cluster E, in the picocyanobacterial analysis (**Fig. 5.1, Chapter 5**). Consequently, in the *Phormidesmis* analysis, Antarctic assemblies are less likely to have matches with vertically inherited genes in non-Antarctic genomes than in the picocyanobacterial analysis. This could explain the higher number of Antarctic-specific GCs identified in the *Phormidesmis* analysis compared to the picocyanobacterial analysis.

The phylogenetic distance threshold used to select samples for pangenomic analyses is also critical. Extending pangenomic analyses beyond the species level to encompass broader taxonomic groups, such as genera or families, poses important methodological challenges, particularly for the identification of orthologous genes through the definition of appropriate clustering thresholds. As phylogenetic distance increases, sequence similarity decreases and synteny becomes less conserved, making orthologs inference more difficult (Sarton-Lohéac et al., 2025). As a result, sequence identity thresholds that are appropriate for species-level analyses (e.g., $\geq 95\%$) may artificially reduce the estimated core genome while inflating the accessory genome when applied to more distantly related taxa (A. Guerra, 2026).

6.3.4. Geographic location, habitat types and strain isolation

When comparing **Chapters 2** and **3**, sampling also has limitations with respect to **geographic location**. Indeed, the lacustrine (**Chapter 2**) and terrestrial (**Chapter 3**) datasets differ in both their geographic spatial **scale** and sampling **resolution**. Consequently, comparisons of the relative effects of environmental parameters (pH and conductivity) and geographic distance on the structure of lacustrine and terrestrial cyanobacterial communities should be interpreted with caution. Furthermore, the sampling strategy should be uniformized between terrestrial and aquatic habitats (**Fig. 6.1**) to enable a more robust comparison of the proportion of endemic taxa. Sampling should target both habitat types across multiple ACBRs. Indeed, our preliminary comparison is based on a single ACBR (Donning Maud Land). Including terrestrial habitats from additional ACBRs would enable comparisons between aquatic and terrestrial communities at the same spatial scale. Sampling resolution should also be comparable among ACBRs, which is not currently the case for our aquatic dataset. Finally, aquatic and terrestrial samples should be collected in close proximity to one another to facilitate habitat comparisons while minimizing the confounding effects of geographic distance and broader climatic differences.

When comparing **Chapters 4** and **5**, the sampling also differs in terms of the Antarctic **habitats** represented, which may lead to different adaptative strategies and thus influence the comparison. The two *Phormidesmis* genomes come from two nearby freshwater lakes in the Larsemann Hills. By contrast, the picocyanobacteria originate from more distinct environments: subgroup D1 comes from brackish lakes in the Vestfold Hills, whereas subgroup D2 comes from freshwater lakes in the South Shetland Islands. The higher similarity in habitat type among *Phormidesmis* genomes than among picocyanobacterial genomes could explain the higher

number of shared genes, as cyanobacterial gene repertoires have been shown to be strongly associated with habitat type and ecological constraints (Chen et al., 2021).

Finally, the analyses presented in **Chapters 4 and 5** should take into account that these Antarctic strains were isolated and have since been maintained in laboratory cultures, a process that may have led to artificial selection, as previously reported for cyanobacteria (Kanesaki et al., 2012; Koskinen et al., 2023). Adaptation to culture conditions may have differed among taxa, and because their genomes were sequenced long after isolation, it is not possible to distinguish adaptations to Antarctic environments from adaptations to laboratory culture conditions. Further studies are therefore needed to assess the extent of this potential bias.

6.3.5. Sources of bias and signal robustness in amplicon sequencing

Efficient **DNA extraction** from environmental microbial cells is essential for molecular ecology studies, as incomplete or selective cell lysis can bias estimates of community composition (Wintzingerode et al., 1997). This issue is particularly important for morphologically diverse cyanobacteria (Fiore et al., 2000), some of which possess mucilaginous envelopes or sheaths that reduce the efficiency of standard cell-lysis protocols. Despite using a previously validated extraction protocol (Pessi et al., 2016), these limitations should be considered when interpreting the results of **Chapters 2 and 3**. In particular, estimates of the relative abundance of different cyanobacterial taxa may be biased by differences in their morphologies and susceptibility to cell lysis. Another consideration is the choice of **PCR primers**, as it influences the phylogenetic breadth of the resulting amplicons. This issue is discussed in **Section 6.4.2**. During **PCR amplification**, low abundant chimeric artefacts may be generated (Haas et al., 2011). In addition, errors also occur during **sequencing**, including substitution miscalls that may artificially inflate inferred sequence diversity (Schirmer et al., 2015), as well as tag switching or index misassignment, which can result in incorrect assignment of sequences to samples, false-positive taxa, and batch effects (Schnell et al., 2015; Jia et al., 2022).

Due to such amplicon sequencing errors, avoiding the artificial inflation of the molecular taxa inferred and the distortion of abundance profiles requires several preprocessing steps (e.g., quality filtering and denoising, detection of chimeric sequences). The inferred molecular taxa depend on the algorithms used for this preprocessing. A wide range of bioinformatic tools has been developed to correct sequencing errors and delineate molecular taxa using two main approaches: distance-based Operational Taxonomic Units (OTUs) (e.g., UPARSE, OptiClust) and denoising-based Amplicon Sequence Variants (ASVs) (e.g., DADA2, UNOISE, Deblur). There is

currently no consensus regarding the relative performance of the different approaches. Several studies have highlighted the respective advantages and limitations of OTU- and ASV-based methods. For example, ASV algorithms produce highly reproducible outputs but may suffer from over-splitting (due to e.g., sequencing errors or intragenomic variation in multiple 16S rRNA genes copies), whereas OTU algorithms can absorb such minor sequence variation but may be more susceptible to over-merging (Fares et al., 2025).

Liu et al. (2023) have shown that OTU- and ASV-based approaches can lead to similar taxonomic profiles and alpha diversity indices, despite considerable differences in the number of inferred OTUs or ASVs (Liu et al., 2023). Conversely, some authors have argued that one approach may be more advantageous than the other. For instance, OTU-based methods have been reported to underestimate diversity and distort dominance and evenness indices compared to ASV-based approaches (Fasolo et al., 2024). Therefore, the use of OTU-based approach in **Chapter 2** and an ASV-based approach in **Chapter 3** constitutes a limitation when comparing the results.

Replicate analyses (e.g., Wang et al., 2018; Barnes et al., 2020; Nilsen et al., 2024) and artificial community (i.e., mock community) analyses (e.g., Wang et al., 2018; Müller and Nebel, 2021; Fares et al., 2025) are often performed to compare algorithm outputs across datasets, ultimately allowing to assess the robustness of the diversity estimates. Accordingly, we performed replicates analyses and artificial community analyses in **Chapters 2** and **3**.

In **Chapter 2**, we used a distance-based approach (UPARSE; Edgar, 2013) to delineate molecular taxa. The replicate analysis shows a robust signal, as replicates shared similar abundances and grouped together in the nMDS ordination (**Fig. S2.1; Chapter 2**). The mock community analysis was less conclusive. It consisted of cyanobacterial strains from the BCCM/ULC collection that were pooled together after PCR amplification. However, the DNA quality was insufficient to obtain reliable results, and the strains were not sequenced independently, preventing us from comparing the abundance profile of the mock community with those of the corresponding individual strains. We learned from this inconclusive experiment and prepared a valid mock community for **Chapter 3**, consisting of six cyanobacterial strains, from five taxonomic orders, pooled into a single sample. Each strain was also sequenced separately. We observed a low level of noise, reflected by the persistence of comparatively low-abundance ASVs in the raw data (**Table S3.2; Chapter 3**). However, the overall signal was robust, as the mock community accurately reflected the expected combination of the individual strains, in the expected proportions. The residual low-abundance ASVs may result from PCR or sequencing biases that have previously been reported in the literature, including batch effects, residual PCR or sequencing errors and tag switching (see **Section 6.3.5**). Consequently, applying an abundance threshold is necessary, together with the

analysis of replicates as an additional quality-control step, and the definition of appropriate limits for interpreting the results. Although this residual noise represents a limitation of **Chapter 3**, applying an appropriate threshold helps prevent the overinterpretation of fine-scale patterns.

As mentioned above, replicate analysis provides a means of assessing the robustness of the signal. In **Chapter 3**, we increased the number of replicate samples compared with **Chapter 2**. The comparison of inter- and intra-replicate variability shows that Bray-Curtis distance was significantly higher between different replicates than within the same replicate group, indicating that the biological signal exceeded technical variability (**Fig. 6.4A**). Using this replicate dataset, we compared two denoising-based approaches implemented in the RASPAM workflow, developed in collaboration with Marie Harmel (Harmel et al., 2025). The first approach relies on ZOTU taxonomic units computed by UNOISE (Edgar, 2016), whereas the second relies on ASV taxonomic units computed by DADA2 (Callahan et al., 2016). The first approach yielded 570 ZOTUs, while the second yielded 312 ASVs. Nevertheless, both approaches produced similar taxonomic profiles at the order level (**Fig. 6.4B**). Although benchmarking studies have reported differences between ASV- and ZOTU-based approaches (Liu et al., 2023; Fares et al., 2025), our results suggest that they can produce comparable patterns depending on the taxonomic resolution considered (**Fig. 6.4B**).

To conclude this discussion on signal robustness, in **Chapter 3**, we used the CyanoSeq database (Lefler et al., 2023) to apply the most detailed and recent assignments, rather than the time-consuming approach used in **Chapter 2**, which relied on BLAST searches followed by manual verification of phylogenetic trees. Thus, the CyanoSeq database combines the speed of automated taxonomic assignment with sufficient resolution for cyanobacteria while relying on an up-to-date taxonomy. In contrast to commonly used databases such as SILVA (Quast et al., 2013) and Greengenes (DeSantis et al., 2006), it incorporates the frequent revisions of cyanobacterial taxonomy and the continuous description of new taxa, thereby improving the accuracy of sequence classification and ecological interpretation.

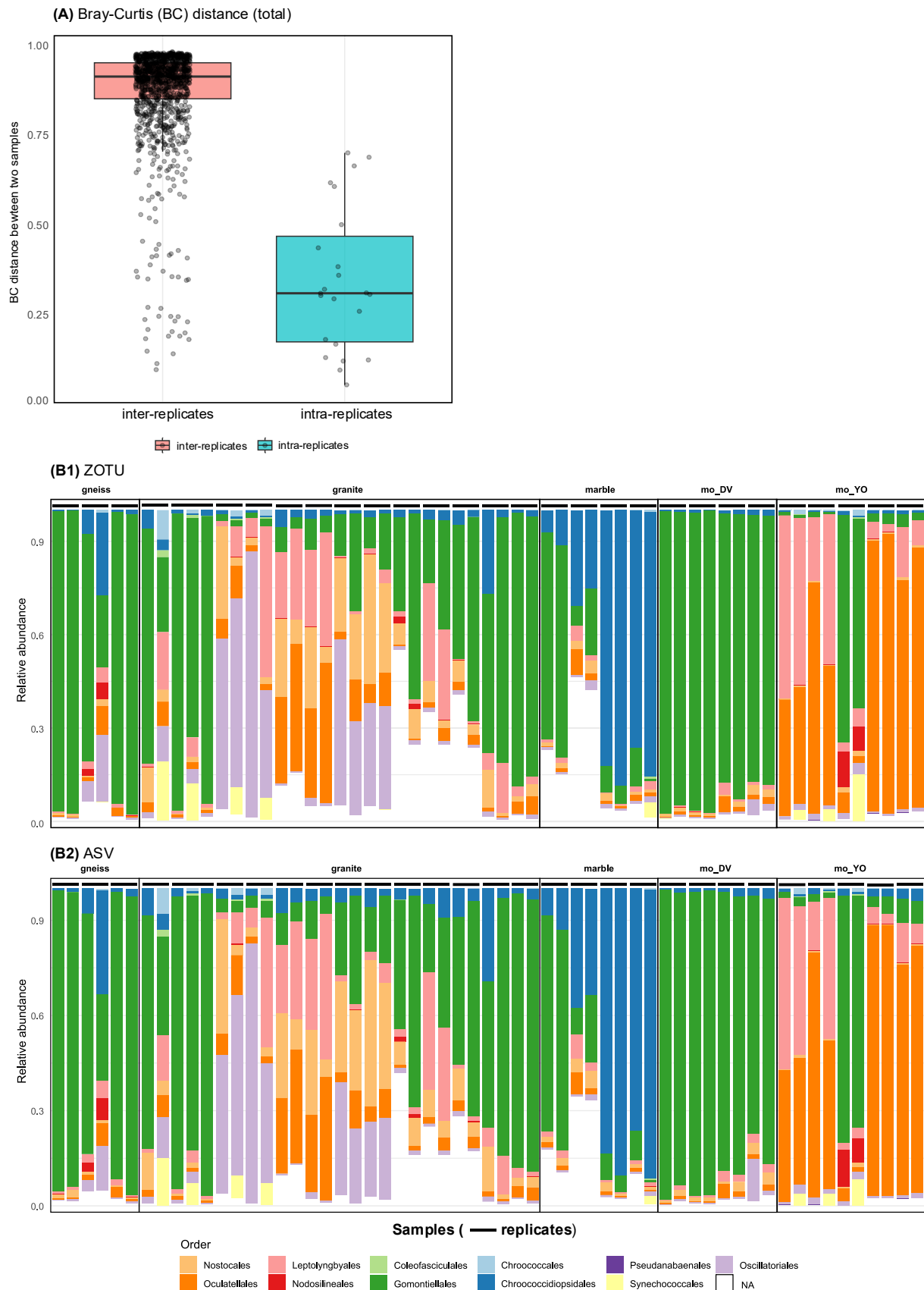


Fig. 6.4. A. Comparison of Bray–Curtis distances between and within replicate samples groups. **B.** Comparison of the relative abundances of ZOTUs and ASVs after pooling the results at the taxonomic order level, highlighting the congruence between replicate samples.

6.4. Final considerations and perspectives

6.4.1. Towards a better understanding of Antarctic uniqueness: complementary sampling and analyses

Further sampling will be required to test the hypotheses proposed in this thesis. In particular, a terrestrial sampling campaign covering the same geographic regions as the lacustrine survey presented in **Chapter 2**, or alternatively the acquisition of samples already collected by other research groups, would enable a more robust comparison of the proportion of endemic taxa between lacustrine (**Chapter 2**) and terrestrial (**Chapter 3**) cyanobacterial communities, as well as a better assessment of the influence of environmental parameters across both habitat types. The availability of RASPAM (Harmel et al., 2025), a bioinformatic pipeline developed in our laboratory, will facilitate this effort. Indeed, RASPAM provides a standardized, reproducible, and user-friendly workflow that enables the analysis of raw amplicon sequencing data without advanced bioinformatics expertise. Thus, it can be used by colleagues while ensuring the quality and comparability of the results, thereby enabling the continuation of this comparative analysis. Similarly, obtaining additional *Phormidesmis* genomes that are closely related to those analyzed in **Chapter 4** would allow the pangenomic analysis to be refined, leading to a more robust identification of Antarctic-specific genes. For example, it would be valuable to acquire and sequence the genomes of *Phormidesmis* strains described by Davydov and Vilnet (2022) or those available from the BCCM/ULC culture collection (ULC026 and ULC147).

Sampling and **metadata** collection are the main aspects that should be improved to address the limitations identified in the diversity analyses (**Chapters 2** and **3**). They are also essential for strengthening the comparative genomic analyses presented in **Chapters 4** and **5**. For the latter, additional analyses could further improve our understanding of the genomic specificities of Antarctic cyanobacteria.

The identification of **Antarctic-specific genes** could, for example, be refined by using OrthoFinder v3 (Emms et al., 2026) instead of the basic Markov Clustering algorithm (van Dongen & Abreu-Goodger, 2012) or anvi'o (Eren et al., 2015) to infer homologous genes clusters (GCs). In addition, the preliminary analysis of the origin of these Antarctic-specific genes (**Section 6.2.3**) could be improved by considering for example more sequence similarity hits rather than only the first 200, an arbitrary heuristic threshold. A more robust identification of Antarctic-specific genes and a more comprehensive characterization of their evolutionary origin, whether by recent horizontal gene transfer or specific retention from an ancestor, constitute the necessary first steps toward a deeper understanding of Antarctic genomic evolution. Also, it would be interesting

to investigate genes showing significantly higher **duplication rates**. Indeed, increased gene copy number may constitute an adaptive mechanism to specific ecological conditions by enabling a faster and/or stronger physiological response to environmental stimuli (Bratlie et al., 2010). To identify gene duplication events, gene and species tree reconciliation methods, such as implemented in AleRax (Morel et al., 2024) could be used.

Secondly, to identify **Antarctic specific functions**, the functional annotation method combining InterProScan (Jones et al., 2014) and AlphaFold (Abramson et al., 2024), as applied in **Section 6.2.3**, should be extended to all Antarctic-specific genes clusters identified in the *Phormidesmis* (37 GCs; **Chapter 4**) and picocyanobacterial (16 GCs; **Chapter 5**). AlphaFold (Abramson et al., 2024) could also be used to improve the functional characterization of proteins currently annotated as having unknown functions. Similarly, approaches based on Large Language Model (LLM), such as ProtNote (Char et al., 2025), could be used to predict the functions of unknown functions proteins. Once these Antarctic-specific genes have been better characterized, **functional enrichment** analyses, which have already been performed for the picocyanobacterial genomes (**Section 5.6.7**), should also be carried out for the *Phormidesmis* genomes. Such analyses would make it possible to determine whether the enriched functions in Antarctic *Phormidesmis* genomes are also associated with Antarctic environmental conditions and to assess the extent to which these patterns resemble those observed in picocyanobacteria.

Third, it should be interesting to determine the **selection pressure** acting on protein-coding genes. To perform this analysis, the ratio of nonsynonymous (K_a) to synonymous (K_s) substitution rates could be used (Rubio & Pérez-Pulido, 2021). A K_a/K_s ratio below 1 ($K_s > K_a$) indicates that the corresponding protein-coding gene is subject to purifying, or negative, selection, which removes deleterious amino-acid substitutions and preserves protein structure and function (Jordan et al., 2002). In contrast, a ratio above 1 ($K_s < K_a$) is generally interpreted as evidence of positive selection, suggesting that amino-acid-changing substitutions have been favored and may have contributed to the evolution of protein structures or functions. Are Antarctic-specific genes subject to stronger selection pressure than other genes? Is the same pattern observed for genes associated with functions enriched in Antarctic genomes?

Following these *in silico* analyses, **ecophysiological experiments** should be performed to validate the genomic findings. In particular, the responses of Antarctic and non-Antarctic strains to environmental stresses, such as desiccation, low temperatures, and UV radiation, can be compared through culture experiments monitoring physiological parameters, including, for example, growth rate, pigment content, and the production of exopolysaccharides and carotenoids. For desiccation and UV-stress experiments, the protocol developed in our laboratory

for a *Stenomitos* strain (Roncero-Ramos et al., 2024) could serve as a basis. Ecophysiology could be complemented by **transcriptomic analyses** (e.g., Kopf et al., 2014). Indeed, transcriptomics would make it possible to identify genes that are differentially expressed between Antarctic and non-Antarctic strains under the experimental conditions. It would also allow us to determine whether, and to what extent, the Antarctic-specific genes identified through comparative genomics (**Chapters 4 and 5**) are expressed.

Given that the Antarctic *Phormidesmis* cluster (**Fig. 6.3A**) and the two Antarctic picocyanobacterial clusters (**Fig. 5.1**) are phylogenetically distant from their closest relatives (~81% ANI), they are likely to represent distinct species. Collaboration with taxonomists would therefore be valuable to formally describe these potential new species.

Finally, it is important to keep in mind that the cyanobacteria studied in this thesis are part of **complex microbial ecosystems** (e.g., biofilms, mats, biocrusts). Interactions with other organisms within these communities have likely driven co-evolutionary processes through a variety of ecological relationships, such as symbioses with mycobionts in lichens (e.g., (Wirtz et al., 2003), associations with bryophytes (e.g., (Adams & Duggan, 2008), and antagonistic interactions with viruses or fungi (e.g., (Cairns et al., 2016). Therefore, the evolution of cyanobacteria cannot be fully understood without considering these biotic interactions. Thus, extending amplicon sequencing to other taxonomic groups using appropriate genetic markers (e.g. the 18S rRNA gene for the characterization of eukaryotes) would provide a more comprehensive view of the diversity and structure of the microbial communities studied in this thesis (**Chapters 2 and 3**). In parallel, **metagenomic** analyses would provide access to the genomic content of these microbial communities and enable it to be linked to the taxonomic diversity inferred from amplicon sequencing. Such analyses would provide valuable insights into the metabolic potential of these microbial communities (Brumfield et al., 2020). Furthermore, these metagenomic datasets could be used for **metapangenomic** and **comparative metagenomic** analyses by comparing Antarctic and non-Antarctic microbial communities. Such analyses would allow us to investigate community-level genomic specificities or functional signatures associated with Antarctic environments (Delmont & Eren, 2018; Flinkstrom et al., 2024).

6.4.2. How could genomics help strengthen the concept of endemism?

Endemism is one evidence showing that organisms exhibit spatial biogeographic patterns. By definition, an endemic taxon is restricted to a particular location and is therefore not widely distributed worldwide (Hanson et al., 2012). Organisms can be endemic to distinct geographic regions (i.e., regional endemism; e.g., Oakley et al., 2010), to broad habitat types (e.g., soils, freshwater lakes; Nemergut et al., 2011) or to specific habitats (e.g., hot springs; Takacs-Vesbach et al., 2008).

Until recently, the existence of microbial endemism remained a matter of debate, despite more than a century of research on microbial biogeography (Hedlund & Staley, 2004). Over the past two decades, however, numerous studies have provided evidence that both contemporary selection (i.e., environmental filtering) and historical processes (e.g., past selection and drift) contribute to shaping biogeographical patterns, including distance-decay and endemism. In polar cyanobacteria, several studies have reported evidence of significant endemism (e.g., Taton et al., 2006; Pessi et al., 2018), whereas others have described bipolar distributions (e.g., Jungblut et al., 2010; Kleinteich et al., 2017). Little is known about the spatial genetic structure of supposedly “cosmopolitan” microbial species or about the evolutionary processes that have shaped their distributions (Hanson et al., 2012; Segawa et al., 2023). Several studies suggest that current patterns of global diversification and regional endemism in microorganisms have been shaped by genetic isolation and local adaptation through microevolutionary processes in distinct environments (Chown et al., 2015; Thompson & Fronhofer, 2019; Segawa et al., 2023). Advances in genetics and genomics have provided new opportunities to investigate these processes, especially in organisms with cryptic morphology. Now that the existence of microbial endemism is no longer in doubt, the interesting challenge is to better characterize it.

The methodology used to assess endemism should be adapted to the taxonomic level under investigation. Although early studies argued that bacterial genera are largely cosmopolitan (e.g., Staley and Gosink, 1999), morphological characteristics may be sufficient to identify endemism when describing a new genus or species (Skoupý et al., 2022). However, below the species level, or in groups comprising cryptic taxa, morphology is often insufficient to distinguish ecotypes (Dvořák et al., 2014). When genetic approaches are used, the ability to infer whether a taxon is endemic or cosmopolitan depends on the taxonomic resolution provided by the genetic marker. The 16S rRNA gene appears too conserved to resolve biogeographic patterns below the genus level (Hedlund & Staley, 2004; Hahn et al., 2016) making the use of more variable markers necessary. For example, the internal transcribed spacer (ITS) region provides greater resolution than the 16S rRNA gene and has been shown to reveal geographic differentiation at the strain level

(Segawa et al., 2017), making it more effective for identifying endemic taxa (Fernandez-Carazo et al., 2011).

Genome-scale approaches can also provide much finer taxonomic resolution and enable the identification of lineage-specific genomic features. Indeed, comparative genomics can reveal genomic specificities that remain undetected by single-marker analyses (**Chapter 5** and **Section 6.2.2**). For example, *Cyanobium* sp. ULC054 (Antarctica) and *Cyanobium* sp. Tobar 12.5g (Spain) share 99.6% identity across their 16S rRNA gene sequences and would therefore be considered as the same species based on this marker (Yarza et al., 2014). However, their Average Nucleotide Identity (ANI) is only 82% (**Chapter 5**), well below the species boundary of 95–96% ANI (Konstantinidis & Tiedje, 2005). This discrepancy illustrates the value of genome-scale analyses for uncovering cryptic genomic divergence. Furthermore, whole-genome analyses at single-nucleotide (nt) resolution have revealed that much of the observed spatial distribution of microorganisms reflects regional endemism, with many taxa being restricted to localized geographic regions (Garcia et al., 2018; Brown & Tucker, 2020). Pangenomic analyses also enable the identification of gene clusters that are specific to a given group of genomes (see **Chapter 5** and **Section 6.2.2**).

Consequently, genomic approaches provide a valuable means of refining the **assessment of endemism**. Like the polyphasic approach used to resolve the taxonomy of certain cyanobacterial taxa, integrating multiple genetic and genomic criteria could improve the sensitivity and robustness of endemism determination. Such criteria could include genetic markers (e.g., 16S rRNA gene and ITS region), complemented by genome-scale characteristics such as Average Nucleotide Identity (ANI), group-specific genes identified through pangenomic analyses, or synteny conservation.

Going one step further, the concept of endemism could be reconsidered within a more integrative evolutionary framework, taking into account not only the focal organism but also the organisms that are in ecological interactions with it. (Cyano)bacteria exchange genetic material with neighboring microorganisms through horizontal gene transfer, potentially generating locally unique genomic characteristics that are unlikely to arise independently elsewhere (Dufresne et al., 2008; Pérez-Carrascal et al., 2021). Beyond the endemism of individual taxa, this raises a broader question: should endemism also be considered at the level of microbial ecosystems? Metagenomic and metapangenomic approaches could, for example, be used to identify groups of genes or genomic functions that are unique to a particular region or habitat. Such approaches may ultimately help determine whether entire microbial communities, rather than only individual taxa, can exhibit endemism.

6.4.3. Assessing and communicating microbial endemism: strategic tools for nature conservation

Microorganisms play pivotal roles in biogeochemical cycles and form the basis of food webs, yet they are still often overlooked in conservation biology and remain largely absent from global conservation frameworks (Cockell & Jones, 2009), even in Antarctica, where nature conservation receives considerable attention (Hughes et al., 2015). The recent establishment of the IUCN Microbial Conservation Specialist Group is a promising development. Its objectives are to (1) build a global network to assess conservation priorities, (2) map threatened microbial ecosystems worldwide, (3) develop microorganism-specific Red List criteria that account for microbial traits such as metabolic and ecological resilience, and (4) assess existing microbial conservation initiatives to facilitate their integration into IUCN conservation efforts (Gilbert et al., 2025). To achieve these objectives, such initiatives require access to data on threatened microbial habitats, as well as microbial biobanking, culture collection resources, and microbial conservation projects.

Why is the assessment of endemism important for conservation strategies? Conservation efforts often **focus** on areas that harbor **high numbers of endemic species** (Stattersfield et al., 1998; Myers et al., 2000), and patterns of endemism within well-characterized taxonomic groups may be useful for guiding conservation planning (Lamoreux et al., 2006). Accordingly, improving the sensitivity of endemism assessment in Antarctic microbial communities, for example by implementing the approaches proposed in **Section 6.4.2**, could provide stronger scientific support for conservation management decisions.

Generating scientific knowledge and evidence in support of conservation will have limited impact unless it is translated into effective conservation policies. This raises the following question: how can scientific knowledge be made accessible to, and taken seriously by, decision-makers?

Several key factors include effective communication, interdisciplinary collaboration, science outreach, and citizen science. **Effective communication** requires presenting research findings in policy-ready formats through established science–policy pathways, such as National Delegations to the Antarctic Treaty Consultative Meetings and the Scientific Committee on Antarctic Research (SCAR) (Hughes et al., 2023). Two-way dialogue between researchers and decision-makers is also essential to align research outputs with policy needs and facilitate their translation into conservation measures (Young et al., 2014). In addition, communication should extend beyond traditional scientific and policy channels to reach broader audiences through media platforms (e.g., opinion articles or editorials). Raising public awareness of conservation issues can help ensure that these challenges receive greater attention from decision-makers. This may

subsequently be complemented by policy plea, for example through environmental conservation organizations.

Interdisciplinary collaboration also represents a promising way for future research. It would provide an opportunity to broaden the interpretation of biological findings by integrating perspectives made from other scientific dimensions (Guerrero et al., 2018). For example, collaboration with sociologists or anthropologists could extend the discussion of endemism and conservation beyond a purely theoretical ecological framework by placing these concepts within a broader societal context. In addition to expanding conservation research across disciplinary boundaries, such collaborations could provide complementary knowledge and arguments to support the development of conservation policies (Bennett et al., 2017).

Science outreach enables scientific knowledge to be shared beyond the scientific community, allowing it to reach society more broadly (Varner, 2014). Despite their fundamental ecological roles, microorganisms remain an abstract concept for much of society. Furthermore, despite its ecological importance, Antarctica is often perceived as a remote and isolated wilderness, potentially distancing its environmental challenges from those experienced more directly by the wider public. It is therefore important to communicate this knowledge to children, students, and other actors of society. This can be achieved by engaging with these audiences through a variety of outreach activities, such as public science events (e.g., *Printemps des Sciences* in Belgium, <https://www.sciences.be/printemps2026/>), educational activities in schools (e.g., outreach activities organized by the Association of Polar Early Career Scientists Belgium, APECS Belgium, <https://www.apecsbelgium.com/>), or science communication through social media (e.g., the Instagram accounts [tardibabe](#) and [microbialeecology](#)).

Citizen science is another effective way to share scientific knowledge with society and raise awareness of conservation issues (Ballard et al., 2017). By involving citizens in data collection, analysis, interpretation, and the dissemination of research findings (e.g., the work of the non-profit organization *LacYme* in Belgium, <https://lacyme.org/>; the citizen science project *The quest for new basal cyanobacteria in caves* implemented by Dr. L. Cornet, <https://lcor.net.github.io/citizen/>), they become active participants in the scientific process and develop a stronger sense of ownership of the topic. As such, citizen science, alongside science outreach, represents a powerful means of communicating scientific issues to society. Ultimately, increasing public awareness of microbial conservation issues can strengthen the impact of scientific arguments on conservation decisions, as public support can be an important factor that influence decision-makers when developing and implementing conservation policies.

References

- Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., ... Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, 630(8016), 493–500. <https://doi.org/10.1038/s41586-024-07487-w>
- Adams, D. G., & Duggan, P. S. (2008). Cyanobacteria–bryophyte symbioses. *Journal of Experimental Botany*, 59(5), 1047–1058. <https://doi.org/10.1093/jxb/ern005>
- Allard, J. B., & Kumar, S. (2026). Treemblem: A graphical tool to generate Newick strings from phylogenetic tree images. *Bioinformatics*, 42(5), btag197. <https://doi.org/10.1093/bioinformatics/btag197>
- Allegretti, G., Carchini, G., Todisco, V., Convey, P., & Sbordoni, V. (2006). A molecular phylogeny of antarctic chironomidae and its implications for biogeographical history. *Polar Biology*, 29(4), 320–326. <https://doi.org/10.1007/s00300-005-0056-7>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Andrew, D. R., Fitak, R. R., Munguia-Vega, A., Racolta, A., Martinson, V. G., & Dontsova, K. (2012). Abiotic Factors Shape Microbial Diversity in Sonoran Desert Soils. *Applied and Environmental Microbiology*, 78(21), 7527–7537. <https://doi.org/10.1128/AEM.01459-12>
- Arvola, L., Leppäranta, & Zhijun, L. (2024). Microscopic analysis on eukaryotic algae and cyanobacteria in nine seasonal lakes and ponds in Vestfjella, Dronning Maud Land, Antarctica. *Advances in Polar Science*, 35(2), 206–218. <https://doi.org/10.12429/j.advps.2023.0029>
- Ballard, H. L., Dixon, C. G. H., & Harris, E. M. (2017). Youth-focused citizen science: Examining the role of environmental science learning and agency for conservation. *Biological Conservation, The Role of Citizen Science in Biological Conservation*, 208, 65–75. <https://doi.org/10.1016/j.biocon.2016.05.024>
- Barnes, C. J., Rasmussen, L., Asplund, M., Knudsen, S. W., Clausen, M.-L., Agner, T., & Hansen, A. J. (2020). Comparing DADA2 and OTU clustering approaches in studying the bacterial communities of atopic dermatitis. *Journal of Medical Microbiology*, 69(11), 1293–1302. <https://doi.org/10.1099/jmm.0.001256>
- Barria, C., Malecki, M., & Arraiano, C. M. (2013). Bacterial adaptation to cold. *Microbiology*, 159(Pt_12), 2437–2443. <https://doi.org/10.1099/mic.0.052209-0>
- Bell, E. M., & Laybourn-Parry, J. (1999). Annual plankton dynamics in an Antarctic saline lake. *Freshwater Biology*, 41(3), 507–519. <https://doi.org/10.1046/j.1365-2427.1999.00396.x>

- Bennett, N. J., Roth, R., Klain, S. C., Chan, K., Christie, P., Clark, D. A., Cullman, G., Curran, D., Durbin, T. J., Epstein, G., Greenberg, A., Nelson, M. P., Sandlos, J., Stedman, R., Teel, T. L., Thomas, R., Veríssimo, D., & Wyborn, C. (2017). Conservation social science: Understanding and integrating human dimensions to improve conservation. *Biological Conservation*, 205, 93–108. <https://doi.org/10.1016/j.biocon.2016.10.006>
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Keira Cheetham, R., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., ... Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*, 456(7218), 53–59. <https://doi.org/10.1038/nature07517>
- Beyer, L., Pingpank, K., Wriedt, G., & Bölter, M. (2000). Soil formation in coastal continental Antarctica (Wilkes Land). *Geoderma*, 95(3), 283–304. [https://doi.org/10.1016/S0016-7061\(99\)00095-6](https://doi.org/10.1016/S0016-7061(99)00095-6)
- Bickel, S., & Or, D. (2021). The chosen few—Variations in common and rare soil bacteria across biomes. *The ISME Journal*, 15(11), 3315–3325. <https://doi.org/10.1038/s41396-021-00981-3>
- Biersma, E. M., Jackson, J. A., Stech, M., Griffiths, H., Linse, K., & Convey, P. (2018). Molecular Data Suggest Long-Term in Situ Antarctic Persistence Within Antarctica's Most Speciose Plant Genus, *Schistidium*. *Frontiers in Ecology and Evolution*, 6. <https://doi.org/10.3389/fevo.2018.00077>
- Blankenship, R. E. (2002). *Molecular Mechanisms of Photosynthesis* (1st ed.). Wiley. <https://doi.org/10.1002/9780470758472>
- Block, W., Lewis Smith, R. I., & Kennedy, A. D. (2009). Strategies of survival and resource exploitation in the Antarctic fellfield ecosystem. *Biological Reviews*, 84(3), 449–484. <https://doi.org/10.1111/j.1469-185X.2009.00084.x>
- Bottos, E. M., Laughlin, D. C., Herbold, C. W., Lee, C. K., McDonald, I. R., & Cary, S. C. (2020). Abiotic factors influence patterns of bacterial diversity and community composition in the Dry Valleys of Antarctica. *FEMS Microbiology Ecology*, 96(5), fiae042. <https://doi.org/10.1093/femsec/fiae042>
- Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., Schulz, F., Jarett, J., Rivers, A. R., Eloie-Fadrosch, E. A., Tringe, S. G., Ivanova, N. N., Copeland, A., Clum, A., Becraft, E. D., Malmstrom, R. R., Birren, B., Podar, M., Bork, P., ... Woyke, T. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nature Biotechnology*, 35(8), 725–731. <https://doi.org/10.1038/nbt.3893>
- Bratlie, M. S., Johansen, J., Sherman, B. T., Huang, D. W., Lempicki, R. A., & Drabløs, F. (2010). Gene duplications in prokaryotes can be associated with environmental adaptation. *BMC Genomics*, 11(1), 588. <https://doi.org/10.1186/1471-2164-11-588>

- Brooks, S. T., Jabour, J., van den Hoff, J., & Bergstrom, D. M. (2019). Our footprint on Antarctica competes with nature for rare ice-free land. *Nature Sustainability*, 2(3), 185–190. <https://doi.org/10.1038/s41893-019-0237-y>
- Brooks, S. T., Tejedo, P., & O'Neill, T. A. (2019). Insights on the environmental impacts associated with visible disturbance of ice-free ground in Antarctica. *Antarctic Science*, 31(6), 304–314. <https://doi.org/10.1017/S0954102019000440>
- Brown, S. P., & Tucker, A. E. (2020). Distribution and biogeography of *Sanguina* snow algae: Fine-scale sequence analyses reveal previously unknown population structure. *Ecology and Evolution*, 10(20), 11352–11361. <https://doi.org/10.1002/ece3.6772>
- Brumfield, K. D., Huq, A., Colwell, R. R., Olds, J. L., & Leddy, M. B. (2020). Microbial resolution of whole genome shotgun and 16S amplicon metagenomic sequencing using publicly available NEON data. *PLOS ONE*, 15(2), e0228899. <https://doi.org/10.1371/journal.pone.0228899>
- Burton-Johnson, A., Black, M., Fretwell, P. T., & Kaluza-Gilbert, J. (2016). An automated methodology for differentiating rock from snow, clouds and sea in Antarctica from Landsat 8 imagery: A new rock outcrop map and area estimation for the entire Antarctic continent. *The Cryosphere*, 10(4), 1665–1677. <https://doi.org/10.5194/tc-10-1665-2016>
- Butterfield, N. J. (2007). Macroevolution And Macroecology Through Deep Time. *Palaeontology*, 50(1), 41–55. <https://doi.org/10.1111/j.1475-4983.2006.00613.x>
- Cabello-Yeves, P. J., Callieri, C., Picazo, A., Schallenberg, L., Huber, P., Roda-Garcia, J. J., Bartosiewicz, M., Belykh, O. I., Tikhonova, I. V., Torcello-Requena, A., De Prado, P. M., Puxty, R. J., Millard, A. D., Camacho, A., Rodriguez-Valera, F., & Scanlan, D. J. (2022). Elucidating the picocyanobacteria salinity divide through ecogenomics of new freshwater isolates. *BMC Biology*, 20(1), 175. <https://doi.org/10.1186/s12915-022-01379-z>
- Cairns, J., Coloma, S., Sivonen, K., & Hiltunen, T. (2016). Evolving interactions between diazotrophic cyanobacterium and phage mediate nitrogen release and host competitive ability. *Royal Society Open Science*, 3(12), 160839. <https://doi.org/10.1098/rsos.160839>
- 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>
- Callieri, C., Cronberg, G., & Stockner, J. G. (2012). Freshwater Picocyanobacteria: Single Cells, Microcolonies and Colonial Forms. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (pp. 229–269). Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_8
- Cameron, K. A., Hodson, A. J., & Osborn, A. M. (2012). Structure and diversity of bacterial, eukaryotic and archaeal communities in glacial cryoconite holes from the Arctic and the Antarctic. *FEMS Microbiology Ecology*, 82(2), 254–267. <https://doi.org/10.1111/j.1574-6941.2011.01277.x>

- Cao, S., Zhang, W., Ding, W., Wang, M., Fan, S., Yang, B., McMinn, A., Wang, M., Xie, B., Qin, Q.-L., Chen, X.-L., He, J., & Zhang, Y.-Z. (2020). Structure and function of the Arctic and Antarctic marine microbiota as revealed by metagenomics. *Microbiome*, 8(1), 47. <https://doi.org/10.1186/s40168-020-00826-9>
- Castenholz, R. W., & Garcia-Pichel, F. (2012). Cyanobacterial Responses to UV Radiation. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (pp. 481–499). Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_19
- Ceres, K. M., Stanhope, M. J., & Gröhn, Y. T. (2022). A critical evaluation of *Mycobacterium bovis* pangenomics, with reference to its utility in outbreak investigation. *Microbial Genomics*, 8(6), 000839. <https://doi.org/10.1099/mgen.0.000839>
- Char, S., Corley, N., Alamdari, S., Yang, K. K., & Amini, A. P. (2025). ProtNote: A multimodal method for protein–function annotation. *Bioinformatics*, 41(5), btaf170. <https://doi.org/10.1093/bioinformatics/btaf170>
- Chase, A. B., & Martiny, J. B. (2018). The importance of resolving biogeographic patterns of microbial microdiversity. *Microbiology Australia*, 39(1), 5–8. <https://doi.org/10.1071/MA18003>
- Chen, M.-Y., Teng, W.-K., Zhao, L., Hu, C.-X., Zhou, Y.-K., Han, B.-P., Song, L.-R., & Shu, W.-S. (2021). Comparative genomics reveals insights into cyanobacterial evolution and habitat adaptation. *The ISME Journal*, 15(1), 211–227. <https://doi.org/10.1038/s41396-020-00775-z>
- Cho, J.-C., & Tiedje, J. M. (2000). Biogeography and Degree of Endemicity of Fluorescent *Pseudomonas* Strains in Soil. *Applied and Environmental Microbiology*, 66(12), 5448–5456. <https://doi.org/10.1128/AEM.66.12.5448-5456.2000>
- Chong, C. W., Goh, Y. S., Convey, P., Pearce, D., & Tan, I. K. P. (2013). Spatial pattern in Antarctica: What can we learn from Antarctic bacterial isolates? *Extremophiles*, 17(5), 733–745. <https://doi.org/10.1007/s00792-013-0555-3>
- Chong, C.-W., Pearce, D. A., & Convey, P. (2015). Emerging spatial patterns in Antarctic prokaryotes. *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.01058>
- Chow, C.-E. T., & Suttle, C. A. (2015). Biogeography of Viruses in the Sea. *Annual Review of Virology*, 2(Volume 2, 2015), 41–66. <https://doi.org/10.1146/annurev-virology-031413-085540>
- Chown, S. L., Clarke, A., Fraser, C. I., Cary, S. C., Moon, K. L., & McGeoch, M. A. (2015). The changing form of Antarctic biodiversity. *Nature*, 522(7557), 431–438. <https://doi.org/10.1038/nature14505>
- Christmas, N. A. M., Anesio, A. M., & Sánchez-Baracaldo, P. (2015). Multiple adaptations to polar and alpine environments within cyanobacteria: A phylogenomic and Bayesian approach. *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.01070>

- Christmas, N. A. M., Anesio, A. M., & Sánchez-Baracaldo, P. (2018). The future of genomics in polar and alpine cyanobacteria. *FEMS Microbiology Ecology*, 94(4). <https://doi.org/10.1093/femsec/fiy032>
- Christmas, N. A. M., Barker, G., Anesio, A. M., & Sánchez-Baracaldo, P. (2016). Genomic mechanisms for cold tolerance and production of exopolysaccharides in the Arctic cyanobacterium *Phormidesmis priestleyi* BC1401. *BMC Genomics*, 17(1), 533. <https://doi.org/10.1186/s12864-016-2846-4>
- Christodoulou, M., Wahlsten, M., & Sivonen, K. (2023). Morphological and Molecular Evaluation of *Pseudanabaena epilithica* sp. Nov. And *P. suomiensis* sp. Nov. (Pseudanabaenaceae, Cyanobacteria) from Finland. *Diversity*, 15(8), 909. <https://doi.org/10.3390/d15080909>
- Cockell, C. S., & Jones, H. L. (2009). Advancing the case for microbial conservation. *Oryx*, 43(4), 520–526. <https://doi.org/10.1017/S0030605309990111>
- Cockell, C. S., & Stokes, M. D. (2004). Widespread colonization by polar hypoliths. *Nature*, 431(7007), 414–414. <https://doi.org/10.1038/431414a>
- Contreras, M. J., Leal, K., Bruna, P., Nuñez-Montero, K., Gómez-Espinoza, O., Santos, A., Bravo, L., Valenzuela, B., Solis, F., Gahona, G., Cayo, M., Dinamarca, M. A., Ibacache-Quiroga, C., Zamorano, P., & Barrientos, L. (2023). Commonalities between the Atacama Desert and Antarctica rhizosphere microbial communities. *Frontiers in Microbiology*, 14. <https://doi.org/10.3389/fmicb.2023.1197399>
- Convey, P., Biersma, E. M., Casanova-Katny, A., & Maturana, C. S. (2020). Refuges of Antarctic diversity. In *Past Antarctica* (pp. 181–200). Elsevier. <https://doi.org/10.1016/B978-0-12-817925-3.00010-0>
- Convey, P., Chown, S. L., Clarke, A., Barnes, D. K. A., Bokhorst, S., Cummings, V., Ducklow, H. W., Frati, F., Green, T. G. A., Gordon, S., Griffiths, H. J., Howard-Williams, C., Huiskes, A. H. L., Laybourn-Parry, J., Lyons, W. B., McMinn, A., Morley, S. A., Peck, L. S., Quesada, A., ... Wall, D. H. (2014). The spatial structure of Antarctic biodiversity. *Ecological Monographs*, 84(2), 203–244. <https://doi.org/10.1890/12-2216.1>
- Cordero, O. X., & Polz, M. F. (2014). Explaining microbial genomic diversity in light of evolutionary ecology. *Nature Reviews Microbiology*, 12(4), 263–273. <https://doi.org/10.1038/nrmicro3218>
- Cornet, L., & Baurain, D. (2022). Contamination detection in genomic data: More is not enough. *Genome Biology*, 23(1), 60. <https://doi.org/10.1186/s13059-022-02619-9>
- Cornet, L., Bertrand, A. R., Hanikenne, M., Javaux, E. J., Wilmotte, A., & Baurain, D. (2018). Metagenomic assembly of new (sub)polar Cyanobacteria and their associated microbiome from non-axenic cultures. *Microbial Genomics*, 4(9). <https://doi.org/10.1099/mgen.0.000212>

Cottingham, K. L., Ewing, H. A., Greer, M. L., Carey, C. C., & Weathers, K. C. (2015). Cyanobacteria as biological drivers of lake nitrogen and phosphorus cycling. *Ecosphere*, 6(1), art1. <https://doi.org/10.1890/ES14-00174.1>

Crump, B. C., Amaral-Zettler, L. A., & Kling, G. W. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *The ISME Journal*, 6(9), 1629–1639. <https://doi.org/10.1038/ismej.2012.9>

Cruz-Martínez, K., Rosling, A., Zhang, Y., Song, M., Andersen, G. L., & Banfield, J. F. (2012). Effect of Rainfall-Induced Soil Geochemistry Dynamics on Grassland Soil Microbial Communities. *Applied and Environmental Microbiology*, 78(21), 7587–7595. <https://doi.org/10.1128/AEM.00203-12>

D’Amico, S., Collins, T., Marx, J., Feller, G., Gerday, C., & Gerday, C. (2006). Psychrophilic microorganisms: Challenges for life. *The EMBO Reports*, 7(4), 385–389. <https://doi.org/10.1038/sj.embor.7400662>

Davydov, D., & Vilnet, A. (2022). Review of the Cyanobacterial Genus *Phormidesmis* (Leptolyngbyaceae) with the Description of *Apatinema* gen. Nov. *Diversity*, 14(9), 731. <https://doi.org/10.3390/d14090731>

De Wever, A., Leliaert, F., Verleyen, E., Vanormelingen, P., Van der Gucht, K., Hodgson, D. A., Sabbe, K., & Vyverman, W. (2009). Hidden levels of phylodiversity in Antarctic green algae: Further evidence for the existence of glacial refugia. *Proceedings of the Royal Society B: Biological Sciences*, 276(1673), 3591–3599. <https://doi.org/10.1098/rspb.2009.0994>

Delmont, T. O., & Eren, A. M. (2018). Linking pangenomes and metagenomes: The *Prochlorococcus* metapangenome. *PeerJ*, 6, e4320. <https://doi.org/10.7717/peerj.4320>

Demoulin, C. F., Lara, Y. J., Cornet, L., François, C., Baurain, D., Wilmotte, A., & Javaux, E. J. (2019). Cyanobacteria evolution: Insight from the fossil record. *Free Radical Biology and Medicine, Early Life on Earth and Oxidative Stress*, 140, 206–223. <https://doi.org/10.1016/j.freeradbiomed.2019.05.007>

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P., & Andersen, G. L. (2006). Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Applied and Environmental Microbiology*, 72(7), 5069–5072. <https://doi.org/10.1128/AEM.03006-05>

Dewar, A. E., Belcher, L. J., & West, S. A. (2025). A phylogenetic approach to comparative genomics. *Nature Reviews Genetics*, 26(6), 395–405. <https://doi.org/10.1038/s41576-024-00803-0>

Dextro, R. B., Delbaje, E., Freitas, P. N. N., Geraldès, V., Pinto, E., Long, P. F., & Fiore, M. F. (2023). Environmental adaptations by the intertidal Antarctic cyanobacterium *Halotia branconii* CENA392 as revealed using long-read genome sequencing. *Limnology and Oceanography Letters*, 8(5), 770–780. <https://doi.org/10.1002/lol2.10337>

- Domingo-Sananes, M. R., & McInerney, J. O. (2021). Mechanisms That Shape Microbial Pangenomes. *Trends in Microbiology*, 29(6), 493–503. <https://doi.org/10.1016/j.tim.2020.12.004>
- Doré, H., Farrant, G. K., Guyet, U., Haguit, J., Humily, F., Ratin, M., Pitt, F. D., Ostrowski, M., Six, C., Brillet-Guéguen, L., Hoebeke, M., Bisch, A., Le Corguillé, G., Corre, E., Labadie, K., Aury, J.-M., Wincker, P., Choi, D. H., Noh, J. H., ... Garczarek, L. (2020). Evolutionary Mechanisms of Long-Term Genome Diversification Associated With Niche Partitioning in Marine Picocyanobacteria. *Frontiers in Microbiology*, 11, 567431. <https://doi.org/10.3389/fmicb.2020.567431>
- Dory, F., Nava, V., Nedbalová, L., Soler, V., Mauro, B. D., Traversa, G., Spreafico, M., & Leoni, B. (2025). Morphological diversity of microalgae and Cyanobacteria of cryoconite holes in Northern Victoria Land, Antarctica. *Journal of Glaciology*, 71, e37. <https://doi.org/10.1017/jog.2025.12>
- Dragone, N. B., Childress, M. K., Mendez, N., Galletta, J., Vanderburgh, C., Mesquita, C. P. B. de, DeAngelis, K. M., Quandt, C. A., Leung, P. M., Greening, C., Adams, B. J., & Fierer, N. (2026). Evidence for endemism and local adaptation in Antarctic soil bacteria (p. 2026.05.01.722257). *bioRxiv*. <https://doi.org/10.64898/2026.05.01.722257>
- Dsouza, M., Taylor, M. W., Turner, S. J., & Aislabie, J. (2015). Genomic and phenotypic insights into the ecology of *Arthrobacter* from Antarctic soils. *BMC Genomics*, 16(1), 36. <https://doi.org/10.1186/s12864-015-1220-2>
- Dudeja, S., Bhattacharjee, A. B., & Chela-Flores, J. (2010). Microbial Mats in Antarctica as Models for the Search of Life on the Jovian Moon Europa. In J. Seckbach & A. Oren (Eds.), *Microbial Mats: Modern and Ancient Microorganisms in Stratified Systems* (pp. 543–561). Springer Netherlands. https://doi.org/10.1007/978-90-481-3799-2_28
- Dufresne, A., Ostrowski, M., Scanlan, D. J., Garczarek, L., Mazard, S., Palenik, B. P., Paulsen, I. T., De Marsac, N. T., Wincker, P., Dossat, C., Ferriera, S., Johnson, J., Post, A. F., Hess, W. R., & Partensky, F. (2008). Unraveling the genomic mosaic of a ubiquitous genus of marine cyanobacteria. *Genome Biology*, 9(5), R90. <https://doi.org/10.1186/gb-2008-9-5-r90>
- Durieu, B., Savaglia, V., Vlierberghe, M. V., Lupo, V., Baurain, D., Wilmotte, A., & Cornet, L. (2025). Out of Antarctica: New insights into Antarctic Subcluster 5.2 picocyanobacteria based on high-quality genomes (p. 2025.12.17.694815). *bioRxiv*. <https://doi.org/10.64898/2025.12.17.694815>
- Dvořák, P., Casamatta, D. A., Pouličková, A., Hašler, P., Ondřej, V., & Sanges, R. (2014). *Synechococcus*: 3 billion years of global dominance. *Molecular Ecology*, 23(22), 5538–5551. <https://doi.org/10.1111/mec.12948>
- Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10(10), 996–998. <https://doi.org/10.1038/nmeth.2604>
- Edgar, R. C. (2016). UNOISE2: Improved error-correction for Illumina 16S and ITS amplicon sequencing (p. 081257). *bioRxiv*. <https://doi.org/10.1101/081257>

- Effendi, D. B., Sakamoto, T., Ohtani, S., Awai, K., & Kanesaki, Y. (2022). Possible involvement of extracellular polymeric substrates of Antarctic cyanobacterium *Nostoc* sp. Strain SO-36 in adaptation to harsh environments. *Journal of Plant Research*, 135(6), 771–784. <https://doi.org/10.1007/s10265-022-01411-x>
- Ellis-Evans, J. C. (1996). Microbial diversity and function in Antarctic freshwater ecosystems. *Biodiversity & Conservation*, 5(11), 1395–1431. <https://doi.org/10.1007/BF00051985>
- Elster, J. (2002). Ecological Classification of Terrestrial Algal Communities in Polar Environments. In L. Beyer & M. Bölker (Eds.), *Geoecology of Antarctic Ice-Free Coastal Landscapes* (pp. 303–326). Springer. https://doi.org/10.1007/978-3-642-56318-8_17
- Emms, D. M., Liu, Y., Belcher, L., Holmes, J., & Kelly, S. (2026). OrthoFinder: Improved phylogenetic orthology inference with enhanced accuracy and scalability. *Nature Methods*, 23(7), 1327–1333. <https://doi.org/10.1038/s41592-026-03126-6>
- Environmental Protocol. (1991). *Protocol on environmental protection to the Antarctic treaty*. Antarctic Treaty Secretariat.
- Eren, A. M., Esen, Ö. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., & Delmont, T. O. (2015). Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ*, 3, e1319. <https://doi.org/10.7717/peerj.1319>
- Fares, M., Tharwat, E. K., Cleenwerck, I., Monsieurs, P., Van Houdt, R., Vandamme, P., El-Hadidi, M., & Mysara, M. (2025). The unresolved struggle of 16S rRNA amplicon sequencing: A benchmarking analysis of clustering and denoising methods. *Environmental Microbiome*, 20(1), 51. <https://doi.org/10.1186/s40793-025-00705-6>
- Fasolo, A., Deb, S., Stevanato, P., Concheri, G., & Squartini, A. (2024). ASV vs OTUs clustering: Effects on alpha, beta, and gamma diversities in microbiome metabarcoding studies. *PLOS ONE*, 19(10), e0309065. <https://doi.org/10.1371/journal.pone.0309065>
- Fernandez-Carazo, R., Hodgson, D. A., Convey, P., & Wilmotte, A. (2011). Low cyanobacterial diversity in biotopes of the Transantarctic Mountains and Shackleton Range (80–82°S), Antarctica. *FEMS Microbiology Ecology*, 77(3), 503–517. <https://doi.org/10.1111/j.1574-6941.2011.01132.x>
- Ferreira, A., Mendes, C. R. B., Costa, R. R., Brotas, V., Tavano, V. M., Guerreiro, C. V., Secchi, E. R., & Brito, A. C. (2024). Climate change is associated with higher phytoplankton biomass and longer blooms in the West Antarctic Peninsula. *Nature Communications*, 15(1), 6536. <https://doi.org/10.1038/s41467-024-50381-2>
- Finn, R. D., Clements, J., & Eddy, S. R. (2011). HMMER web server: Interactive sequence similarity searching. *Nucleic Acids Research*, 39(suppl_2), W29–W37. <https://doi.org/10.1093/nar/gkr367>
- Fiore, M. F., Moon, D. H., Tsai, S. M., Lee, H., & Trevors, J. T. (2000). Miniprep DNA isolation from unicellular and filamentous cyanobacteria. *Journal of Microbiological Methods*, 39(2), 159–169. [https://doi.org/10.1016/S0167-7012\(99\)00110-4](https://doi.org/10.1016/S0167-7012(99)00110-4)

- Flinkstrom, Z., Bryson, S., Candry, P., & Winkler, M.-K. H. (2024). Metagenomic clustering links specific metabolic functions to globally relevant ecosystems. *mSystems*, 9(8), e00573-24. <https://doi.org/10.1128/msystems.00573-24>
- Flombaum, P., Gallegos, J. L., Gordillo, R. A., Rincón, J., Zabala, L. L., Jiao, N., Karl, D. M., Li, W. K. W., Lomas, M. W., Veneziano, D., Vera, C. S., Vrugt, J. A., & Martiny, A. C. (2013). Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. *Proceedings of the National Academy of Sciences*, 110(24), 9824–9829. <https://doi.org/10.1073/pnas.1307701110>
- Fraser, C. I., Terauds, A., Smellie, J., Convey, P., & Chown, S. L. (2014). Geothermal activity helps life survive glacial cycles. *Proceedings of the National Academy of Sciences*, 111(15), 5634–5639. <https://doi.org/10.1073/pnas.1321437111>
- Fritsen, C. H., & Priscu, J. C. (1998). CYANOBACTERIAL ASSEMBLAGES IN PERMANENT ICE COVERS ON ANTARCTIC LAKES: DISTRIBUTION, GROWTH RATE, AND TEMPERATURE RESPONSE OF PHOTOSYNTHESIS. *Journal of Phycology*, 34(4), 587–597. <https://doi.org/10.1046/j.1529-8817.1998.340587.x>
- Garcia, S. L., Stevens, S. L. R., Crary, B., Martinez-Garcia, M., Stepanauskas, R., Woyke, T., Tringe, S. G., Andersson, S. G. E., Bertilsson, S., Malmstrom, R. R., & McMahon, K. D. (2018). Contrasting patterns of genome-level diversity across distinct co-occurring bacterial populations. *The ISME Journal*, 12(3), 742–755. <https://doi.org/10.1038/s41396-017-0001-0>
- Garczarek, L., Guyet, U., Doré, H., Farrant, G. K., Hoebeke, M., Brillet-Guéguen, L., Bisch, A., Ferrieux, M., Siltanen, J., Corre, E., Le Corguillé, G., Ratin, M., Pitt, F. D., Ostrowski, M., Conan, M., Siegel, A., Labadie, K., Aury, J.-M., Wincker, P., ... Partensky, F. (2021). Cyanorak v2.1: A scalable information system dedicated to the visualization and expert curation of marine and brackish picocyanobacteria genomes. *Nucleic Acids Research*, 49(D1), D667–D676. <https://doi.org/10.1093/nar/gkaa958>
- Gilbert, J. A., Peixoto, R. S., Scholz, A. H., Dominguez Bello, M. G., Korsten, L., Berg, G., Singh, B., Boetius, A., Wang, F., Greening, C., Wrighton, K., Bordenstein, S., Jansson, J. K., Lennon, J. T., Souza, V., Thomas, T., Cowan, D., Crowther, T. W., Nguyen, N., ... Redford, K. (2025). Launching the IUCN Microbial Conservation Specialist Group as a global safeguard for microbial biodiversity. *Nature Microbiology*, 10(10), 2359–2360. <https://doi.org/10.1038/s41564-025-02113-5>
- Golubic, S., & Schneider, J. (2003). Microbial Endoliths as Internal Biofilms. In W. E. Krumbein, D. M. Paterson, & G. A. Zavarzin (Eds.), *Fossil and Recent Biofilms: A Natural History of Life on Earth* (pp. 249–263). Springer Netherlands. https://doi.org/10.1007/978-94-017-0193-8_16
- Greco, C., Andersen, D. T., Hawes, I., Bowles, A. M. C., Yallop, M. L., Barker, G., & Jungblut, A. D. (2020). Microbial Diversity of Pinnacle and Conical Microbial Mats in the Perennially Ice-Covered Lake Untersee, East Antarctica. *Frontiers in Microbiology*, 11. <https://doi.org/10.3389/fmicb.2020.607251>

- Green, J. L., Holmes, A. J., Westoby, M., Oliver, I., Briscoe, D., Dangerfield, M., Gillings, M., & Beattie, A. J. (2004). Spatial scaling of microbial eukaryote diversity. *Nature*, 432(7018), 747–750. <https://doi.org/10.1038/nature03034>
- Grettenberger, C. L., Sumner, D. Y., Wall, K., Brown, C. T., Eisen, J. A., Mackey, T. J., Hawes, I., Jospin, G., & Jungblut, A. D. (2020). A phylogenetically novel cyanobacterium most closely related to *Gloeobacter*. *The ISME Journal*, 14(8), 2142–2152. <https://doi.org/10.1038/s41396-020-0668-5>
- Griffiths, R. I., Thomson, B. C., James, P., Bell, T., Bailey, M., & Whiteley, A. S. (2011). The bacterial biogeography of British soils. *Environmental Microbiology*, 13(6), 1642–1654. <https://doi.org/10.1111/j.1462-2920.2011.02480.x>
- Guerra, A. (2026). The pangenome: A statistical model, not a fixed biological property. *Bioinformatics Advances*, 6(1), vbag069. <https://doi.org/10.1093/bioadv/vbag069>
- Guerra, C. A., Berdugo, M., Eldridge, D. J., Eisenhauer, N., Singh, B. K., Cui, H., Abades, S., Alfaro, F. D., Bamigboye, A. R., Bastida, F., Blanco-Pastor, J. L., de los Ríos, A., Durán, J., Grebenc, T., Illán, J. G., Liu, Y.-R., Makhalanyane, T. P., Mamet, S., Molina-Montenegro, M. A., ... Delgado-Baquerizo, M. (2022). Global hotspots for soil nature conservation. *Nature*, 610(7933), 693–698. <https://doi.org/10.1038/s41586-022-05292-x>
- Guerrero, A., Bennett, N., Wilson, K., Carter, N., Gill, D., Mills, M., Ives, C., Selinske, M., Larrosa, C., Bekessy, S., Januchowski-Hartley, F., Travers, H., Wyborn, C., & Nuno, A. (2018). Achieving the promise of integration in social-ecological research: A review and prospectus. *Ecology and Society*, 23(3). <https://doi.org/10.5751/ES-10232-230338>
- Haas, B. J., Gevers, D., Earl, A. M., Feldgarden, M., Ward, D. V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S. K., Sodergren, E., Methé, B., DeSantis, T. Z., Consortium, T. H. M., Petrosino, J. F., Knight, R., & Birren, B. W. (2011). Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. *Genome Research*, 21(3), 494–504. <https://doi.org/10.1101/gr.112730.110>
- Hahn, M. W., Jezberová, J., Koll, U., Saueressig-Beck, T., & Schmidt, J. (2016). Complete ecological isolation and cryptic diversity in Polynucleobacter bacteria not resolved by 16S rRNA gene sequences. *The ISME Journal*, 10(7), 1642–1655. <https://doi.org/10.1038/ismej.2015.237>
- Hambücken, L., Baurain, D., & Cornet, L. (2026). Exploring thylakoid emergence: Evolution of membrane biogenesis and photosystem II assembly in early-diverging cyanobacteria. *New Phytologist*, 251(3), 1329–1343. <https://doi.org/10.1111/nph.71284>
- Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., & Martiny, J. B. H. (2012). Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nature Reviews Microbiology*, 10(7), 497–506. <https://doi.org/10.1038/nrmicro2795>
- Harmel, M., Durieu, B., Savaglia, V., Tytgat, B., Baurain, D., Wilmotte, A., Verleyen, E., & Cornet, L. (2025). A FAIR Amplicon Sequencing Workflow for Long-term Environmental Monitoring (p. 2025.12.04.692289). *bioRxiv*. <https://doi.org/10.64898/2025.12.04.692289>

- Hawes, I., & Schwarz, A.-M. J. (2001). Absorption and Utilization of Irradiance by Cyanobacterial Mats in Two Ice-Covered Antarctic Lakes with Contrasting Light Climates. *Journal of Phycology*, 37(1), 5–15. <https://doi.org/10.1046/j.1529-8817.1999.014012005.x>
- Hedlund, B. P., Chuvoshina, M., Hugenholtz, P., Konstantinidis, K. T., Murray, A. E., Palmer, M., Parks, D. H., Probst, A. J., Reysenbach, A.-L., Rodriguez-R, L. M., Rossello-Mora, R., Sutcliffe, I. C., Venter, S. N., & Whitman, W. B. (2022). SeqCode: A nomenclatural code for prokaryotes described from sequence data. *Nature Microbiology*, 7(10), 1702–1708. <https://doi.org/10.1038/s41564-022-01214-9>
- Hedlund, B. P., & Staley, J. T. (2004). Microbial Endemism and Biogeography. In A. T. Bull (Ed.), *Microbial Diversity and Bioprospecting* (pp. 225–231). ASM Press. <https://doi.org/10.1128/9781555817770.ch22>
- Henshaw, T., & Laybourn-Parry, J. (2002). The annual patterns of photosynthesis in two large, freshwater, ultra-oligotrophic Antarctic lakes. *Polar Biology*, 25(10), 744–752. <https://doi.org/10.1007/s00300-002-0402-y>
- Herrero, A., Stavans, J., & Flores, E. (2016). The multicellular nature of filamentous heterocyst-forming cyanobacteria. *FEMS Microbiology Reviews*, 40(6), 831–854. <https://doi.org/10.1093/femsre/fuw029>
- Hodgson, D. A. (2012). Antarctic Lakes. In L. Bengtsson, R. W. Herschy, & R. W. Fairbridge (Eds.), *Encyclopedia of Lakes and Reservoirs* (pp. 26–31). Springer Netherlands. https://doi.org/10.1007/978-1-4020-4410-6_38
- Hofmann, H. J. (1976). Precambrian Microflora, Belcher Islands, Canada: Significance and Systematics. *Journal of Paleontology*, 50(6), 1040–1073.
- Hughes, K. A., Cowan, D. A., & Wilmotte, A. (2015). Protection of Antarctic microbial communities – ‘out of sight, out of mind.’ *Frontiers in Microbiology*, 6. <https://doi.org/10.3389/fmicb.2015.00151>
- Hughes, K. A., Ireland, L. C., Convey, P., & Fleming, A. H. (2016). Assessing the effectiveness of specially protected areas for conservation of Antarctica’s botanical diversity. *Conservation Biology*, 30(1), 113–120. <https://doi.org/10.1111/cobi.12592>
- Hughes, K. A., Lowther, A., Gilbert, N., Waluda, C. M., & Lee, J. R. (2023). Communicating the best available science to inform Antarctic policy and management: A practical introduction for researchers. *Antarctic Science*, 35(6), 438–472. <https://doi.org/10.1017/S095410202300024X>
- Ji, M., Greening, C., Vanwonderghem, I., Carere, C. R., Bay, S. K., Steen, J. A., Montgomery, K., Lines, T., Beardall, J., van Dorst, J., Snape, I., Stott, M. B., Hugenholtz, P., & Ferrari, B. C. (2017). Atmospheric trace gases support primary production in Antarctic desert surface soil. *Nature*, 552(7685), 400–403. <https://doi.org/10.1038/nature25014>
- Jia, Y., Zhao, S., Guo, W., Peng, L., Zhao, F., Wang, L., Fan, G., Zhu, Y., Xu, D., Liu, G., Wang, R., Fang, X., Zhang, H., Kristiansen, K., Zhang, W., & Chen, J. (2022). Sequencing introduced false

positive rare taxa lead to biased microbial community diversity, assembly, and interaction interpretation in amplicon studies. *Environmental Microbiome*, 17(1), 43. <https://doi.org/10.1186/s40793-022-00436-y>

Johnson, J. S., Spakowicz, D. J., Hong, B.-Y., Petersen, L. M., Demkowicz, P., Chen, L., Leopold, S. R., Hanson, B. M., Agresta, H. O., Gerstein, M., Sodergren, E., & Weinstock, G. M. (2019). Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. *Nature Communications*, 10(1), 5029. <https://doi.org/10.1038/s41467-019-13036-1>

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., & Hunter, S. (2014). InterProScan 5: Genome-scale protein function classification. *Bioinformatics*, 30(9), 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>

Jordan, I. K., Rogozin, I. B., Wolf, Y. I., & Koonin, E. V. (2002). Microevolutionary Genomics of Bacteria. *Theoretical Population Biology*, 61(4), 435–447. <https://doi.org/10.1006/tpbi.2002.1588>

Jungblut, A. D., & Hawes, I. (2017). Using Captain Scott's Discovery specimens to unlock the past: Has Antarctic cyanobacterial diversity changed over the last 100 years? *Proceedings of the Royal Society B: Biological Sciences*, 284(1857), 20170833. <https://doi.org/10.1098/rspb.2017.0833>

Jungblut, A. D., Hawes, I., Mackey, T. J., Krusor, M., Doran, P. T., Sumner, D. Y., Eisen, J. A., Hillman, C., & Goroncy, A. K. (2016). Microbial Mat Communities along an Oxygen Gradient in a Perennially Ice-Covered Antarctic Lake. *Applied and Environmental Microbiology*, 82(2), 620–630. <https://doi.org/10.1128/AEM.02699-15>

Jungblut, A. D., Lovejoy, C., & Vincent, W. F. (2010). Global distribution of cyanobacterial ecotypes in the cold biosphere. *The ISME Journal*, 4(2), 191–202. <https://doi.org/10.1038/ismej.2009.113>

Kanesaki, Y., Shiwa, Y., Tajima, N., Suzuki, M., Watanabe, S., Sato, N., Ikeuchi, M., & Yoshikawa, H. (2012). Identification of Substrain-Specific Mutations by Massively Parallel Whole-Genome Resequencing of *Synechocystis* sp. PCC 6803. *DNA Research*, 19(1), 67–79. <https://doi.org/10.1093/dnares/dsr042>

Kaštovský, J. (2024). Welcome to the jungle!: An overview of modern taxonomy of cyanobacteria. *Hydrobiologia*, 851(4), 1063–1077. <https://doi.org/10.1007/s10750-023-05356-7>

Katoh, K., & Standley, D. M. (2013). MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution*, 30(4), 772–780. <https://doi.org/10.1093/molbev/mst010>

Kaur, H., Shannon, L. M., & Samac, D. A. (2024). A stepwise guide for pangenome development in crop plants: An alfalfa (*Medicago sativa*) case study. *BMC Genomics*, 25(1), 1022. <https://doi.org/10.1186/s12864-024-10931-w>

- Keshari, N., Zhao, Y., Das, S. K., Zhu, T., & Lu, X. (2022). Cyanobacterial Community Structure and Isolates From Representative Hot Springs of Yunnan Province, China Using an Integrative Approach. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.872598>
- Kleinteich, J., Hildebrand, F., Bahram, M., Voigt, A. Y., Wood, S. A., Jungblut, A. D., Küpper, F. C., Quesada, A., Camacho, A., Pearce, D. A., Convey, P., Vincent, W. F., Zarfl, C., Bork, P., & Dietrich, D. R. (2017). Pole-to-Pole Connections: Similarities between Arctic and Antarctic Microbiomes and Their Vulnerability to Environmental Change. *Frontiers in Ecology and Evolution*, 5. <https://doi.org/10.3389/fevo.2017.00137>
- Kociolek, J. P., Kopalová, K., Hamsher, S. E., Kohler, T. J., Van de Vijver, B., Convey, P., & McKnight, D. M. (2017). Freshwater diatom biogeography and the genus *Luticola*: An extreme case of endemism in Antarctica. *Polar Biology*, 40(6), 1185–1196. <https://doi.org/10.1007/s00300-017-2090-7>
- Komárek, J. (2016). A polyphasic approach for the taxonomy of cyanobacteria: Principles and applications. *European Journal of Phycology*, 51(3), 346–353. <https://doi.org/10.1080/09670262.2016.1163738>
- Komárek, J., Kaštovský, J., Mareš, J., & Johansen, J. R. (2014). Taxonomic classification of cyanoprokaryotes (cyanobacterial genera) 2014, using a polyphasic approach. *Preslia*, 86, 295–335.
- Komárek, O., & Komárek, J. (2010). Diversity and Ecology of Cyanobacterial Microflora of Antarctic Seepage Habitats: Comparison of King George Island, Shetland Islands, and James Ross Island, NW Weddell Sea, Antarctica. In J. Seckbach & A. Oren (Eds.), *Microbial Mats: Modern and Ancient Microorganisms in Stratified Systems* (pp. 515–539). Springer Netherlands. https://doi.org/10.1007/978-90-481-3799-2_27
- Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences*, 102(7), 2567–2572. <https://doi.org/10.1073/pnas.0409727102>
- Koo, H., Mojib, N., Hakim, J. A., Hawes, I., Tanabe, Y., Andersen, D. T., & Bej, A. K. (2017). Microbial Communities and Their Predicted Metabolic Functions in Growth Laminae of a Unique Large Conical Mat from Lake Untersee, East Antarctica. *Frontiers in Microbiology*, 8. <https://doi.org/10.3389/fmicb.2017.01347>
- Kopf, M., Klähn, S., Scholz, I., Matthiessen, J. K. F., Hess, W. R., & Voß, B. (2014). Comparative Analysis of the Primary Transcriptome of *Synechocystis* sp. PCC 6803. *DNA Research*, 21(5), 527–539. <https://doi.org/10.1093/dnares/dsu018>
- Koskinen, S., Kurkela, J., Linhartová, M., & Tyystjärvi, T. (2023). The genome sequence of *Synechocystis* sp. PCC 6803 substrain GT-T and its implications for the evolution of PCC 6803 substrains. *FEBS Open Bio*, 13(4), 701–712. <https://doi.org/10.1002/2211-5463.13576>

- Lamoreux, J. F., Morrison, J. C., Ricketts, T. H., Olson, D. M., Dinerstein, E., McKnight, M. W., & Shugart, H. H. (2006). Global tests of biodiversity concordance and the importance of endemism. *Nature*, 440(7081), 212–214. <https://doi.org/10.1038/nature04291>
- Lara, Y., Durieu, B., Cornet, L., Verlaine, O., Rippka, R., Pessi, I. S., Misztak, A., Joris, B., Javaux, E. J., Baurain, D., & Wilmotte, A. (2017). Draft Genome Sequence of the Axenic Strain *Phormidesmis priestleyi* ULC007, a Cyanobacterium Isolated from Lake Bruehwiler (Larsemann Hills, Antarctica). *Genome Announcements*, 5(7), e01546-16. <https://doi.org/10.1128/genomeA.01546-16>
- Larsson, J., Nylander, J. A., & Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evolutionary Biology*, 11(1), 187. <https://doi.org/10.1186/1471-2148-11-187>
- Laybourn-Parry, J., & Bayliss, P. (1996). Seasonal dynamics of the planktonic community in Lake Druzhby, Princess Elizabeth Land, Eastern Antarctica. *Freshwater Biology*, 35(1), 57–67. <https://doi.org/10.1046/j.1365-2427.1996.00480.x>
- Laybourn-Parry, J., Quayle, W., & Henshaw, T. (2002). The biology and evolution of Antarctic saline lakes in relation to salinity and trophy. *Polar Biology*, 25(7), 542–552. <https://doi.org/10.1007/s00300-002-0383-x>
- Lebre, P. H., Bottos, E., Makhalanyane, T. P., Hogg, I., & Cowan, D. A. (2021). Islands in the sand: Are all hypolithic microbial communities the same? *FEMS Microbiology Ecology*, 97(1), fiae216. <https://doi.org/10.1093/femsec/fiae216>
- Lee, J. R., Raymond, B., Bracegirdle, T. J., Chadès, I., Fuller, R. A., Shaw, J. D., & Terauds, A. (2017). Climate change drives expansion of Antarctic ice-free habitat. *Nature*, 547(7661), 49–54. <https://doi.org/10.1038/nature22996>
- Lefler, F. W., Berthold, D. E., & Laughinghouse IV, H. D. (2023). Cyanoseq: A database of cyanobacterial 16S rRNA gene sequences with curated taxonomy. *Journal of Phycology*, 59(3), 470–480. <https://doi.org/10.1111/jpy.13335>
- Letunic, I., & Bork, P. (2024). Interactive Tree of Life (iTOL) v6: Recent updates to the phylogenetic tree display and annotation tool. *Nucleic Acids Research*, 52(W1), W78–W82. <https://doi.org/10.1093/nar/gkae268>
- Liu, G., Li, T., Zhu, X., Zhang, X., & Wang, J. (2023). An independent evaluation in a CRC patient cohort of microbiome 16S rRNA sequence analysis methods: OTU clustering, DADA2, and Deblur. *Frontiers in Microbiology*, 14. <https://doi.org/10.3389/fmicb.2023.1178744>
- Lomolino, M. V., Riddle, B. R., & Brown, J. H. (2006). *Biogeography* (3rd ed). Sinauer Associates.
- Lozupone, C. A., & Knight, R. (2007). Global patterns in bacterial diversity. *Proceedings of the National Academy of Sciences*, 104(27), 11436–11440. <https://doi.org/10.1073/pnas.0611525104>

- Lumian, J. E., Jungblut, A. D., Dillion, M. L., Hawes, I., Doran, P. T., Mackey, T. J., Dick, G. J., Grettenberger, C. L., & Sumner, D. Y. (2021). Metabolic Capacity of the Antarctic Cyanobacterium *Phormidium pseudopriestleyi* That Sustains Oxygenic Photosynthesis in the Presence of Hydrogen Sulfide. *Genes*, 12(3), 426. <https://doi.org/10.3390/genes12030426>
- Lumian, J., Grettenberger, C., Jungblut, A. D., Mackey, T. J., Hawes, I., Alatorre-Acevedo, E., & Sumner, D. Y. (2024). Genomic profiles of four novel cyanobacteria MAGs from Lake Vanda, Antarctica: Insights into photosynthesis, cold tolerance, and the circadian clock. *Frontiers in Microbiology*, 14, 1330602. <https://doi.org/10.3389/fmicb.2023.1330602>
- Lumian, J., Sumner, D. Y., Grettenberger, C. L., Jungblut, A. D., Irber, L., Pierce-Ward, N. T., & Brown, C. T. (2024). Biogeographic distribution of five Antarctic cyanobacteria using large-scale k-mer searching with sourmash branchwater. *Frontiers in Microbiology*, 15, 1328083. <https://doi.org/10.3389/fmicb.2024.1328083>
- Luz, R., Cordeiro, R., Kaštovský, J., Johansen, J. R., Dias, E., Fonseca, A., Urbatzka, R., Vasconcelos, V., & Gonçalves, V. (2023). Description of four new filamentous cyanobacterial taxa from freshwater habitats in the Azores Archipelago. *Journal of Phycology*, 59(6), 1323–1338. <https://doi.org/10.1111/jpy.13396>
- Ma, B., Dai, Z., Wang, H., Dsouza, M., Liu, X., He, Y., Wu, J., Rodrigues, J. L. M., Gilbert, J. A., Brookes, P. C., & Xu, J. (2017). Distinct Biogeographic Patterns for Archaea, Bacteria, and Fungi along the Vegetation Gradient at the Continental Scale in Eastern China. *mSystems*, 2(1), 10.1128/msystems.00174-16. <https://doi.org/10.1128/msystems.00174-16>
- Maistrenko, O. M., Mende, D. R., Luetge, M., Hildebrand, F., Schmidt, T. S. B., Li, S. S., Rodrigues, J. F. M., von Mering, C., Pedro Coelho, L., Huerta-Cepas, J., Sunagawa, S., & Bork, P. (2020). Disentangling the impact of environmental and phylogenetic constraints on prokaryotic within-species diversity. *The ISME Journal*, 14(5), 1247–1259. <https://doi.org/10.1038/s41396-020-0600-z>
- Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., Berka, J., Braverman, M. S., Chen, Y.-J., Chen, Z., Dewell, S. B., Du, L., Fierro, J. M., Gomes, X. V., Godwin, B. C., He, W., Helgesen, S., Ho, C. H., Irzyk, G. P., ... Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, 437(7057), 376–380. <https://doi.org/10.1038/nature03959>
- Matsen, F. A., Kodner, R. B., & Armbrust, E. V. (2010). pplacer: Linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics*, 11(1), 538. <https://doi.org/10.1186/1471-2105-11-538>
- Matthews, C. A., Watson-Haigh, N. S., Burton, R. A., & Sheppard, A. E. (2024). A gentle introduction to pangenomics. *Briefings in Bioinformatics*, 25(6), bbae588. <https://doi.org/10.1093/bib/bbae588>
- McInerney, J. O., McNally, A., & O'Connell, M. J. (2017). Why prokaryotes have pangenomes. *Nature Microbiology*, 2(4), 17040. <https://doi.org/10.1038/nmicrobiol.2017.40>

- Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., & Ferrin, T. E. (2023). UCSF ChimeraX: Tools for structure building and analysis. *Protein Science*, 32(11), e4792. <https://doi.org/10.1002/pro.4792>
- Méric, G., Yahara, K., Mageiros, L., Pascoe, B., Maiden, M. C. J., Jolley, K. A., & Sheppard, S. K. (2014). A Reference Pan-Genome Approach to Comparative Bacterial Genomics: Identification of Novel Epidemiological Markers in Pathogenic *Campylobacter*. *PLOS ONE*, 9(3), e92798. <https://doi.org/10.1371/journal.pone.0092798>
- Milici, M., Tomasch, J., Wos-Oxley, M. L., Decelle, J., Jáuregui, R., Wang, H., Deng, Z.-L., Plumeier, I., Giebel, H.-A., Badewien, T. H., Wurst, M., Pieper, D. H., Simon, M., & Wagner-Döbler, I. (2016). Bacterioplankton Biogeography of the Atlantic Ocean: A Case Study of the Distance-Decay Relationship. *Frontiers in Microbiology*, 7. <https://doi.org/10.3389/fmicb.2016.00590>
- Morel, B., Williams, T. A., Stamatakis, A., & Szöllősi, G. J. (2024). AleRax: A tool for gene and species tree co-estimation and reconciliation under a probabilistic model of gene duplication, transfer, and loss. *Bioinformatics*, 40(4), btae162. <https://doi.org/10.1093/bioinformatics/btae162>
- Müller, R., & Nebel, M. (2021). On the use of sequence-quality information in OTU clustering. *PeerJ*, 9, e11717. <https://doi.org/10.7717/peerj.11717>
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, 403(6772), 853–858. <https://doi.org/10.1038/35002501>
- Namsaraev, Z., Mano, M.-J., Fernandez, R., & Wilmotte, A. (2010). Biogeography of terrestrial cyanobacteria from Antarctic ice-free areas. *Annals of Glaciology*, 51(56), 171–177. <https://doi.org/10.3189/172756411795931930>
- Nemergut, D. R., Costello, E. K., Hamady, M., Lozupone, C., Jiang, L., Schmidt, S. K., Fierer, N., Townsend, A. R., Cleveland, C. C., Stanish, L., & Knight, R. (2011). Global patterns in the biogeography of bacterial taxa. *Environmental Microbiology*, 13(1), 135–144. <https://doi.org/10.1111/j.1462-2920.2010.02315.x>
- Niederberger, T. D., Sohmi, J. A., Tirindelli, J., Gunderson, T., Capone, D. G., Carpenter, E. J., & Cary, S. C. (2012). Diverse and highly active diazotrophic assemblages inhabit ephemerally wetted soils of the Antarctic Dry Valleys. *FEMS Microbiology Ecology*, 82(2), 376–390. <https://doi.org/10.1111/j.1574-6941.2012.01390.x>
- Nilsen, T., Snipen, L.-G., Angell, I. L., Keeley, N. B., Majaneva, S., Pettersen, R., & Rudi, K. (2024). Swarm and UNOISE outperform DADA2 and Deblur for denoising high-diversity marine seafloor samples. *ISME Communications*, 4(1), ycae071. <https://doi.org/10.1093/ismeco/ycae071>
- Nübel, U., Garcia-Pichel, F., & Muyzer, G. (1997). PCR primers to amplify 16S rRNA genes from cyanobacteria. *Applied and Environmental Microbiology*, 63(8), 3327–3332. <https://doi.org/10.1128/aem.63.8.3327-3332.1997>

- Oakley, B. B., Carbonero, F., van der Gast, C. J., Hawkins, R. J., & Purdy, K. J. (2010). Evolutionary divergence and biogeography of sympatric niche-differentiated bacterial populations. *The ISME Journal*, 4(4), 488–497. <https://doi.org/10.1038/ismej.2009.146>
- Obbels, D., Verleyen, E., Mano, M.-J., Namsaraev, Z., Sweetlove, M., Tytgat, B., Fernandez-Carazo, R., De Wever, A., D'hondt, S., Ertz, D., Elster, J., Sabbe, K., Willems, A., Wilmotte, A., & Vyverman, W. (2016). Bacterial and eukaryotic biodiversity patterns in terrestrial and aquatic habitats in the Sør Rondane Mountains, Dronning Maud Land, East Antarctica. *FEMS Microbiology Ecology*, 92(6), fiw041. <https://doi.org/10.1093/femsec/fiw041>
- O'Mullan, G. D., & Ward, B. B. (2005). Relationship of Temporal and Spatial Variabilities of Ammonia-Oxidizing Bacteria to Nitrification Rates in Monterey Bay, California. *Applied and Environmental Microbiology*, 71(2), 697–705. <https://doi.org/10.1128/AEM.71.2.697-705.2005>
- Pearce, D. A. (2017). Extremophiles in Antarctica: Life at Low Temperatures. In H. Stan-Lotter & S. Fendrihan (Eds.), *Adaption of Microbial Life to Environmental Extremes: Novel Research Results and Application* (pp. 99–131). Springer International Publishing. https://doi.org/10.1007/978-3-319-48327-6_5
- Pérez-Carrascal, O. M., Tromas, N., Terrat, Y., Moreno, E., Giani, A., Corrêa Braga Marques, L., Fortin, N., & Shapiro, B. J. (2021). Single-colony sequencing reveals microbe-by-microbiome phyllosymbiosis between the cyanobacterium *Microcystis* and its associated bacteria. *Microbiome*, 9(1), 194. <https://doi.org/10.1186/s40168-021-01140-8>
- Pessi, I. (2017). The cyanobacterial biota of polar regions: A molecular approach. *PhD Thesis*.
- Pessi, I. S., Lara, Y., Durieu, B., Maalouf, P. D. C., Verleyen, E., & Wilmotte, A. (2018). Community structure and distribution of benthic cyanobacteria in Antarctic lacustrine microbial mats. *FEMS Microbiology Ecology*, 94(5). <https://doi.org/10.1093/femsec/fiy042>
- Pessi, I. S., Maalouf, P. D. C., Laughinghouse IV, H. D., Baurain, D., & Wilmotte, A. (2016). On the use of high-throughput sequencing for the study of cyanobacterial diversity in Antarctic aquatic mats. *Journal of Phycology*, 52(3), 356–368. <https://doi.org/10.1111/jpy.12399>
- Pessi, I. S., Popin, R. V., Durieu, B., Lara, Y., Tytgat, B., Savaglia, V., Roncero-Ramos, B., Hultman, J., Verleyen, E., Vyverman, W., & Wilmotte, A. (2023). Novel diversity of polar Cyanobacteria revealed by genomeresolved metagenomics. *Microbial Genomics*, 9(7). <https://doi.org/10.1099/mgen.0.001056>
- Pisa, S., Biersma, E. M., Convey, P., Patiño, J., Vanderpoorten, A., Werner, O., & Ros, R. M. (2014). The cosmopolitan moss *Bryum argenteum* in Antarctica: Recent colonisation or in situ survival? *Polar Biology*, 37(10), 1469–1477. <https://doi.org/10.1007/s00300-014-1537-3>
- Pittera, J., Jouhet, J., Breton, S., Garczarek, L., Partensky, F., Maréchal, É., Nguyen, N. A., Doré, H., Ratin, M., Pitt, F. D., Scanlan, D. J., & Six, C. (2018). Thermoacclimation and genome adaptation of the membrane lipidome in marine *Synechococcus*. *Environmental Microbiology*, 20(2), 612–631. <https://doi.org/10.1111/1462-2920.13985>

- Pointing, S. B., Bollard-Breen, B., & Gillman, L. N. (2014). Diverse cryptic refuges for life during glaciation. *Proceedings of the National Academy of Sciences*, 111(15), 5452–5453. <https://doi.org/10.1073/pnas.1403594111>
- Prieto-Barajas, C. M., Valencia-Cantero, E., & Santoyo, G. (2018). Microbial mat ecosystems: Structure types, functional diversity, and biotechnological application. *Electronic Journal of Biotechnology*, 31, 48–56. <https://doi.org/10.1016/j.ejbt.2017.11.001>
- Proteau, P. J., Gerwick, W. H., Garcia-Pichel, F., & Castenholz, R. (1993). The structure of scytonemin, an ultraviolet sunscreen pigment from the sheaths of cyanobacteria. *Experientia*, 49(9), 825–829. <https://doi.org/10.1007/BF01923559>
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glöckner, F. O. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Research*, 41(D1), D590–D596. <https://doi.org/10.1093/nar/gks1219>
- Quesada, A., & Vincent, W. F. (1997). Strategies of adaptation by Antarctic cyanobacteria to ultraviolet radiation. *European Journal of Phycology*, 32(4), 335–342. <https://doi.org/10.1080/09670269710001737269>
- 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>
- Rego, A., Raio, F., Martins, T. P., Ribeiro, H., Sousa, A. G. G., Séneca, J., Baptista, M. S., Lee, C. K., Cary, S. C., Ramos, V., Carvalho, M. F., Leão, P. N., & Magalhães, C. (2019). Actinobacteria and Cyanobacteria Diversity in Terrestrial Antarctic Microenvironments Evaluated by Culture-Dependent and Independent Methods. *Frontiers in Microbiology*, 10. <https://doi.org/10.3389/fmicb.2019.01018>
- Robinson, S. A., & Waterman, M. J. (2014). Sunsafe Bryophytes: Photoprotection from Excess and Damaging Solar Radiation. In D. T. Hanson & S. K. Rice (Eds.), *Photosynthesis in Bryophytes and Early Land Plants* (Vol. 37, pp. 113–130). Springer Netherlands. https://doi.org/10.1007/978-94-007-6988-5_7
- Roncero-Ramos, B., Savaglia, V., Durieu, B., Van de Vreken, I., Richel, A., & Wilmotte, A. (2024). Ecophysiological and genomic approaches to cyanobacterial hardening for restoration. *Journal of Phycology*, 60(2), 465–482. <https://doi.org/10.1111/jpy.13436>
- Roos, J. C., & Vincent, W. F. (1998). Temperature Dependence of Uv Radiation Effects on Antarctic Cyanobacteria. *Journal of Phycology*, 34(1), 118–125. <https://doi.org/10.1046/j.1529-8817.1998.340118.x>
- Rossi, F., & De Philippis, R. (2015). Role of Cyanobacterial Exopolysaccharides in Phototrophic Biofilms and in Complex Microbial Mats. *Life*, 5(2), 1218–1238. <https://doi.org/10.3390/life5021218>

- Rousk, J., Bååth, E., Brookes, P. C., Lauber, C. L., Lozupone, C., Caporaso, J. G., Knight, R., & Fierer, N. (2010). Soil bacterial and fungal communities across a pH gradient in an arable soil. *The ISME Journal*, 4(10), 1340–1351. <https://doi.org/10.1038/ismej.2010.58>
- Rubio, A., & Pérez-Pulido, A. J. (2021). Protein-Coding Genes of *Helicobacter pylori* Predominantly Present Purifying Selection though Many Membrane Proteins Suffer from Selection Pressure: A Proposal to Analyze Bacterial Pangenomes. *Genes*, 12(3), 377. <https://doi.org/10.3390/genes12030377>
- Sánchez-Baracaldo, P., Bianchini, G., Wilson, J. D., & Knoll, A. H. (2022). Cyanobacteria and biogeochemical cycles through Earth history. *Trends in Microbiology*, 30(2), 143–157. <https://doi.org/10.1016/j.tim.2021.05.008>
- Sanger, F., Nicklen, S., & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74(12), 5463–5467. <https://doi.org/10.1073/pnas.74.12.5463>
- Santini, T. C., Gramenz, L., Southam, G., & Zammit, C. (2022). Microbial Community Structure Is Most Strongly Associated With Geographical Distance and pH in Salt Lake Sediments. *Frontiers in Microbiology*, 13. <https://doi.org/10.3389/fmicb.2022.920056>
- Sarasa-Buisán, C., Nieves-Mori6n, M., Lindblad, P., Nierzwicki-Bauer, S., Schluepmann, H., & Flores, E. (2025). Intercellular communication in the fern endosymbiotic cyanobacterium *Nostoc azollae*. *mBio*, 16(9), e01187-25. <https://doi.org/10.1128/mbio.01187-25>
- Sarton-Loh6ac, G., Romashchenko, N., Train, C. M., Majidian, S., & Glover, N. (2025). Reconstructing Evolutionary Histories with Hierarchical Orthologous Groups. *Journal of Molecular Evolution*, 93(6), 740–764. <https://doi.org/10.1007/s00239-025-10277-1>
- Sayers, E. W., Beck, J., Bolton, E. E., Brister, J. R., Chan, J., Connor, R., Feldgarden, M., Fine, A. M., Funk, K., Hoffman, J., Kannan, S., Kelly, C., Klimke, W., Kim, S., Lathrop, S., Marchler-Bauer, A., Murphy, T. D., O’Sullivan, C., Schmieder, E., ... Pruitt, K. D. (2025). Database resources of the National Center for Biotechnology Information in 2025. *Nucleic Acids Research*, 53(D1), D20–D29. <https://doi.org/10.1093/nar/gkae979>
- Scambos, T. A., Campbell, G. G., Pope, A., Haran, T., Muto, A., Lazzara, M., Reijmer, C. H., & van den Broeke, M. R. (2018). Ultralow Surface Temperatures in East Antarctica From Satellite Thermal Infrared Mapping: The Coldest Places on Earth. *Geophysical Research Letters*, 45(12), 6124–6133. <https://doi.org/10.1029/2018GL078133>
- Schirmer, M., Ijaz, U. Z., D’Amore, R., Hall, N., Sloan, W. T., & Quince, C. (2015). Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. *Nucleic Acids Research*, 43(6), e37. <https://doi.org/10.1093/nar/gku1341>
- Schirrmeister, B. E., Antonelli, A., & Bagheri, H. C. (2011). The origin of multicellularity in cyanobacteria. *BMC Evolutionary Biology*, 11(1), 45. <https://doi.org/10.1186/1471-2148-11-45>

- Schnell, I. B., Bohmann, K., & Gilbert, M. T. P. (2015). Tag jumps illuminated – reducing sequence-to-sample misidentifications in metabarcoding studies. *Molecular Ecology Resources*, 15(6), 1289–1303. <https://doi.org/10.1111/1755-0998.12402>
- Schopf, J. W., Kudryavtsev, A. B., Czaja, A. D., & Tripathi, A. B. (2007). Evidence of Archean life: Stromatolites and microfossils. *Precambrian Research, Earliest Evidence of Life on Earth*, 158(3), 141–155. <https://doi.org/10.1016/j.precamres.2007.04.009>
- Segawa, T., Yonezawa, T., Edwards, A., Akiyoshi, A., Tanaka, S., Uetake, J., Irvine-Fynn, T., Fukui, K., Li, Z., & Takeuchi, N. (2017). Biogeography of cryoconite forming cyanobacteria on polar and Asian glaciers. *Journal of Biogeography*, 44(12), 2849–2861. <https://doi.org/10.1111/jbi.13089>
- Segawa, T., Yonezawa, T., Matsuzaki, R., Mori, H., Akiyoshi, A., Navarro, F., Fujita, K., Aizen, V. B., Li, Z., Mano, S., & Takeuchi, N. (2023). Evolution of snow algae, from cosmopolitans to endemics, revealed by DNA analysis of ancient ice. *The ISME Journal*, 17(4), 491–501. <https://doi.org/10.1038/s41396-023-01359-3>
- Shaw, J. D., Terauds, A., Riddle, M. J., Possingham, H. P., & Chown, S. L. (2014). Antarctica's Protected Areas Are Inadequate, Unrepresentative, and at Risk. *PLOS Biology*, 12(6), e1001888. <https://doi.org/10.1371/journal.pbio.1001888>
- Singh, S. M., & Elster, J. (2007). Cyanobacteria in Antarctic Lake Environments: A mini_review. In *Algae and Cyanobacteria in Extreme Environments* (Springer, Vol. 11). https://doi.org/https://doi.org/10.1007/978-1-4020-6112-7_16 (Original work published Seckbach, J.)
- Skoupý, S., Stanojković, A., Pavlíková, M., Poulíčková, A., & Dvořák, P. (2022). New cyanobacterial genus *Argonema* is hiding in soil crusts around the world. *Scientific Reports*, 12(1), 7203. <https://doi.org/10.1038/s41598-022-11288-4>
- Smillie, C. S., Smith, M. B., Friedman, J., Cordero, O. X., David, L. A., & Alm, E. J. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature*, 480(7376), 241–244. <https://doi.org/10.1038/nature10571>
- Stahl, D. A., Flowers, J. J., Hullar, M., & Davidson, S. (2013). Structure and Function of Microbial Communities. In *The Prokaryotes* (pp. 3–30). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-30123-0_34
- Staley, J. T., & Gosink, J. J. (1999). Poles Apart: Biodiversity and Biogeography of Sea Ice Bacteria. *Annual Review of Microbiology*, 53(1), 189–215. <https://doi.org/10.1146/annurev.micro.53.1.189>
- Stamatakis, A. (2006). RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21), 2688–2690. <https://doi.org/10.1093/bioinformatics/btl446>

- Stattersfield, A. J., Crosbv, M. J., Long, A. J., & Wege, D. C. (1998). Endemic Bird Areas of the World: Priorities for Biodiversity Conservation. *The Condor: Ornithological Applications*, 100(4), 770–771. <https://doi.org/10.2307/1369765>
- Stevens, M. I., & D’Haese, C. A. (2014). Islands in ice: Isolated populations of *Cryptopygus sverdrupi* (Collembola) among nunataks in the Sør Rondane Mountains, Dronning Maud Land, Antarctica. *Biodiversity*, 15(2–3), 169–177. <https://doi.org/10.1080/14888386.2014.928791>
- Sudianto, E., Shlafstein, M. D., Durieu, B., Harmel, M., Cornet, L., & Saw, J. H. (2026). Taxonomic description of cyanobacteria from extreme habitats through genome-based classification. *Frontiers in Microbiology*, 17. <https://doi.org/10.3389/fmicb.2026.1824103>
- Sumner, D. Y., Jungblut, A. D., Hawes, I., Andersen, D. T., Mackey, T. J., & Wall, K. (2016). Growth of elaborate microbial pinnacles in Lake Vanda, Antarctica. *Geobiology*, 14(6), 556–574. <https://doi.org/10.1111/gbi.12188>
- Takacs-Vesbach, C., Mitchell, K., Jackson-Weaver, O., & Reysenbach, A.-L. (2008). Volcanic calderas delineate biogeographic provinces among Yellowstone thermophiles. *Environmental Microbiology*, 10(7), 1681–1689. <https://doi.org/10.1111/j.1462-2920.2008.01584.x>
- Takacs-Vesbach, C., Zeglin, L., Barrett, J. E., Gooseff, M. N., & Priscu, J. C. (2010). Factors promoting microbial diversity in the McMurdo Dry Valleys, Antarctica. In D. M. McKnight, P. T. Doran, & W. B. Lyons (Eds.), *Life in Antarctic Deserts and other Cold Dry Environments: Astrobiological Analogs* (pp. 221–257). Cambridge University Press. <https://doi.org/10.1017/CBO9780511712258.008>
- Tallada, S., Hall, G., Barich, D., Morgan-Kiss, R. M., & Slonczewski, J. L. (2022). Antibiotic resistance genes and taxa analysis from mat and planktonic microbiomes of Antarctic perennial ice-covered Lake Fryxell and Lake Bonney. *Antarctic Science*, 34(6), 408–422. <https://doi.org/10.1017/S0954102022000360>
- Tanabe, Y., & Kudoh, S. (2012). Possible ecological implications of floating microbial assemblages lifted from the lakebed on an Antarctic lake. *Ecological Research*, 27(2), 359–367. <https://doi.org/10.1007/s11284-011-0907-3>
- Tang, E. P. Y., & Vincent, W. F. (1999). Strategies of thermal adaptation by high-latitude cyanobacteria. *New Phytologist*, 142(2), 315–323. <https://doi.org/10.1046/j.1469-8137.1999.00385.x>
- Tang, J., Du, L.-M., Liang, Y.-M., & Daroch, M. (2019). Complete Genome Sequence and Comparative Analysis of *Synechococcus* sp. CS-601 (SynAce01), a Cold-Adapted Cyanobacterium from an Oligotrophic Antarctic Habitat. *International Journal of Molecular Sciences*, 20(1), 152. <https://doi.org/10.3390/ijms20010152>
- Taton, A., Grubisic, S., Ertz, D., Hodgson, D. A., Piccardi, R., Biondi, N., Tredici, M. R., Mainini, M., Losi, D., Marinelli, F., & Wilmotte, A. (2006). Polyphasic study of Antarctic cyabacterial strains. *Journal of Phycology*, 42(6), 1257–1270. <https://doi.org/10.1111/j.1529-8817.2006.00278.x>

- Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N. S., Wijesundera, R., Ruiz, L. V., Vasco-Palacios, A. M., Thu, P. Q., Suija, A., Smith, M. E., Sharp, C., Saluveer, E., Saitta, A., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa, K., ... Abarenkov, K. (2014). Global diversity and geography of soil fungi. *Science*, 346(6213), 1256688. <https://doi.org/10.1126/science.1256688>
- Terauds, A., & Lee, J. R. (2016). Antarctic biogeography revisited: Updating the Antarctic Conservation Biogeographic Regions. *Diversity and Distributions*, 22(8), 836–840. <https://doi.org/10.1111/ddi.12453>
- Tettelin, H., & Medini, D. (Eds.). (2020). *The Pangenome: Diversity, Dynamics and Evolution of Genomes*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-38281-0>
- Thomas, D. N., Fogg, G. E., Convey, P., Fritsen, C. H., Gili, J.-M., Gradinger, R., Laybourn-Parry, J., Reid, K., Walton, D. W. H., Thomas, D. N., Fogg, G. E., Convey, P., Fritsen, C. H., Gili, J.-M., Gradinger, R., Laybourn-Parry, J., Reid, K., & Walton, D. W. H. (2008). *The Biology of Polar Regions* (Second Edition, Second Edition). Oxford University Press.
- Thompson, P. L., & Fronhofer, E. A. (2019). The conflict between adaptation and dispersal for maintaining biodiversity in changing environments. *Proceedings of the National Academy of Sciences*, 116(42), 21061–21067. <https://doi.org/10.1073/pnas.1911796116>
- Tin, T., Liggett, D., Maher, P. T., & Lamers, M. (Eds.). (2014). *Antarctic Futures: Human Engagement with the Antarctic Environment*. Springer Netherlands. <https://doi.org/10.1007/978-94-007-6582-5>
- Tringe, S. G., & Hugenholtz, P. (2008). A renaissance for the pioneering 16S rRNA gene. *Current Opinion in Microbiology, Antimicrobials/Genomics*, 11(5), 442–446. <https://doi.org/10.1016/j.mib.2008.09.011>
- Turnbull, J. D., & Robinson, S. A. (2009). Accumulation of DNA damage in Antarctic mosses: Correlations with ultraviolet-B radiation, temperature and turf water content vary among species. *Global Change Biology*, 15(2), 319–329. <https://doi.org/10.1111/j.1365-2486.2008.01739.x>
- Turner, J., Barrand, N. E., Bracegirdle, T. J., Convey, P., Hodgson, D. A., Jarvis, M., Jenkins, A., Marshall, G., Meredith, M. P., Roscoe, H., Shanklin, J., French, J., Goosse, H., Guglielmin, M., Gutt, J., Jacobs, S., Ji, M. C. K., Masson-Delmotte, V., Mayewski, P., ... Klepikov, A. (2014). Antarctic climate change and the environment: An update. *Polar Record*, 50(3), 237–259. <https://doi.org/10.1017/S0032247413000296>
- Tytgat, B., Verleyen, E., Sweetlove, M., Van Den Berge, K., Pinseel, E., Hodgson, D. A., Chown, S. L., Sabbe, K., Wilmotte, A., Willems, A., The Polar Lake Sampling Consortium, & Vyverman, W. (2023). Polar lake microbiomes have distinct evolutionary histories. *Science Advances*, 9(46), eade7130. <https://doi.org/10.1126/sciadv.ade7130>
- V. Wintzingerode, F., Göbel, U. B., & Stackebrandt, E. (1997). Determination of microbial diversity in environmental samples: Pitfalls of PCR-based rRNA analysis. *FEMS Microbiology Reviews*, 21(3), 213–229. <https://doi.org/10.1111/j.1574-6976.1997.tb00351.x>

- van Dongen, S., & Abreu-Goodger, C. (2012). Using MCL to Extract Clusters from Networks. In J. van Helden, A. Toussaint, & D. Thieffry (Eds.), *Bacterial Molecular Networks: Methods and Protocols* (pp. 281–295). Springer. https://doi.org/10.1007/978-1-61779-361-5_15
- Van Goethem, M. W., Vikram, S., Cowan, D. A., & Makhalanyane, T. P. (2025). Comparative genomics reveals adaptive traits in novel Antarctic lithic cyanobacteria. *BMC Genomics*, 26(1), 994. <https://doi.org/10.1186/s12864-025-12203-7>
- Van Kranendonk, M. J., Contributors:, Altermann, W., Beard, B. L., Hoffman, P. F., Johnson, C. M., Kasting, J. F., Melezhik, V. A., Nutman, A. P., Papineau, D., & Pirajno, F. (2012). Chapter 16 - A Chronostratigraphic Division of the Precambrian: Possibilities and Challenges. In F. M. Gradstein, J. G. Ogg, M. D. Schmitz, & G. M. Ogg (Eds.), *The Geologic Time Scale* (pp. 299–392). Elsevier. <https://doi.org/10.1016/B978-0-444-59425-9.00016-0>
- Varin, T., Lovejoy, C., Jungblut, A. D., Vincent, W. F., & Corbeil, J. (2012). Metagenomic Analysis of Stress Genes in Microbial Mat Communities from Antarctica and the High Arctic. *Applied and Environmental Microbiology*, 78(2), 549–559. <https://doi.org/10.1128/AEM.06354-11>
- Varliero, G., Lebre, P. H., Adams, B., Chown, S. L., Convey, P., Dennis, P. G., Fan, D., Ferrari, B., Frey, B., Hogg, I. D., Hopkins, D. W., Kong, W., Makhalanyane, T., Matcher, G., Newsham, K. K., Stevens, M. I., Weigh, K. V., & Cowan, D. A. (2024). Biogeographic survey of soil bacterial communities across Antarctica. *Microbiome*, 12(1), 9. <https://doi.org/10.1186/s40168-023-01719-3>
- Varner, J. (2014). Scientific Outreach: Toward Effective Public Engagement with Biological Science. *BioScience*, 64(4), 333–340. <https://doi.org/10.1093/biosci/biu021>
- Velichko, N., Smirnova, S., Averina, S., & Pinevich, A. (2021). A survey of Antarctic cyanobacteria. *Hydrobiologia*, 848(11), 2627–2652. <https://doi.org/10.1007/s10750-021-04588-9>
- Verleyen, E., Van de Vijver, B., Tytgat, B., Pinseel, E., Hodgson, D. A., Kopalová, K., Chown, S. L., Van Ranst, E., Imura, S., Kudoh, S., Van Nieuwenhuyze, W., Consortium, A., Sabbe, K., & Vyverman, W. (2021). Diatoms define a novel freshwater biogeography of the Antarctic. *Ecography*, 44(4), 548–560. <https://doi.org/10.1111/ecog.05374>
- Vincent, W. F. (1988). *Microbial ecosystems of Antarctica*. Cambridge University Press.
- Vincent, W. F. (2002). Cyanobacterial Dominance in the Polar Regions. In B. A. Whitton & M. Potts (Eds.), *The Ecology of Cyanobacteria: Their Diversity in Time and Space* (pp. 321–340). Springer Netherlands. https://doi.org/10.1007/0-306-46855-7_12
- Vincent, W. F. (2007). Cold Tolerance in Cyanobacteria and Life in the Cryosphere. In J. Seckbach (Ed.), *Algae and Cyanobacteria in Extreme Environments* (pp. 287–301). Springer Netherlands. https://doi.org/10.1007/978-1-4020-6112-7_15

- Vincent, W. F., & Quesada, A. (2012). Cyanobacteria in High Latitude Lakes, Rivers and Seas. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (pp. 371–385). Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_13
- Vyverman, W., Verleyen, E., Wilmotte, A., Hodgson, D. A., Willems, A., Peeters, K., Van de Vijver, B., De Wever, A., Leliaert, F., & Sabbe, K. (2010). Evidence for widespread endemism among Antarctic micro-organisms. *Polar Science, Antarctic Biology in the 21st Century - Advances in and beyond IPY*, 4(2), 103–113. <https://doi.org/10.1016/j.polar.2010.03.006>
- Wang, J., Zhang, Q., Wu, G., Zhang, C., Zhang, M., & Zhao, L. (2018). *Minimizing spurious features in 16S rRNA gene amplicon sequencing* (E26872v1). PeerJ Inc. <https://doi.org/10.7287/peerj.preprints.26872v1>
- Wang, X., Ren, Y., Yu, Z., Shen, G., Cheng, H., & Tao, S. (2022). Effects of environmental factors on the distribution of microbial communities across soils and lake sediments in the Hoh Xil Nature Reserve of the Qinghai-Tibetan Plateau. *Science of The Total Environment*, 838, 156148. <https://doi.org/10.1016/j.scitotenv.2022.156148>
- Wauchope, H. S., Shaw, J. D., & Terauds, A. (2019). A snapshot of biodiversity protection in Antarctica. *Nature Communications*, 10(1), 946. <https://doi.org/10.1038/s41467-019-08915-6>
- Welsh, D. T. (2000). Ecological significance of compatible solute accumulation by micro-organisms: From single cells to global climate. *FEMS Microbiology Reviews*, 24(3), 263–290. <https://doi.org/10.1111/j.1574-6976.2000.tb00542.x>
- Wen, W., Lee, J. R., Shaw, J. D., & Wilson, K. A. (2026). Comprehensive and representative conservation planning for biodiversity of ice-free Antarctica. *Conservation Biology*, n/a(n/a), e70352. <https://doi.org/10.1111/cobi.70352>
- Wharton Jr, R. A., Parker, B. C., & Simmons Jr, G. M. (1983). Distribution, species composition and morphology of algal mats in Antarctic dry valley lakes. *Phycologia*, 22(4), 355–365. <https://doi.org/10.2216/i0031-8884-22-4-355.1>
- Whitaker, R. J., Grogan, D. W., & Taylor, J. W. (2003). Geographic Barriers Isolate Endemic Populations of Hyperthermophilic Archaea. *Science*, 301(5635), 976–978. <https://doi.org/10.1126/science.1086909>
- Whitton, B. A., & Potts, M. (2012). Introduction to the Cyanobacteria. In B. A. Whitton (Ed.), *Ecology of Cyanobacteria II: Their Diversity in Space and Time* (pp. 1–13). Springer Netherlands. https://doi.org/10.1007/978-94-007-3855-3_1
- Wirtz, N., Lumbsch, H. T., Green, T. G. A., Türk, R., Pintado, A., Sancho, L., & Schroeter, B. (2003). Lichen fungi have low cyanobiont selectivity in maritime Antarctica. *New Phytologist*, 160(1), 177–183. <https://doi.org/10.1046/j.1469-8137.2003.00859.x>
- Woese, C. R. (1987). Bacterial evolution. *Microbiological Reviews*, 51(2), 221–271. <https://doi.org/10.1128/mr.51.2.221-271.1987>

- Wong, T. K. F., Ly-Trong, N., Ren, H., Demotte, P., Baños, H., Roger, A. J., Susko, E., Bielow, C., De Maio, N., Goldman, N., Hahn, M. W., dos Reis, M., Vinh, L. S., Huttley, G., Lanfear, R., & Minh, B. Q. (2026). IQ-TREE 3: Phylogenomic inference software using complex evolutionary models. *Molecular Biology and Evolution*, 43(5), msag117. <https://doi.org/10.1093/molbev/msag117>
- Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F. O., Ludwig, W., Schleifer, K.-H., Whitman, W. B., Euzéby, J., Amann, R., & Rosselló-Móra, R. (2014). Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nature Reviews Microbiology*, 12(9), 635–645. <https://doi.org/10.1038/nrmicro3330>
- Young, J. C., Waylen, K. A., Sarkki, S., Albon, S., Bainbridge, I., Balian, E., Davidson, J., Edwards, D., Fairley, R., Margerison, C., McCracken, D., Owen, R., Quine, C. P., Stewart-Roper, C., Thompson, D., Tinch, R., Van den Hove, S., & Watt, A. (2014). Improving the science-policy dialogue to meet the challenges of biodiversity conservation: Having conversations rather than talking at one-another. *Biodiversity and Conservation*, 23(2), 387–404. <https://doi.org/10.1007/s10531-013-0607-0>
- Zachos, J., Pagani, M., Sloan, L., Thomas, E., & Billups, K. (2001). Trends, Rhythms, and Aberrations in Global Climate 65 Ma to Present. *Science*, 292(5517), 686–693. <https://doi.org/10.1126/science.1059412>
- Zakhia, F., Jungblut, A.-D., Taton, A., Vincent, W. F., & Wilmotte, A. (2008). Cyanobacteria in Cold Ecosystems. In R. Margesin, F. Schinner, J.-C. Marx, & C. Gerday (Eds.), *Psychrophiles: From Biodiversity to Biotechnology* (pp. 121–135). Springer. https://doi.org/10.1007/978-3-540-74335-4_8
- Zhang, L., Jungblut, A. D., Hawes, I., Andersen, D. T., Sumner, D. Y., & Mackey, T. J. (2015). Cyanobacterial diversity in benthic mats of the McMurdo Dry Valley lakes, Antarctica. *Polar Biology*, 38(8), 1097–1110. <https://doi.org/10.1007/s00300-015-1669-0>
- Zheng, Y.-M., Cao, P., Fu, B., Hughes, J. M., & He, J.-Z. (2013). Ecological Drivers of Biogeographic Patterns of Soil Archaeal Community. *PLOS ONE*, 8(5), e63375. <https://doi.org/10.1371/journal.pone.0063375>

Publications

Peer-review journals

Sudianto, E., Shlafstein, M. D., **Durieu, B.**, Harmel, M., Cornet, L., & Saw, J. H. (2026). Taxonomic description of cyanobacteria from extreme habitats through genome-based classification. *Frontiers in Microbiology*, 17, 1824103. doi:10.3389/fmicb.2026.1824103

Durieu, B., Savaglia, V., Lara, Y., Lambion, A., Pessi, I. S., Vyverman, W., Verleyen, E., & Wilmotte, A. (2025). (Sub-)Antarctic endemic cyanobacteria from benthic mats are rare and have restricted geographic distributions. *Ecography*. doi:10.1111/ecog.07489

Roncero Ramos, B., Savaglia, V., **Durieu, B.**, Van De Vreken, I., Richel, A., & Wilmotte, A. (2024). Ecophysiological and genomic approaches to cyanobacterial hardening for restoration. *Journal of Phycology*, 1-18. doi:10.1111/jpy.13436

Pessi, I. S., Popin, R. V., **Durieu, B.**, Lara, Y., Tytgat, B., Savaglia, V., Roncero-Ramos, B., Hultman, J., Verleyen, E., Vyverman, W., & Wilmotte, A. (2023). Novel diversity of polar Cyanobacteria revealed by genome-resolved metagenomics. *Microbial Genomics*, 9 (7), 001056. doi:10.1099/mgen.0.001056

Cornet, L., **Durieu, B.**, Baert, F., D'hooge, E., Colignon, D., Meunier, L., Lupo, V., Cleenwerck, I., Daniel, H.-M., Rigouts, L., Sirjacobs, D., Declerck, S., Vandamme, P., Wilmotte, A., Baurain, D., & Becker, P. (28 December 2022). The GEN-ERA toolbox: unified and reproducible workflows for research in microbial genomics. *GigaScience*, 12. doi:10.1093/gigascience/giad022

Van Hassel, W., Andjelkovic, M., **Durieu, B.**, Marroquin, V. A., Huybrechts, B., Masquelier, J., & Wilmotte, A. (16 January 2022). A Summer of Cyanobacterial Blooms in Belgian Waterbodies: Microcystin Quantification and Molecular Characterizations. *Toxins*, 14 (1), 61. doi:10.3390/toxins14010061

Beyer, T., Mélanie, C., **Durieu, B.**, & Savaglia, V. (2021). Chercheurs, médiateurs et musées universitaires : une relation à (dé) construire, des pratiques à systématiser. *Les Cahiers de Muséologie*.

Cornet, L., Meunier, L., Van Vlierberghe, M., Léonard, R., **Durieu, B.**, Lara, Y., Misztak, A., Sirjacobs, D., Javaux, E., Philippe, H., Wilmotte, A., & Baurain, D. (2018). Consensus assessment of the contamination level of publicly available cyanobacterial genomes. *PLoS ONE*, 13 (7), 0200323. doi:10.1371/journal.pone.0200323

Pessi, I. S., Lara, Y., **Durieu, B.**, Maalouf, P., Verleyen, E., & Wilmotte, A. (2018). Community structure and distribution of benthic cyanobacteria in Antarctic lacustrine microbial mats. *FEMS Microbiology Ecology*, 94 (5), 042. doi:10.1093/femsec/fiy042

Lara, Y., **Durieu, B.**, Cornet, L., Verlaine, O., Rippka, R., Stelmach Pessi, I., Misztak, A., Joris, B., Javaux, E., Baurain, D., & Wilmotte, A. (2017). Draft genome sequence of the axenic strain

Phormidesmis priestleyi ULC007, a cyanobacterium isolated from Lake Bruehwiler (Larsemann Hills, Antarctica). *Genome Announcements*, 01546-16. doi:10.1128/genomeA.01546-16

Publications submitted or in preparation

Durieu, B., Savaglia, V., Van Vlierberghe, M., Lupo, V., Baurain, D., Wilmotte, A., & Cornet, L. (2025). Out of Antarctica: new insights into Antarctic Subcluster 5.2 picocyanobacteria based on high-quality genomes. *bioRxiv*, 2025-12.

Harmel, M., **Durieu, B.**, Savaglia, V., Tytgat, B., Baurain, D., Wilmotte, A., ... & Cornet, L. (2025). A FAIR Amplicon Sequencing Workflow for Long-term Environmental Monitoring. *bioRxiv*, 2025-12.

Savaglia, V., Roncero-Ramos, B., **Durieu, B.**, Dextro, R. B., Cornet, L., Verleyen, E., Wilmotte, A., & Pessi, I. S. (submitted). (Near-)complete genomes of three Antarctic *Nostoc* strains expand genomic resources for polar cyanobacteria.

Savaglia, V.*, **Durieu, B.***, Tytgat, B., Harmel, M., Elster, J., Vyverman, W., Verleyen, E., & Wilmotte, A. (in prep.). From single-family dominance to balanced assemblages: substrate-driven soil cyanobacterial communities in the Sør Rondane Mountains, East Antarctica.

Reports

Durieu, B., Lara, Y., Pessi, I. S., Wilmotte, A., Willems, A., Tytgat, B., Sweetlove, M., Pinseel, E., Verleyen, E., Vyverman, W., Van de Vijver, B., Van de Putte, A. & Convey, P. (2019). Climate change and Antarctic microbial diversity 'CCAMBIO'. Final report.

Tytgat, B., Willems, A., Sweetlove, M., Pinseel, E., Verleyen, E., Vyverman, W., Stelmach Pessi, I., Lara, Y., **Durieu, B.**, Wilmotte, A., Van de Vijver, B., & Van De Putte, A. (2015). ANNUAL SCIENTIFIC REPORT - Climate change and Antarctic microbial biodiversity (CCAMBIO). (SD/BA/03A, Belgian Science Policy (Belspo)). Bruxelles, Belgium: Belgian Science Policy (Belspo). <https://orbi.uliege.be/handle/2268/189790>

Conferences: oral presentations posters presented as first author

Durieu, B., Savaglia, V., Van Vlierberghe, M., Lupo, V., Baurain, D., Cornet, L. & Wilmotte, A. (2026). New insights into Antarctic picocyanobacteria of the Subcluster 5.2 based in high-quality genomes [Poster presentation]. 23RD Symposium of the International Association for Cyanophyte/Cyanobacteria Research, IAC 2026, Liège, Belgium.

Durieu, B., Savaglia, V., Van Vlierberghe, M., Lupo, V., Baurain, D., Cornet, L. & Wilmotte, A. (2026). New insights into Antarctic picocyanobacteria of the Subcluster 5.2 based in high-quality

genomes [Poster presentation]. 10TH International Conference on Polar and Alpine Microbiology, PAM 2026, Copenhagen, Denmark.

Harmel, M., **Durieu, B.**, Savaglia, V., Tytgat, B., Baurain, D., Wilmotte, A., Verleyen, E. & Cornet, L. (2026). RASPAM: A FAIRE Amplicon Sequencing Workflow for Long-term Environmental Monitoring [Poster presentation]. 10TH International Conference on Polar and Alpine Microbiology, PAM 2026, Copenhagen, Denmark.

Durieu, B., Baurain, D., Wilmotte, A., & Lara, Y. (2018). Cold Adaptation Strategy of the Antarctic Cyanobacterium *P. priestleyi* ULC007 [Poster presentation]. Polar 2018, Where the Poles come together, Open Science Conference, Davos, Switzerland.

Durieu, B., Lara, Y., Stelmach Pessi, I., Lambion, A., Verleyen, E., Baurain, D., & Wilmotte, A. (2018). Biogeography of Cyanobacteria in Antarctic Mats and Implication for Conservation [Oral presentation]. Polar 2018, Where the Poles come together, Open Science Conference, Davos, Switzerland.

Durieu, B., Stelmach Pessi, I., Lara, Y., C Maalouf, P., Lambion, A., & Wilmotte, A. (2017). Biogeographic patterns and genomic adaptation of benthic cyanobacteria in Antarctic lakes [Oral presentation]. In XIIth SCAR Biology Symposium, Leuven, 10-14 July 2017.

Durieu, B., Wilmotte, A., & Lara, Y. (10 November 2017). *Cold adaptation strategy of the Antarctic cyanobacterium P. priestleyi ULC007* [Poster presentation]. SFMBBM Annual PhD day, Liège, Belgium.

Durieu, B., Lara, Y., Obbels, D., Pinseels, E., Stelmach Pessi, I., Sweetlove, M., Tytgat, B., Van De Putte, A., Van De Vijver, B., Verleyen, E., Vyverman, W., Willems, A., & Wilmotte, A. (29 April 2016). *The CCAMBIO project to characterize the biodiversity and distribution of microorganisms in microbial mats of Antarctic lakes* [Poster presentation]. Belgian National Committee for Geodesy and Geophysics (BNCGG) - Belgian National Committee on Antarctic Research (BNCAR) 2016 Symposium, Brussels, Belgium.

Durieu, B., & Michaux, J. (12 December 2014). Evolutionary history of the wood mouse (*Apodemus sylvaticus*) in the Palearctic region, with emphasis on the colonization of the Orkney Islands and Iceland [Poster presentation]. Zoology 2014 - The 21st Benelux Congress of Zoology, Liège, Belgium.

Durieu, B. (2014). Evolutionary history of the wood mouse (*Apodemus sylvaticus*) in the Palearctic region, with emphasis on the colonization of the Orkney Islands and Iceland [Paper presentation]. Jacques Kets Award for Zoology 2014, Antwerp, Belgium.

Book

Savaglia, V., Christiansen, H., Gan, M. Y., Pinseel, E., Jacques, C. & **Durieu, B.** (2019). Imagine l'Antarctique, Denkbeeldig Antarctica. APECS Belgium