Science of the Total Environment 985 (2025) 179699

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Short Communication

Characterization of known and novel clinically important antibiotic resistance genes and novel microbes from wastewater-impacted high Arctic fjord sediments

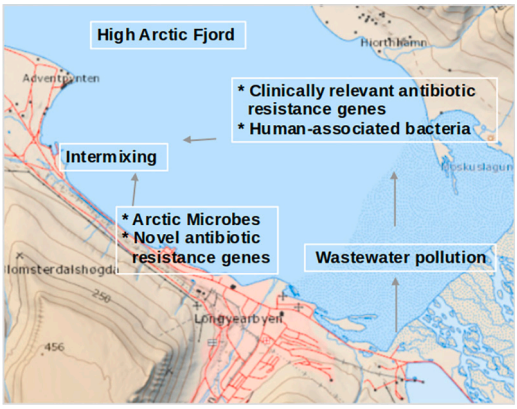
Manish P. Victor [a], Lise Øvreås [b,c], Nachiket P. Marathe [a,*]

[a] Department of Contaminants and Biohazards, Institute of Marine Research (IMR), Bergen, Norway
[b] Department of Biological Sciences, University of Bergen (UiB), Bergen, Norway
[c] The University Centre in Svalbard (UNIS), Norway

HIGHLIGHTS

- High Arctic fjord sediments were studied using metagenomics.
- Clinically important resistance genes like bla_{NDM}, bla_{VIM}, *mcr-1*, *mcr-2*, *cfr* were detected.
- 644 MAGs were assembled from the sediment metagenomes, with >97 % representing novel taxa.
- 478 novel β-lactamases belonging to 217 novel β-lactamase families were discovered.
- The study demonstrates intermixing of human-associated bacteria and novel Arctic bacteria.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Ewa Korzeniewska

Keywords:
Metagenomics
Hidden Markov models
Carbapenamases
Svalbard
Arctic microbiota
β-lactamases

ABSTRACT

Arctic microbiota is enigmatic and highly underexplored. With the aim of understanding the resistome and microbiota of high-Arctic fjord sediments and the effect of wastewater discharge on sediment microbiota, we analyzed sediments from Advent fjord in Svalbard using metagenomics. We show the presence of 888 clinically relevant antibiotic resistance genes including genes coding resistance against last-resort antibiotics such as carbapenems, colistin, vancomycin, linezolid and tigecycline in the sediment microbiota. Using computational models, 478 novel β-lactamases belonging to 217 novel β-lactamase families were revealed in the sediment microbiota. Further, we identified hosts for 69 novel families and showed that these genes are widespread in the Arctic environment. We assembled 644 metagenome-assembled genomes (MAGs) from sediment metagenomes. Of these >97 % belonged to novel taxa with 89 bacterial MAGs representing seven putative novel phyla. These MAGs encoded important functions like nutrient cycling and methane metabolism etc. Our study demonstrated mixing of human associated bacteria and Arctic sediment microbiota. It provides the first comprehensive dataset

* Corresponding author.
E-mail address: nachiket.marathe@hi.no (N.P. Marathe).

https://doi.org/10.1016/j.scitotenv.2025.179699
Received 2 April 2025; Received in revised form 12 May 2025; Accepted 15 May 2025
Available online 26 May 2025
0048-9697/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

of the distribution and diversity of novel microbes and β-lactamases in the wastewater-impacted high Arctic fjord sediments.

1. Introduction

The Polar Regions are geographically distant ecotypes on Earth with extreme environmental conditions like low temperature, low nutrient availability and highly variable light cycles. Microbes are one of the major contributors to total ecosystem biomass, biodiversity, and biogeochemical nutrient cycling. Although climate change and human impact have severe effects on the Arctic, little is known about its impacts on Arctic microbial communities (Cavicchioli et al., 2019; Verde et al., 2016). Arctic sediments still represent huge diversity of novel and hitherto uncharacterized microbes. Understanding the diversity and role of microbes in the Arctic is thus of utmost importance for designing mitigation strategies for preservation of the polar biodiversity (Kleinteich et al., 2017; Edwards et al., 2020). The Advent fjord in Svalbard is exposed to human activities like wastewater discharge and increasing tourism (Kalinowska et al., 2020; Granberg et al., 2017). The aim of this study was to understand the impact of wastewater pollution on sediment microbiota with respect to microbial diversity and antimicrobial resistance (AMR) by characterizing microbiota, genomes of novel bacteria and resistome of the Arctic sediment microbiota. Further we aimed at characterizing novel β-lactamases present in the microbiota as well as identifying their potential hosts.

2. Materials and methods

2.1. Sample collection, sequencing, quality filtering and taxonomy

Sample Collection: Sediment samples were collected from nine sites in Advent fjord in Svalbard in duplicates using a sediment grab and a boat (Supplementary Fig. 1). The samples were collected in sterile 50 ml falcon tubes, frozen and transported to the laboratory. The total DNA was extracted from the sediment samples using DNeasy PowerSoil kit (Qiagen Germany), following the manufacturer's protocol. The extracted DNA was quantified using QubitTM 2000 Spectrophotometer (Thermo Scientific, USA). Sequencing libraries were prepared, using Nextera Library Prep Kit (Illumina, USA), with 2 × 150 bp chemistry, and the sequencing was carried out using Illumina NovsSeq platform at the Norwegian Sequencing Centre, Oslo.

2.2. Quality filtering of the reads and taxonomic classification based on SSU

The FASTQ reads obtained were trimmed with Trim Galore (Version 0.6.10) (trim_galore –quality 30 –phred33 –fastqc –trim-n –cores {} –paired {Forward Sequence} {Reverse Sequence}) (https://github.com/FelixKrueger/TrimGalore). Metaxa2 (Bengtsson-Palme et al., 2015) (V 2.2.3) pipeline was used for the taxonomic assignment of the trimmed reads using small subunit ribosomal ribonucleic acid. The basis of the Metaxa2 SSU database is the SILVA2 reference release 111 (Quast et al., 2012) and Mitozoa3 release10 (D'Onorio De Meo et al., 2012). Metaxa2 -o metaxa_out −1 {Forward Sequence} -2 {Reverse Sequence} -f fastq -g ssu –mode metagenome –usearch v11.0.67_i86linux64 –usearch_bin usearch –cpu {number of cores} –table T –reltax T and thereafter, Metaxa taxonomy traversal tool metaxa2_ttt -i {out_taxonomy.txt} -o {out file} for filtering out taxonomy from phylum to species).

2.3. Detection of metal, biocide, antibiotic resistance genes, mobile genetic elements, and virulence factor

A local database for antimicrobial resistance genes (ARGs) was prepared by combining datasets from NCBI AMRFinderPlus (Feldgarden et al., 2021), CARD (Jia et al., 2017), ARG-ANNOT (Gupta et al., 2014). Biocide and metal resistance genes were obtained from BacMet database (Pal et al., 2014). Sequences for mobile genetic elements were obtained from Pärnänen et al. (Pärnänen et al., 2018). Sequences for virulence factors were retrived from VFDb database (Chen et al., 2005). The quality trimmed reads were analyzed for the respective genes using Usearch v11.0.667_i86linux64 (Edgar, 2010), command line: usearch -userach_global {forward.fastq} -db {local database} -evalue 0.00001 -id 0.9 -query_cov 1 -blast6out {out file} -strand positive -maxaccepts 1 -threads {number of threads}. Obtained hits were length normalized and abundance was expressed as gene copies/16S rDNA copy.

2.4. Assembly and characterization of β-lactamases

fARGene (Berglund et al., 2019) (Fragmented Antibiotic Resistance Gene identifier) (Version v.0.1) was used for assembly of ARGs from the metagenome. It is based upon hidden Markov model (Eddy, 2011) (HMM), we used optimized models for class A β-lactamases, subclass B1 and B2 β-lactamases, subclass B3 β-lactamases, class C and class D β-lactamases. Command line (fargene –infiles {Forward Sequence} {Reverse Sequence} –hmm-model {model name} –score None –meta –output {folder out} –force –processes {cpu} –rerun –amino-dir class_a_out/tmpdir/ –orf-finder) has been used.

2.5. Clustering known and new β-lactamase families from Svalbard and phylogeny

To group the assembled β-lactamase and to identify the novel β-lactamase families from Svalbard (NSFs), clustering was carried out using USEARCH. Initially for each of the classes the known gene sequences were collected from Berglund et al. (Berglund et al., 2017) and β-Lactamase database (BLDB) (Naas et al., 2017) and redundant sequences were removed by clustering at 100 % identity using usearch -cluster_fast {in_file.fasta} -id 1 -centroids {unique_file.fasta} to obtain unique sequences. These sequences were further clustered at 70 % amino acid identity for identifying novel β-lactamases families as suggested in Berglund et al. (2019). Thereafter, an in-house python script was written to extract novel β-lactamases from each cluster, these extracted sequences were aligned with MAFF (Katoh, 2002) and a Newick tree file was generated with FASTTREE (Price et al., 2010) (Version 2.1.11). The phylogenetic trees were reconstructed with TreeDyn in phylogeny.fr (Dereeper et al., 2008).

2.6. Assembly and pathway prediction of MAGs

Assembly was performed using SqueezeMeta (Tamames and Puente-Sánchez, 2019) (V1.6.2). Assembly for the quality trimmed reads fastq files was done using Megahit (Li et al., 2015). Contigs (<1000 bps) were removed and contig statistics were done using prinsEq. (Schmieder and Edwards, 2011). RNAs were predicted using Barrnap (Seemann, 2014) (V 0.9). 16S rRNA sequences were taxonomically classified using the RDP classifier (Wang and Cole, 2024). Binning was done using Metabat2 (Kang et al., 2019) (V2.12.1). Bin statistics were computed using CheckM (Parks et al., 2015). In the end the MAGs obtained were filtered based upon completeness ≥50 % and contamination ≤10 % as suggested by Benoit et al. (Benoit et al., 2024). The metabolic pathways were predicted using a python (Version 2.7.15) based pipeline Fun4me (Sharifi and Ye, 2017), executed by python fun4me.py -i {in_file.fasta} -o {out file}.

2.7. Core-genome analysis of the MAGs

The genomic relatedness of the obtained MAGs was performed using the core genes extracted by the UBCG pipeline (Na et al., 2018) (V2) using default parameters and a maximum-likelihood tree was constructed. The core genes were extracted for all the 644 MAGs using java -jar UBCG.jar extract -bcg_dir core_genome_dir_name -i input_scaffold. fasta -label output_name then all the obtained core genes were aligned with java -jar UBCG.jar align -bcg_dir {core_genome_dir_name} -prefix {out name}. Finally, a newick file is generated with the name out_name. UBCG_gsi(92). codon.50.label.nwk which is used to reconstruct the phylogenetic tree.

2.8. Taxonomic classification of unclassified MAGs

Eighty-nine MAGs not assigned to any phylum were recognized through core-genome (Na et al., 2018) analysis of MAGS and the entire NCBI's dataset for bacteria (esearch -db assembly -query {scaffold_id} | elink -target nuccore | efetch -format fasta > {scaffold name}). An in-house python script was written for extracting the representative bacteria of each validly published genus from the NCBI dataset. Finally, a phylogenetic tree was reconstructed exhibiting the clustering of the 89 MAGs with different genomes representing 3540 bacterial genera.

2.9. Identification of putative hosts for the novel β-lactamases families

For identification of hosts for NSFs we used nR database and DIAMOND (Buchfink et al., 2021) query coverage 95 % and later filtered through percentage identity ≥ 70 % under each matching NSF. For each hit, the metadata in the NCBI was retrieved for finding the habitats of the host bacteria. Similarly, NSF sequences were checked against MAGs obtained in our study to identify putative hosts for the NSFs.

2.10. Distribution of NSF in different niches and statistical analysis

For understanding the distribution of NSF in different niches we used previously published available sequence data from Svalbard for permafrost (Xue et al., 2020), seawater and sediment (Zhang et al., 2022) using parameters described in section 2.4. Further a heatmap and circos plot (Gu et al., 2014) was made for describing 217 NSFs and putative hosts and niches.

3. Result and discussion

3.1. Microbial diversity and MAGs

We collected and sequenced sediments from nine different sites in Advent fjord in duplicates generating a total of ~450 GB of sequences (SRA accession: SRR28980414–31). The sediment microbiota was dominated by *Psuedomonadota* (45–55 % of the total 16S rDNA reads), followed by *Bacteroidota* (10–21 %). Around 9 % of the 16S rDNA reads could not be assigned to any phylum and around 8 % could not be classified beyond phylum level, thus representing a huge diversity of novel bacteria present in the Arctic sediments (Fig. 1a). We did observe variation in microbiota between the sampling sites suggesting micro niches present in high Arctic Fjord sediments. To characterize the novel bacteria, we assembled 644 mid-quality MAGs. The taxonomy of these 644 MAGs was consistent with the 16S rDNA gene-based microbiota analysis, with *Psuedomonadota* (50 %) representing most of the genomes followed by *Bacteroidota* (17 %). Five of the 644 genomes belonged to Archaea with one genome representing a putative novel phylum (Supplementary Table 1). Only 23 MAGs could be classified to either genus level (n = 16) or species level (n = 7), thus demonstrating the high abundance of novel taxa in the Arctic sediments (Fig. 1b). Around 14 % (n = 89) of MAGs were classified into seven putative novel phyla. (Fig. 1b, Supplementary Fig. 2). The metabolic functions deciphered from these genomes show that they perform active roles in nutrient cycling and methane metabolism in the Arctic sediments and thus represent prominent members of the Arctic microbiota (Supplementary Table 2).

3.2. Resistome of sediment microbiota

To understand the anthropogenic impact on the Arctic microbiota we analyzed the abundance of known antibiotic resistance genes (ARGs). We detected a total of 888 different ARGs in these samples encompassing clinically important genes conferring resistance against last resort antibiotics, including carbapenems NDM, VIM, IMP, KPC and OXA-181, mobile colistin resistance genes *mcr-1* and *mcr-2*, tigecycline resistance gene *tet*(X) as well as linezolid resistance genes *poxtA, cfr and optrA* (Supplementary Table 3). Thus, suggesting the presence of human associated bacteria, most likely pathogens, in the sediment microbiota owing to continuous wastewater discharge into the fjord (Marathe et al., 2019; Zhang et al., 2021). Most of the clinically relevant ARGs have originated in environmental bacteria, they serve as a source for acquiring new ARGs for pathogens (Forsberg et al., 2014). The presence of human associated bacteria in the Arctic coupled with climate change, makes the genetic exchange between Arctic microbiota and human pathogens a likely scenario (Edwards et al., 2020). With the high abundance of novel microbes present in the Arctic sediments, we investigated novel β-lactamases in the sediment microbiota using computational models. We discovered 478 unique novel β-lactamases belonging to class A (n = 52), subclass B1/B2 (n = 79), subclass B3 (n = 115) and class D (n = 232), respectively (Supplementary Table 4, Supplementary Fig. 1). Using gene sequences from β-lactamase database (BLDB) and previously published novel β-lactamases, we classified the genes into gene families with cutoff of amino acid identity of 70 % as suggested by Berglund et al. (2019). We report a total of 217 novel β-lactamase families with 27, 49, 45 and 96 novel families belonging to class A, subclass B1/2, subclass B3 and class D respectively (Supplementary Table 4, Supplementary Fig. 3 a-d). Our study thus almost doubles the number of β-lactamase families present in BLDB. Class B (subclass B1, B2 and B3) are particularly clinically important and represent carbapenem-hydrolyzing metallo-β-lactamases (MBLs) that utilize zinc in the active site (Berglund et al., 2019). The presence of 94 novel families of MBLs in Arctic microbiota thus, suggests the potential of Arctic microbiota as a source of clinically important ARGs for pathogens.

3.3. Hosts and distribution of NSFs in Arctic samples

Consistent with differences in microbial community composition we detected differences in the number of novel ARGs assembled from 9 sites in Advent fjord. Although, most of the gene families were detected, no full-length β-lactamase was assembled from sampling site 7. All 217 novel gene families were present with varying abundances in the sediments combined, with some families found to be absent in some of the samples (Supplementary Fig. 4).

Hosts for 58 NSFs were identified using NCBI nR database with 16 families in class A, 10 in subclass B1/2, 11 in B3 and 21 in class D, while MAGs carried a total of 23 novel families, 12 families overlapped between nR database and MAGs. We have thus successfully identified host species for 69 NSFs (Fig. 2, Supplementary Table 5a-c). Based on metadata all the host organisms represent marine bacteria, with majority of novel gene families detected in the phylum *Bacteroidota* (34) followed by *Pseudomonadota* (28). Thus, supporting the notion that less-abundant microbes may play a dominant role in the different ecological functions in each environment (Jousset et al., 2017).

To understand the presence and abundance of novel NSFs in different niches from Svalbard, we used previously published metagenomes from sediments (Zhang et al., 2022), seawater (Zhang et al., 2022) and permafrost (Xue et al., 2020) samples. A total of 168 gene families were

detected in sediments from Kongs fjord, while only 65 from permafrost and 46 from seawater metagenomes respectively (Supplementary Fig. 5). Our data suggests that many of these novel β-lactamases are prevalent in high Arctic niches, while some may be more niche specific.

3.4. Limitations of the study

The study is focused on sediments from Advent fjord in Svalbard. Advent fjord is well documented to be polluted from anthropogenic inputs like untreated wastewater and tourism (Kalinowska et al., 2020; Granberg et al., 2017). While this study provides valuable insights into the impact of wastewater pollution in this specific region, the outcome of this research cannot be generalizable to other Arctic regions or environments without further research. The presence of the novel bacteria and β-lactamases in these Arctic sediments may not directly translate to other environments, and further studies would be needed to understand their global implications. The data presented is from a specific time point, which may not capture seasonal or long-term variations in the microbiota and resistome. Follow-up longitudinal studies may provide a more dynamic understanding of the impacts and dynamics of clinically important resistance markers in the Arctics. To the best of our knowledge this is the first comprehensive study from this fjord addressing novel microbes and novel antimicrobial resistance genes in the high Arctic fjords impacted by wastewater. This study will serve as a baseline study from the area for future studies focusing on pollution and its impact on sediment microbiota.

4. Conclusion

Our study provides the most comprehensive analyses of novel microbes and novel β-lactamases present in the Arctic fjord sediments. It demonstrates the effect of mixing novel indigenous microbes and human associated microbiota in the Arctic sediments and further demonstrates presence of clinically important ARGs and novel ARGs in the Arctic microbiota. With a warming Arctic and continuing anthropogenic impact, Arctic microbiota may serve as a source for emergence of superbug, thus challenging public health in the future.

Supplementary data to this article can be found online at https://doi.

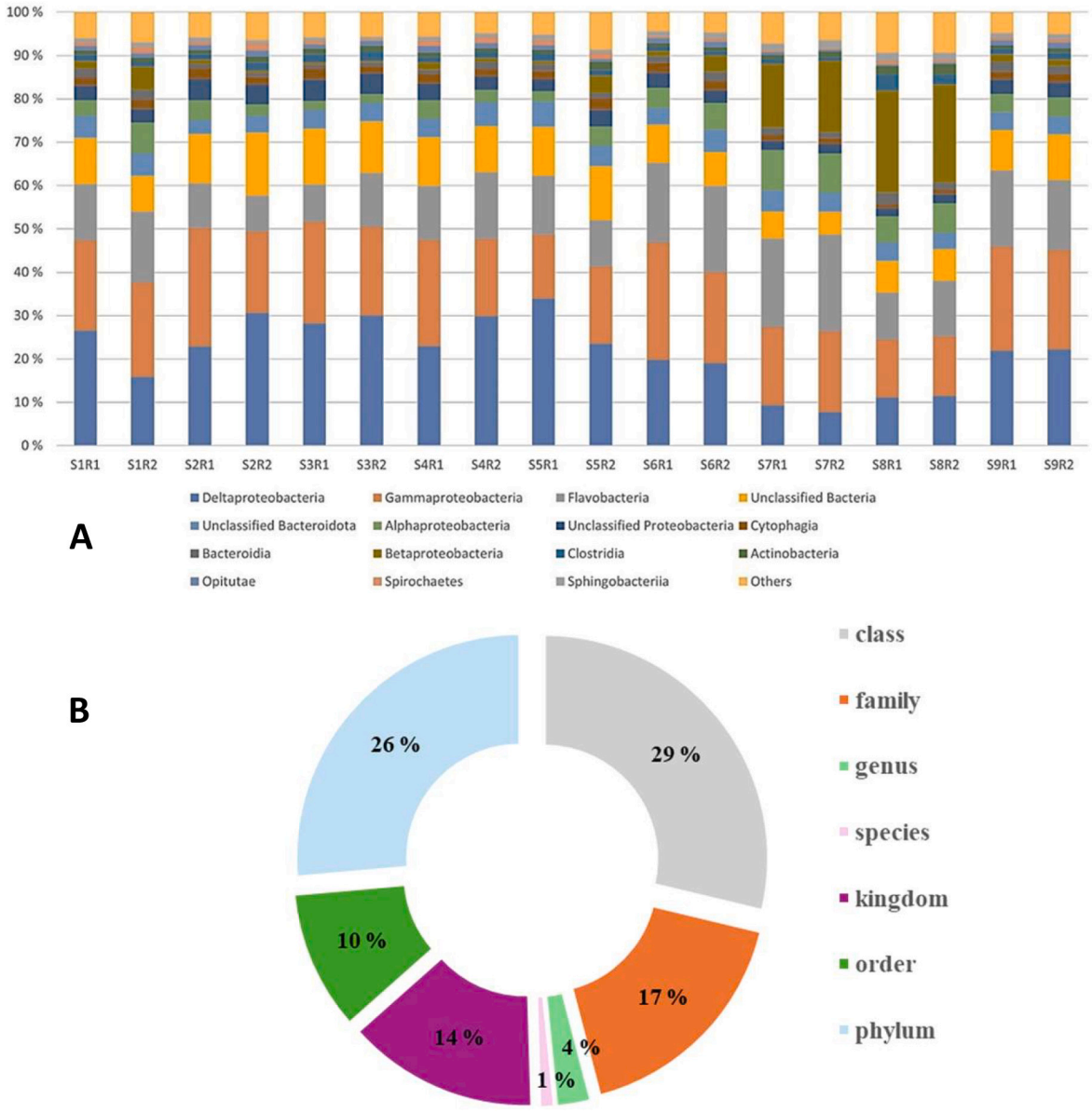


Fig. 1. a) Relative abundance of different phyla in the sediment samples. Legend: S- sampling sites and R-replicates eg. S1R1- site 1 replicate 1. b) Taxonomic classification of 639 MAGs (excluding Archea) at the lowest taxonomic unit possible.

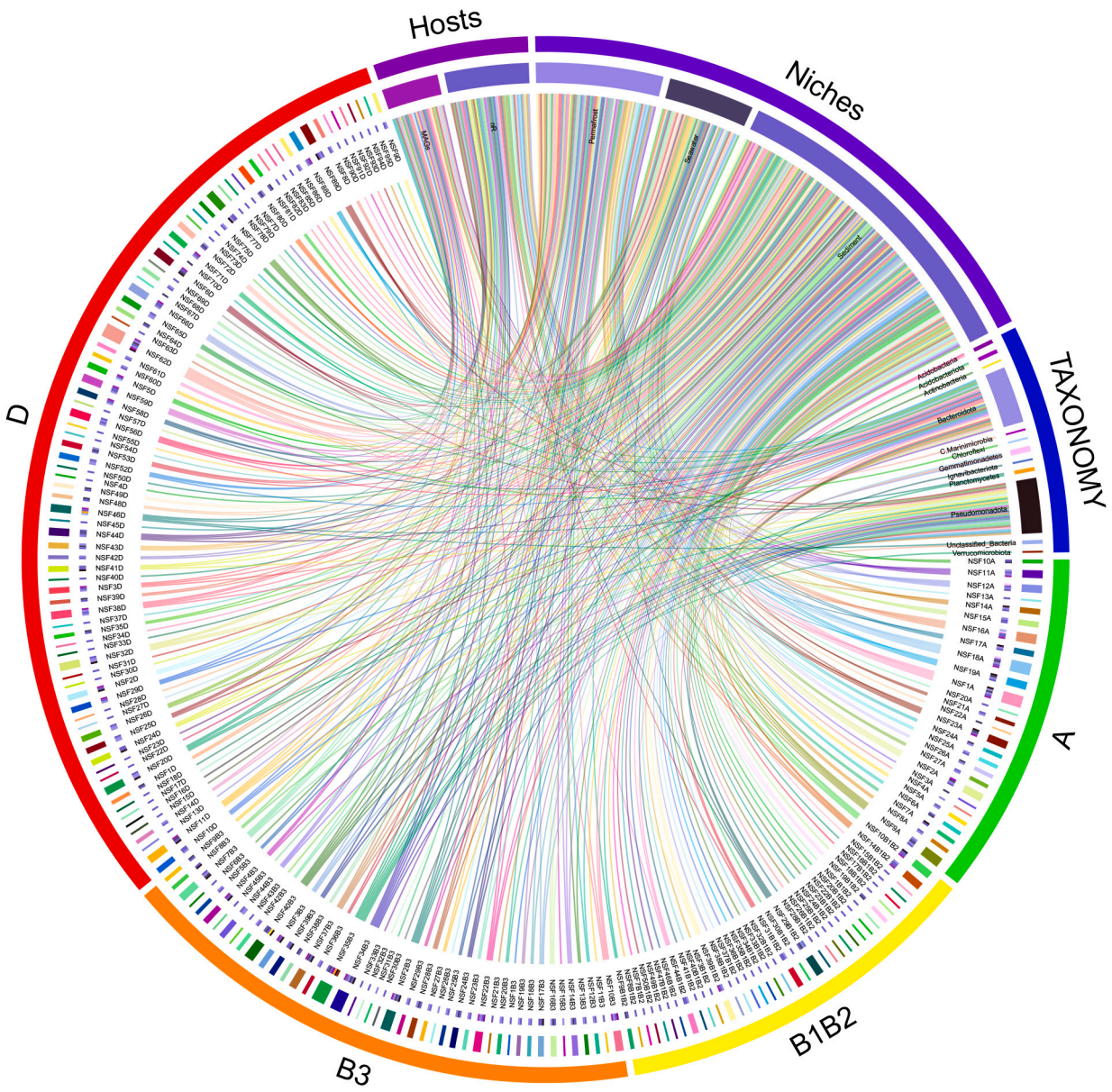


Fig. 2. Hosts and niches of novel β-lactamase families detected in our study. NSF denotes novel β-lactamase families from Svalbard. The figure represents the presence of β-lactamases across niches (sediment, seawater, permafrost samples) as well as taxonomy of the host species and the source (nR database or MAGs). Links are drawn when a certain gene is found in a certain sample and/or in certain host species representing either MAGs or nR database. Different classes of β-lactamases are represented as A, B1B2, B3, D.

org/10.1016/j.scitotenv.2025.179699.

CRediT authorship contribution statement

Manish P. Victor: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation. **Lise Øvreås:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Resources, Investigation, Formal analysis, Data curation. **Nachiket P. Marathe:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge Institute of Marine research (project number 15930) and the Research council of Norway for funding (Res-Marine project grant number 315266) for funding this work. We thank Nadja Junghardt at IMR and University of Bergen for her help with sample collection. The sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the South-Eastern Regional Health Authorities.

Data availabilityThe short reads are deposited at NCBI SRA under accession numbers SRR28980414–31. The amino acid sequences for assembled beta lactamases are presented in Supplementary Table 4.

The short reads are deposited at NCBI SRA under accession numbers SRR28980414–31. The amino acid sequences for assembled beta lactamases are presented in Supplementary Table 4.

References

Bengtsson-Palme, J., et al., 2015. METAXA2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. Mol. Ecol. Resour. 15 (6), 1403–1414. https://doi.org/10.1111/1755-0998.12399.

Benoit, G., Raguideau, S., James, R., Phillippy, A.M., Chikhi, R., Quince, C., 2024. High-quality metagenome assembly from long accurate reads with metaMDBG. Nat. Biotechnol. https://doi.org/10.1038/s41587-023-01983-6.

Berglund, F., Österlund, T., Boulund, F., Marathe, N.P., Larsson, D.G.J., Kristiansson, E., 2019. Identification and reconstruction of novel antibiotic resistance genes from metagenomes. Microbiome 7 (1), 52. https://doi.org/10.1186/s40168-019-0670-1.

Berglund, F., et al., 2017. Identification of 76 novel B1 metallo-β-lactamases through large-scale screening of genomic and metagenomic data. Microbiome 5 (1), 134. https://doi.org/10.1186/s40168-017-0353-8.

Buchfink, B., Reuter, K., Drost, H.-G., 2021. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nat. Methods 18 (4), 366–368. https://doi.org/10.1038/s41592-021-01101-x.

Cavicchioli, R., et al., 2019. Scientists' warning to humanity: microorganisms and climate change. Nat. Rev. Microbiol. 17 (9), 569–586. https://doi.org/10.1038/s41579-019-0222-5.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., Jin, Q., 2005. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 33, D325–D328. https://doi.org/10.1093/nar/gki008.

Dereeper, A., et al., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 36. https://doi.org/10.1093/nar/gkn180 no. Web Server, pp. W465–W469.

D'Onorio De Meo, P., et al., 2012. MitoZoa 2.0: a database resource and search tools for comparative and evolutionary analyses of mitochondrial genomes in Metazoa. Nucleic Acids Res. 40 (D1), D1168–D1172. https://doi.org/10.1093/nar/gkr1144.

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

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26 (19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461.

Edwards, A., Cameron, K.A., Cook, J.M., Debbonaire, A.R., Furness, E., Hay, M.C., Rassner, S.M.E., 2020. Microbial genomics amidst the arctic crisis. Microb. Genom. 6 (5), e000375. https://doi.org/10.1099/mgen.0.000375.

Feldgarden, M., et al., 2021. AMRFinderPlus and the reference gene catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. Sci. Rep. 11 (1), 12728. https://doi.org/10.1038/s41598-021-91456-0.

Forsberg, K.J., et al., 2014. Bacterial phylogeny structures soil resistomes across habitats. Nature 509 (7502), 612–616. https://doi.org/10.1038/nature13377.

Granberg, M.E., Ask, A., Gabrielsen, G.W., 2017. Local contamination in Svalbard: overview and suggestions for remediation actions. Norsk Polarinstitutt. Brief Report no. 44 Accessed: May 09, 2025. [Online]. Available: https://brage.npolar.no/npolar-xmlui/handle/11250/2456307.

Gu, Z., Gu, L., Eils, R., Schlesner, M., Brors, B., 2014. Circlize implements and enhances circular visualization in R. Bioinformatics 30 (19), 2811–2812. https://doi.org/10.1093/bioinformatics/btu393.

Gupta, S.K., et al., 2014. ARG-ANNOT, a new Bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob. Agents Chemother. 58 (1), 212–220. https://doi.org/10.1128/AAC.01310-13.

Jia, B., et al., 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. Nucleic Acids Res. 45 (D1), D566–D573. https://doi.org/10.1093/nar/gkw1004.

Jousset, A., Bienhold, C., Chatzinotas, A., et al., 2017. Where less may be more: how the rare biosphere pulls ecosystems strings. ISME J 11, 853–862. https://doi.org/10.1038/ismej.2016.174.

Kalinowska, A., Szopińska, Małgorzata, Chmiel, Stanisław, Kończak, Magdalena, Polkowska, Żaneta, Artichowicz, Wojciech, Jankowska, Katarzyna, Nowak, Aga, Łuczkiewicz, Aneta, 2020. Heavy metals in a high arctic fiord and their introduction with the wastewater: a case study of Adventfjorden-Longyearbyen system, Svalbard. Water 12 (3), 794. https://doi.org/10.3390/w12030794.

Kang, D.D., et al., 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ 7, e7359. https://doi.org/10.7717/peerj.7359.

Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30 (14), 3059–3066. https://doi.org/10.1093/nar/gkf436.

Kleinteich, J., et al., 2017. Pole-to-pole connections: similarities between Arctic and Antarctic microbiomes and their vulnerability to environmental change. Front. Ecol. Evol. 5. https://doi.org/10.3389/fevo.2017.00137.

Li, D., Liu, C.-M., Luo, R., Sadakane, K., Lam, T.-W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct *de Bruijn* graph. Bioinformatics 31 (10), 1674–1676. https://doi.org/10.1093/bioinformatics/btv033.

Marathe, N.P., et al., 2019. Sewage effluent from an Indian hospital harbors novel carbapenemases and integron-borne antibiotic resistance genes. Microbiome 7 (1), 97. https://doi.org/10.1186/s40168-019-0710-x.

Na, S.-I., Kim, Y.O., Yoon, S.-H., Ha, S., Baek, I., Chun, J., 2018. UBCG: up-to-date bacterial core gene set and pipeline for phylogenomic tree reconstruction. J. Microbiol. 56 (4), 280–285. https://doi.org/10.1007/s12275-018-8014-6.

Naas, T., et al., 2017. Beta-lactamase database (BLDB) – structure and function. J. Enzyme Inhib. Med. Chem. 32 (1), 917–919. https://doi.org/10.1080/14756366.2017.1344235.

Pal, C., Bengtsson-Palme, J., Rensing, C., Kristiansson, E., Larsson, D.G.J., 2014. BacMet: antibacterial biocide and metal resistance genes database. Nucleic Acids Res. 42 (D1), D737–D743. https://doi.org/10.1093/nar/gkt1252.

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 Res. 25 (7), 1043–1055. https://doi.org/10.1101/gr.186072.114.

Pärnänen, K., et al., 2018. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. Nat. Commun. 9 (1), 3891. https://doi.org/10.1038/s41467-018-06393-w.

Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2 – approximately maximum-likelihood trees for large alignments. PLoS One 5 (3), e9490. https://doi.org/10.1371/journal.pone.0009490.

Quast, C., et al., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41 (D1), D590–D596. https://doi.org/10.1093/nar/gks1219.

Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27 (6), 863–864. https://doi.org/10.1093/bioinformatics/btr026.

Seemann, Torsten, 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30 (14), 2068–2069.

Sharifi, F. and Ye, Y. 'From Gene Annotation to Function Prediction for Metagenomics'. Kihara, D. *Protein Function Prediction*, vol. 1611, Springer New York, New York, NY, 2017, pp. 27–34. [Online]. Available: http://link.springer.com/10.1007/978-1-4939-7015-5_3.

Tamames, J., Puente-Sánchez, F., 2019. SqueezeMeta, a highly portable, fully automatic metagenomic analysis pipeline. Front. Microbiol. 9, 3349. https://doi.org/10.3389/fmicb.2018.03349.

Verde, C., Giordano, D., Bellas, C. M., Di Prisco, G., and Anesio, A. M. 'Polar marine microorganisms and climate change', in Advances in microbial physiology, vol. 69, Elsevier, 2016, pp. 187–215. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0065291116300236.

Wang, Q., Cole, J.R., 2024. Updated RDP taxonomy and RDP classifier for more accurate taxonomic classification. Microbiol. Resour. Announc. 13 (4). https://doi.org/10.1128/mra.01063-23 e01063 23.

Xue, Y., Jonassen, I., Øvreås, L., Taş, N., 2020. Metagenome-assembled genome distribution and key functionality highlight importance of aerobic metabolism in Svalbard permafrost. FEMS Microbiol. Ecol. 96 (5), fiaa057. https://doi.org/10.1093/femsec/fiaa057.

Zhang, A.-N., et al., 2021. An omics-based framework for assessing the health risk of antimicrobial resistance genes. Nat. Commun. 12 (1), 4765. https://doi.org/10.1038/s41467-021-25096-3.

Zhang, T., Li, J., Wang, N., Wang, H., Yu, L., 2022. Metagenomic analysis reveals microbiome and resistome in the seawater and sediments of Kongsfjorden (Svalbard, high Arctic). Sci. Total Environ. 809, 151937. https://doi.org/10.1016/j.scitotenv.2021.151937.