Mapping the microbial diversity and natural resistome of north Antarctica soils

Andrés E. Marcoleta[1,*,+], Macarena A Varas[1,+], José Costa[1], Johanna Rojas-Salgado[1], Patricio Arros[1], Camilo Berríos-Pastén[1], Sofía Tapia, Daniel Silva, José Fierro[1], Nicolás Canales[1], Francisco P Chávez[2], Alexis Gaete[3], Mauricio González[3], Miguel A. Allende[4], Rosalba Lagos[1]*.

[1]Grupo de Microbiología Integrativa, Laboratorio de Biología Estructural y Molecular BEM, Departamento de Biología, Facultad de Ciencias, Universidad de Chile. Santiago, Chile.

[2]Laboratorio de Microbiología de Sistemas, Departamento de Biología, Facultad de Ciencias, Universidad de Chile. Santiago, Chile.

3 Laboratorio de Bioinformática y Expresión Génica, Instituto de Nutrición y Tecnología de los Alimentos, Universidad de Chile, Santiago, Chile.

4 FONDAP Center for Genome Regulation, Facultad de Ciencias, Universidad de Chile. Santiago, Chile.

*Corresponding authors
[+]These authors contributed equally to this work.

Running title: Antarctic Peninsula soils natural resistome

Keywords: Antimicrobial resistance, natural resistome, Antarctic Peninsula, mobile genetic elements, soil microbial communities, anthropogenic intervention

ABSTRACT

The rising of multi-resistant bacterial pathogens is currently one of the most critical threats to global health and requires urgent actions for its control and a better understanding of the origin, evolution, and spread of antibiotic resistance. It is of consensus that this must be faced under a one-health prism that includes the environment and non-human animals as possible sources and media for resistance amplification and evolution. In this regard, the resistome extension and diversity present in natural and remote environments remain largely unexplored. Moreover, little is known about the availability of antimicrobial resistance genes (ARGs) from these environments to be disseminated through horizontal transfer, potentially mediating the rise of new resistance mechanisms among clinically relevant pathogenic microorganisms. Furthermore, the impact of the anthropogenic intervention on the presence of antibiotic resistant bacteria and the underlying genes is still a matter of controversy. In this context, the Antarctic Peninsula soils are attractive remote environments to study, since their high microbial diversity, the presence of ice-covered soils sheltering ancestral microorganisms that are being exposed as cause of the global warming, and because it is one of the most transited routes between Antarctica and the rest of the world, thus permitting genetic and microbial carriage among those places. Also, it harbors both human bases and places without noticeable human intervention. In this work, we explored the culturable resistome of soils from different places of North Antarctica, including human bases and non-intervened areas. We identified bacterial isolates resistant to a wide array of antibiotics, harboring up to 10 simultaneous resistances, most of them belonging to the genus *Pseudomonas*. Genomic analyses of the two top resistant bacteria revealed a dominant presence of efflux pumps although an unexpectedly low abundance of known

resistance genes, suggesting the presence of unidentified new mechanisms. Moreover, using 16S rRNA amplicon and metagenomic sequencing we explored the microbial diversity in the sampled soils and evaluated the presence and abundance of antimicrobial resistance genes. Proteobacteria, Bacteroidota, Acidobacteriota, and Verrucomicrobiota corresponded to the most abundant Phyla in most of the soils, while among the most abundant genera were *Candidatus Udaeobacter*, RB41, *Polaromonas*, and *Ferruginibacter*. A similarly high microbial diversity was observed when comparing humanized with non-intervened sites, although beta diversity analysis, as well as sequence composition analysis revealed significant clustering of non-intervened apart from humanized areas. We identified a variety of genes potentially involved in resistance to more than 15 drug classes in both short reads based analyses and ARGs detection among assembled contigs. Further, through performing hybrid assembly by combining short and long read sequence data we searched for resistance genes located inside possible mobile genetic elements and identified the source taxa. *Polaromonas*, *Pseudomonas*, *Streptomyces*, *Variovorax*, *Bhurkolderia, and Gemmatimonas* were the host taxa of most of the identified ARGs. Plasmid prediction among the assembled metagenomes led to the identification of a putative OXA-like beta-lactamase from *Polaromonas* spp. located inside a putative plasmid, which included all the key conserved residues for the activity of this kind of beta-lactamases. Taken together, this evidence indicates that resident North Antarctica soil microbial communities harbor a highly diverse natural resistome, part of which is located inside mobile genetic elements that could act disseminating these ARGs to other bacteria.

INTRODUCTION

Currently, antimicrobial therapies' effectiveness is rapidly declining worldwide, mainly because of the increasing detection of multi-resistant bacterial strains. This situation could lead us back to the pre-antibiotic era, where bacterial diseases represented one of the leading causes of death globally (Friedman et al., 2016). Antimicrobial resistance is now considered among the most pressing health crises of the 21st Century, demanding urgent actions to prevent the dissemination of resistant pathogens and encouraging the research to understand the origin, evolution, and spread of antibiotic resistance (Hutchings et al., 2019; Smith et al., 2015). As with other global health issues, this resistance crisis should be faced following the so-called "One health" directive, which stresses human health's interconnected nature with animal health and the environment. All these three entities play critical roles in antibiotic resistance, being of utmost importance to uncover key interconnections among them.

Antibiotic resistance should be studied in the environment, not only in populated areas or in environments subjected to anthropogenic contamination but also in the wild. In this regard, recent reports have evidenced the presence, in remote areas, of antibiotic resistance genes typically detected among bacterial pathogens infecting humans. This is the case of the bla_{NDM-1} carbapenemase, commonly associated with multi-resistant *Enterobacteriaceae* causing outbreaks, which gene was detected by quantitative PCR among soil samples collected from the High Arctic (northern Norway). Here, a gradient of potential human and wildlife impacts across zones correlated positively with the presence of this and other foreign resistance genes (McCann et al., 2019). A similar observation was made by Tan et al., studying Arctic marine sediments collected from the Bering Sea, which found significant correlations between the copy number of resistance genes and the human mitochondrial gene Hmt as a marker of human contamination, as well as with the concentration of different antibiotics found in the sediments. This suggests a concerning spread of resistance genes from pathogens contaminating diverse natural and remote environments, possibly through bacteria's carriage in migrating birds and animals connecting these environments with humanized areas. Also, it points out the potential influence of anthropogenic antibiotic release on the presence of antibiotic-resistant bacteria in the environment.

On the other hand, it was also documented the presence of potential resistance genes from autochthonous bacteria living in the wild, even in remote areas (Goethem et al., 2018; McCann et al., 2019; Tan et al., 2018). Indeed, the evidence suggests that the development of antibiotic resistance is a natural phenomenon occurring widespread on our planet, which as for other metabolic features, developed through billions of years of evolution (Blair et al., 2015; Forsberg et al., 2012; Hutchings et al., 2019). Most clinically used antibiotics derive from antibiotic-producing microorganisms that naturally expose other species to these substances in their local environment. This would favor the development and selection of resistance in both the affected environmental species and the producer organisms. Also, proteins such as efflux pumps, allowing environmental bacteria to deal with toxic compounds produced on their niche, can be leveraged by pathogens to resist clinically used antibiotics sharing structural or chemical similarity with these natural compounds. Thus, a wide diversity of genes from microbes living in natural environments, conferring resistance to potentially all the known antibiotics, is yet to be characterized. However, there is limited information regarding the mobility of these genes and possible dissemination mechanisms to reach pathogenic bacteria. In particular, little is known regarding the natural resistome found in extreme environments, where the resident microbiota is known to develop remarkable adaptations to thrive with harsh conditions, including toxic substances.

In this work, we used the bacterial communities living in the soil from different north Antarctica locations as model ecosystems to study the presence of antibiotic-resistant bacteria and resistance genes, comparing areas intervened by human activities with protected areas of restricted access without evident human intervention. We tested the resistance to a variety of antibiotics of over 200 Antarctic isolates, identifying various multi-resistant strains and performing genome analysis of two of them, searching for the genetic basis of the resistance phenotype, which suggested the presence of new resistance mechanisms. Furthermore, we conducted metagenomic approaches combining illumina and nanopore sequencing to evaluate the abundance of resistance genes in humanized and non-intervened zones, determine which bacteria host these genes, and search for resistance genes mobile genetic elements genomic islands, transposons and plasmids.

METHODS

Antarctic soil sample collection and bacterial isolation

Soil samples were collected during the 53th and 55th Chilean Antarctic Scientific Expedition, during January and February 2017 and 2019. Samples were collected in triplicates introducing 500 g of soil into Whirl-Pak sterile bags using an aseptic spade. Sampled zones covered the following islands: King George, Greenwich, Barrientos, Doumer, Deception, and Robert I, and included humanized areas (Gabriel de Castilla Base, Luis Risopatron Refugee, Yelcho Base, Hennryk Arctowski Base, and Prat Base), as well as areas devoid of evident human intervention (mainly protected zones) (Fumarola Bay, Coppermine Peninsula, South Bay, Ecology Glacier, Air Force Glacier, and Barrientos Island). More details can be found in Supplementary Table 1. Soil bacteria were isolated by mixing 10 g of soil with 10-15 mL sterile PBS buffer, shaking in vortex during 5 min and then let stand to allow the sedimentation of higher solids. Then, 1 ml of the supernatant was centrifuged at 1500 rpm for 1 min and 100 µl of the supernatant was plated on agar plates with growth media supplemented with 100 µg/mL Cycloheximide. We used a group of different culture media to increase the recovered diversity, including: Marine agar, Casein Starch agar, Luria-Bertani (LB) agar, Tripctic Soy agar, Streptomyces Agar, and Actinomycete agar. Upon inoculation, the plates were incubated at 10-15ºC during approximately 3 weeks or until noticeable colonies formed. The obtained colonies were named and classified according the site of sampling, Gram-staining results, and the morphology of the colonies and cells.

Antimicrobial sensitivity testing

To carry out antibiotic sensitivity tests with Antarctic bacteria, we adapted the disk diffusion method recommended by the European Committee for testing Antimicrobial Susceptibility (EUCAST, 2017). First, cultures of the strains to be tested were prepared in 2 mL of Mueller-Hinton broth (MH) and growth at 10° C with shaking for 48 h or until reaching the desired turbidity. Subsequently, the $OD^{625\ nm}$ of each culture was adjusted to 0.3 diluting with the same MH medium to obtain 2 mL of the adjusted culture. The adjusted culture was torula plated on MH agar to form a lawn. Subsequently, antibiotics discs (Oxoid) were applied on the inoculated plates using a disk dispenser (Oxoid). The plates were incubated at 10 °C for one week, after which the presence of growth inhibition halo was evaluated. The following disks were used: Amikacin 30 µg, Cyprofloxacin 5 µg, Ceftaroline 5 µg, Colistin 10 µg, Clindamycin 2 µg, Erythromycin 15 µg, Cefepime 30 µg, Fosfomycin/Trometramol 200 µg, Linezolid 30 µg, Meropenem 10 µg, Rifampicin 5 µg, Trimethoprim/sulfamethoxazole 25 µg, Tigecycline 15 µg, Piperacillin/Tazobactam 110 µg, and Vancomycin 30 µg. A strain was

considered sensitive when a growth inhibition halo with a diameter equal to or greater than 10 mm was recorded.

16S rRNA-based taxonomic identification of isolates

Taxonomic classification of selected isolates was performed by PCR amplifying the full-lenght 16S rRNA gene using the primers 16SF (5'AGA GTT TGA TCC TGG CTC AG) and 16SR (5' ACG GCT ACC TTG TTA CGA CTT). The purified amplicons were subjected to Sanger sequencing hiring the services of Macrogen Inc. (Korea) and the resulting sequence was compared to the NCBI 16S rRNA sequence database.

Whole genome sequencing and analysis of Antarctic Isolates

Genomic DNA was extracted from selected isolates using the GenJet genomic DNA purification kit (ThermoFisher), following the manufacturer's manual. DNA integrity and purity was asessed by agarose gel electrophoresis and measuring the ratio of absorbance at 260 and 280 nm. Prior to sequencing, DNA was quantified using a Qubit fluorimeter (ThermoFisher). To obtain complete genomes with closed chromosome and plasmids, we sequenced each genome using both illumina and Oxford Nanopore technologies. illumina sequencing was performed in a HiSeq 2500 platform using the Truseq DNA PCR Free library Kit (350 bp), at >50X coverage (Macrogen Inc., Korea). Nanopore sequencing was performed at Laboratorio de Expresión Génica (INTA, Santiago, Chile) using a MinION device (Nanopore, Oxford, UK) and with FLO-MIN106 (R9.4) flow cells, according to the manufacturer's instructions. Sequencing libraries were prepared from 1 µg of gDNA using the 1D Genomic DNA by ligation kit (SQK-LSK108). Base-calling was performed using Guppy software (Nanopore, Oxford, UK). Genome assembly was performed using CANU (Koren et al., 2017) and Unicycler (Wick et al., 2017). The quality of the illumina reads was assessed using FASTQC (Andrews et al., 2012), and filtered using Trimommatic (Bolger et al., 2014). Adaptor sequences were removed from Nanopore reads using the tool Porechop (https://github.com/rrwick/Porechop). Genome annotation was performed using Prokka (Seemann, 2014). Antibiotic resistance gene identification was performed using the RGI-tool of the Comprehensive Antibiotic Resistance Database (Jia et al., 2017). Genome-based taxonomic assignment was performed using the rMLST (Ribosomal Multilocus Sequence Typing) analysis tool (Jolley et al., 2012) available at PubMLST.org (Jolley & Maiden, 2010) and the GTDB-Tk tool (Chaumeil et al., 2019). The phylogenomic tree showing the relationships between the Antarctic multi-resistant isolates ArH3a and YeP6b, with other *Pseudomonas* was built calculating the Mash distances among the whole chromosomes Mash distances using Mashtree (Katz et al., 2019) with 1,000 bootstrap iterations. The accession numbers and additional information about the *Pseudomonas* genomes used are listed in Supplementary Table 2.

DNA extraction from Antarctic soil samples

Metagenomic DNA from soil samples was extracted using the DNeasy PowerSoil DNA isolation Kit (Qiagen), starting from 0.5-1 g of soil and following the manufacturer's guidelines. Alternatively, soil DNA was extracted using a custom protocol. Briefly, 10 g of soil was mixed with 50 mL of Soil Lysis Buffer pre-heated to 65ºC (7.1 g Na_2HPO_4, 43.8 g NaCl, 5 g CTAB, 50 mL 1M Tris pH=8.0, 100 mL 0.5 M EDTA, nanopure water to complete 1 L), and Proteinase K. The mix was vortexed 30s and incubated at 60ºC for 2 h, mixing in vortex every 30 min. Then, 200 µg RNAse A were added and the mix was incubated at room temperature during 45 min shaking in vortex every 15 min, then centrifuging at 4000 x g

during 2 min. The supernatant was transferred to a clean tube and then centrifuged 20 min at 11.000 x g. The supernatant was transferred to a clean tube and mixed with 1 vol. of 100% ethanol. The mixture was then sucesively loaded into a microfuge silica column, centrifuging during 2 min at 11000 x g and discarding the flow-through. Then, the DNA was eluted from the column using 100 µL of nanopure water, and then further purified using the reagents included in the DNeasy PowerSoil kit, mixing the DNA with 300 µL of the PowerBead buffer and performing the following steps recommended by the manufacturer. Metagenomic DNA integrity and purity was asessed by agarose gel electrophoresis and measuring the ratio of absorbance at 260 and 280 nm, respectively. Prior to sequencing, metagenomic DNA samples were quantified using a Qubit fluorimeter (ThermoFisher).

16S-rRNA amplicon sequencing and microbial diversity analyses

16S amplicon libraries were constructed and sequenced hiring the services of Macrogen Inc. (Korea). Upon quality control, metagenomic DNA extracted from soil samples was used along with the Herculase II Fusion DNA Polymerase Nextera XT Index Kit V2 (illumina), in order to amplify the V3-V4 region of the 16S rRNA gene using the primersBakt_341F: CCTACGGGNGGCWGCAG and Bakt_805R: GACTACHVGGGTATCTAATCC. The libraries were sequenced using an illumina Miseq sequencer, obtaining approximately 166,000-245,000 300 bp paired-end reads (50-73 Mbp) per sample. Microbial diversity analyses including relative abundance, and alpha- and beta-diversity calculations were performed using QIIME2 platform v. 2019.7.0 (Bolyen et al., 2019). The 16S rRNA amplicon reads were imported into the Qiime2 pipeline, performing quality assessment and filtering using q2-dada2. The result obtained was explored using featureTable function. A rooted phylogenetic tree required for alpha and beta diversity analyses was generated with the q2 -diversity plugin, performing PERMANOVA statistical analysis for beta diversity. On the other hand, the compositional analysis of beta diversity was carried out using the DEICODE complement. Finally, q2-feature-classifier was used to carry out the taxonomic assignment of the microorganisms present in the samples using the SILVA 138 database.

Shotgun metagenomic sequencing

For illumina shotgun metagenomic sequencing of samples collected at A. Prat base, Air Force glacier, Gabriel de Castilla base, and Barrientos island (during ECA53), the libraries were prepared from soil DNA using the Truseq DNA PCR Free library Kit (350 bp) and were sequenced in an illumina Hiseq2500 platform, obtaining a total of 4.9 to 6.7 Gbp per sample (100 bp paired-end reads). Additionally, the libraries from samples collected at H. Arctowski Base, Coppermine Peninsula, and Fumarola Bay (during ECA55) were prepared using the Truseq nano DNA library Kit (350bp), and were sequenced in an illumina Novaseq sequencer, obtaining a total of 11.8-12.7 Gbp per sample (150 bp paired-end reads). Nanopore sequencing libraries were prepared starting with 1 µg of soil DNA using the Ligation Sequencing kit LSK-109, following the manufacturer's guidelines. For each library, 5-50 fmol were loaded into a R9 flow cell and sequenced using a MinION device (Nanopore Technologies).

Metagenome assembly

Prior to assembly, illumina reads were quality-assessed using FastQC and trimmed (adaptors and low quality bases) using Trimmomatic (Bolger et al., 2014). The quality of Nanopore reads was evaluated using NanoPlot, while adaptor removal was performed using Porechop (https://github.com/rrwick/Porechop). Hybrid metagenome assembly was

performed combining pre-processed illumina and Nanopore data using metaSPAdes (Nurk et al., 2017). The quality of the obtained assemblies was assessed using metaQUAST (Mikheenko et al., 2016).

ARGs relative abundance calculations based on illumina reads

Antibiotic resistance gene identification starting from unassembled illumina data was performed using the tool deepARG (Arango-Argoty et al., 2018), recovering the "mapping.ARG" output file, which contains alignments that have a probability higher or equal to 0.8, which according to the authors, is reliable enough to be considered as a true positive alignment. To construct the sunburst chart showing the most abundant ARG in the sum of the sampled sites, all the ARG alignments from all sites were added in a single table. Then, ARGs with less than 1000 alignments were filtered out, and the number of alignments for each surviving ARG was normalized by the total number of alignments for all of the ARGs. For relative abundance normalization among different samples, the illumina reads from each metagenome were mapped to a custom *rpoB* protein sequence database (kindly provided by Dr. Luis Orellana) using DIAMOND BlastX, with cutoffs of 80% identity and 1E-10 e-value for positive alignments. ARG abundances for each metagenome were calculated considering the total amount of *rpoB* alignments as a normalizing factor, the known length of each ARG in the deepARG database, and the median length for *rpoB* sequences (4029 bp). The deepARG-abundance module from the genetools software was used for this task. For constructing the heatmaps showing the ARGs relative abundance among different sites, seriation distances were determined using the GW method and the seriation R library. The final Heatmap vector graphics were created using the ComplexHeatmap and circlize R libraries. Discriminant ARGs analysis was performed using the Extrarg tool, using as input a table comprising the known amount of raw ARG alignments in every metagenome and the corresponding normalized amount. This table also included a group designation (humanized or non-intervened). ARG categorization by resistance mechanism, drug class, and ARG gene family was performed by searching in the CARD database each of the matched genes from deepARG analysis. NMDS plots were created using the Vegan and ggplot2 R packages. Permanova analysis was performed using the pairwiseAdonis R package.

ARGs identification in metagenomic assemblies and mobile genetic elements prediction

ARG identification among assembled metagenomic contigs was performed using NanoARG (Arango-Argoty et al., 2019). Genetools was used to convert the .json output file into a table that includes information of all the hits (i.e., identity, coverage, e-value). To construct the treemap showing the most abundant ARGs detected in the sum of the assemblies, we concatenated the output tables obtained for each sample. Next, we examined the distribution of the percentage identity and e-value among all the hits to establish appropriate cutoffs, which allow minimizing false positives but permitting the detection of novel resistance genes with limited identity to those present in the database. Based on the observed data distribution, we established a 40% identity and 1E-10 e-value cutoff. The ARG relative abundance was calculated by dividing the number of hits obtained for each ARG by the total number of hits for all the ARGs. Taxonomic assignment of contigs carrying ARGs was performed using Centrifuge (Kim et al., 2016). Plasmid prediction among assembled metagenomic contigs was conducted using PlasFlow (Krawczyk et al., 2018) and ViralVerify (Antipov et al., 2020).

Metagenome-assembled genomes (MAGs) isolation and characterization

MAGs recovery was performed as follows. The hybrid metagenomic assembly generated using metaSPAdes, as well as the short and long read sets from each metagenome, were used as input for MetaWRAP (Uritskiy et al., 2018) in order to perform contig binning and initial bins refinement using a combination of MetaBAT2 (Kang et al., 2019), CONCOCT (Alneberg et al., 2014), and MaxBin2 (Wu et al., 2016) binning algorithms. Afterward, the contigs from the obtained bins were used as a reference to map the whole set of long and short reads using minimap2 v2.17 (Li, 2018) and Bowtie2 (Langmead & Salzberg, 2012), respectively. Upon recovering the mapping reads, they were combined and de novo assembled using Unicycler. The completeness and contamination of the MAGs was evaluated using CheckM (Parks et al., 2015). MAG taxonomic assignment was conducted using GTDB-Tk. Phylogenetic tree of MAGs was constructed based on Mash distances calculation using Mashtree. MAGs annotation was performed using Prokka, and the functional categorization of the predicted protein-coding genes was made using EggNOG mapper (Huerta-Cepas et al., 2017) v2.1.2 and the EggNOG database v.5 (Huerta-Cepas et al., 2019). ARG identification in MAGs was performed using RGI-CARD tool allowing loose hits. The phylogenetic tree and the heatmaps of functional categories were integrated using iTOL v.5 (Letunic & Bork, 2021).

RESULTS AND DISCUSSION

Antibiotic-resistant bacteria in Antarctic soil

We examined the presence of antibiotic-resistant bacteria in soil samples collected from different locations of north Antarctica during the 53rd and 55th Chilean Scientific Antarctic Expedition (January-February 2017 and 2019, respectively), covering areas intervened by human activities (i.e. the surroundings of human bases) as well as restricted areas devoid of evident human intervention (Figure 1A, Supplementary Table 1). Using different culture media and growth temperatures to encourage diversity, we obtained 229 different bacterial isolates from a total of eight areas, which were named and classified based on their zone of origin, Gram staining, and the morphology of the cells and colonies. Then, we used adapted disk diffusion assays to test the isolates' sensitivity to fifteen clinically used antibiotics of diverse chemical nature and mechanism of action. We found at least one resistant isolate to 13 out of 15 antibiotics tested (Figure 1B), while meropenem and tigecycline were able to kill all the isolates. A high proportion of resistant isolates was observed for Fosfomycin and Colistin, followed by Clindamycin and Vancomycin.

Next, we assessed if the resistance to each antibiotic could be preferentially associated with isolates obtained from either humanized or non-intervened areas, looking for possible enrichment of resistant bacteria due to anthropogenic activities. Upon constructing contingency tables from the proportion of resistant isolates coming from each kind of area, we calculated the odds ratio representing the fold-change probability of finding resistance to a defined antibiotic in a soil sample coming from non-intervened areas compared to humanized areas. We found significant associations for eight of the fifteen antibiotics tested ($p<0.05$, two-tailed Fischer's exact test) (Table 1). Remarkably, in seven of them, the resistance had 3- to 6-fold increased probability among non-intervened areas, arguing against the hypothesis of increased resistance among soil bacteria due to anthropogenic contamination. The only exception was ciprofloxacin, with a 10-fold increased probability of resistance in humanized areas.

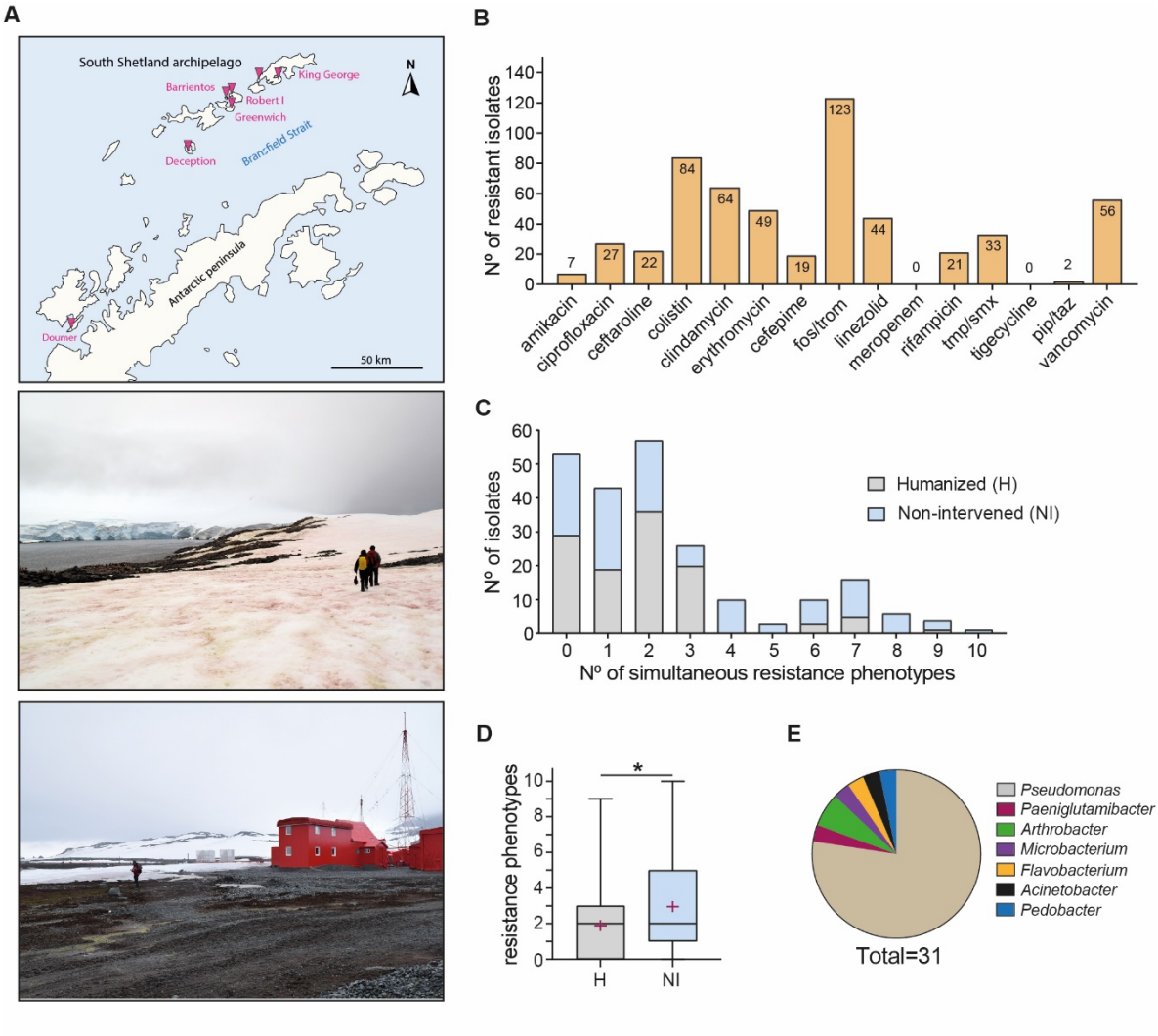

Figure 1. Isolation of antibiotic resistant bacteria from North Antarctica soils. (A) Soil samples were collected at different locations of North Antarctica (pink triangles), including areas devoid of noticeable human intervention (upper picture) and areas harbouring human settlements (lower picture). (B) Antibiotic sensitivity testing of Antarctic bacterial isolates using adapted disk fifussion assays. (C) Recount of isolates showing simultaneous resistance to one or more antibiotics, both from humanized or non-intervened areas. (D) Average number of simpultaneos resistances displayed by Antarctic isolates coming either from humanized or non-intervened areas. (E) Taxonomic classification of selected Antarctic isolates showing resistance to 7 or more antibiotics, or showing resistance to antibiotics that killed most of the isolates (e.g. piperacillin/tazobactam).

Table 1. Comparison of antibiotic resistance among soil bacteria isolated from either humanized or non-intervened areas from north Antarctica.

Antibiotic	Non-intervened (n=116)		Humanized (n=113)		Statistics[1]		
	n	%	n	%	p-value	OR	95% CI
Amikacin	6	5.2	1	0.9	0.1195	-	-
Ciprofloxacin	3	2.6	24	21.2	<0.0001	**0.10**	0.03 - 0.33
Ceftaroline	17	14.7	5	4.4	0.0123	**3.71**	1.32 - 9.42
Colistin	38	32.8	46	40.7	0.2205	-	-
Clindamycin	45	38.8	19	16.8	0.0002	**3.14**	1.70 - 5.66
Erithromycin	39	33.6	10	8.8	<0.0001	**5.22**	2.41 - 10.54
Cefepime	16	13.8	3	2.7	0.0031	**5.87**	1.76 - 19.35
Fosfomycin/ trometamol	60	51.7	63	55.8	0.5966	-	-
Linezolid	34	29.3	10	8.8	<0.0001	**4.27**	2.05 - 8.69
Meropenem	0	0.0	0	0.0	-	-	-
Rifampicin	15	12.9	6	5.3	0.0654	-	-
Trimethoprim/ sulfamethoxazol	24	20.7	9	8.0	0.0079	**3.01**	1.39 - 7.03
Tigecycline	0	0.0	0	0.0	-	-	-
Piperacillin/ tazobactam	2	1.7	0	0.0	0.4979	-	-
Vancomycin	42	36.2	14	12.4	<0.0001	**4.01**	2.09 - 7.89

[1]p-value calculated using a two-tailed Fisher's exact test; OR: odds ratio calculated for significant differences (p<0.05); CI: OR confidence interval calculated following the Baptista-Pike method.

Next, we evaluated the occurrence of resistance to multiple antibiotics among the isolates, which resisted up to ten different antibiotics (Figure 1C). From the total isolates, 50 (~22%) were resistant to four or more of these compounds, from which 41 came from non-intervened areas. Conversely, most of the isolates from humanized areas showed resistance to three or fewer antibiotics. In this direction, isolates from non-intervened areas averaged a significantly higher number of simultaneous resistance phenotypes ($p<0.05$, two-tailed Mann-Whitney test) (Figure 1D). Twenty-seven isolates showing resistance to 7 or more antibiotics were classified taxonomically based on 16S-rRNA analysis. Twenty-four of them were assigned to the genus *Pseudomona*s (Proteobacteria), one to *Pedobacter* (Bacteroidetes), and one to *Paeniglutamicibacter* (Actinobacteria) (Figure 1E). On the other hand, isolates showing resistance to amikacin or tazobactam/piperacillin (antibiotics with the lowest proportion of resistance among the whole collection) were classified as *Microbacterium* (Actinobacteria), *Acinetobacter* (Proteobacteria), *Flavobacterium* (Bacteroidetes), and *Arthrobacter* (Actinobacteria).

Taken together, these results indicate that there is a high proportion of antibiotic-resistant isolates in the sampled Antarctic soils, with no evidence of anthropogenic-derived enrichment. Additionally, most of the multi-resistant isolates came from non-intervened areas and corresponded to *Pseudomonas*.

Genomic features of multi-resistant Antarctic isolates

To investigate possible genetic determinants behind the multi-resistance phenotype, we selected two Antarctic isolates as models for genomic analyses: *Pseudomonas* YeP6b and ArH3a, showing resistance to 10 and 9 antibiotics, respectively (Figure 2A). Using both Illumina and Nanopore sequencing, we generated ~5.7 Gbp short reads (150 bp, paired-end) and 800 Mbp long reads for YeP6b, and ~6.7 Gbp short reads plus ~5.7 Gbp long reads for ArH3a, which were combined in each case to build hybrid assemblies using Canu (Koren et al., 2017) and Unicycler (Wick et al., 2017). This way, YeP6b showed a ~6.66 Mbp closed chromosome with no detected plasmids, while ArH3a showed a ~6.77 Mbp closed chromosome and a 2,628 bp small plasmid (pArH3a). Although coming from locations roughly 420 km far apart (Doumer and King George islands), both strains shared 99.46% average nucleotide identity (ANI), indicating they are members of the same species. Genome-based taxonomic placement using GTDB-Tk (Chaumeil et al., 2019) indicated *Pseudomonas* sp. IB20 (GCF_002263605.1) as the closest genome for both YeP6b and ArH3a, showing an ANI value of 99.28% and 99.36%, respectively. *P*. sp. IB20 was previously isolated from King George Island and classified inside the *Pseudomonas fluorescens* species complex, within the *P. antarctica* group and distant from animals or plants pathogens such as *P. aeruginosa* and *P. syringae* (Vásquez-Ponce et al., 2018). Phylogenomic analysis using Mashtree (Katz et al., 2019) supported this grouping (Supplementary Figure 1). These results support that the multi-resistant isolates YeP6b and ArH3a are part of a ubiquitous bacterial species naturally inhabiting North Antarctica's soils.

Genome annotation using RAST-Tk (Brettin et al., 2015) indicated 6,338 protein-coding genes in the *Pseudomonas* sp. ArH3a chromosome, among other features (Supplementary Table 4). Additionally, pArH3a had 5 CDS encoding a putative plasmid replication protein and four hypothetical proteins without predicted function. Likewise, we identified 6,141 CDS in the *Pseudomonas* YeP6b chromosome. Prophage content analysis using Phaster (Arndt et al., 2016) revealed eight of them in the chromosome of ArH3a and seven in YeP6b, sharing only one of them, reflecting intraspecies genome diversity due to horizontal gene transfer. Next, we used the tool Resistance Gene Identifier (RGI) from the Comprehensive

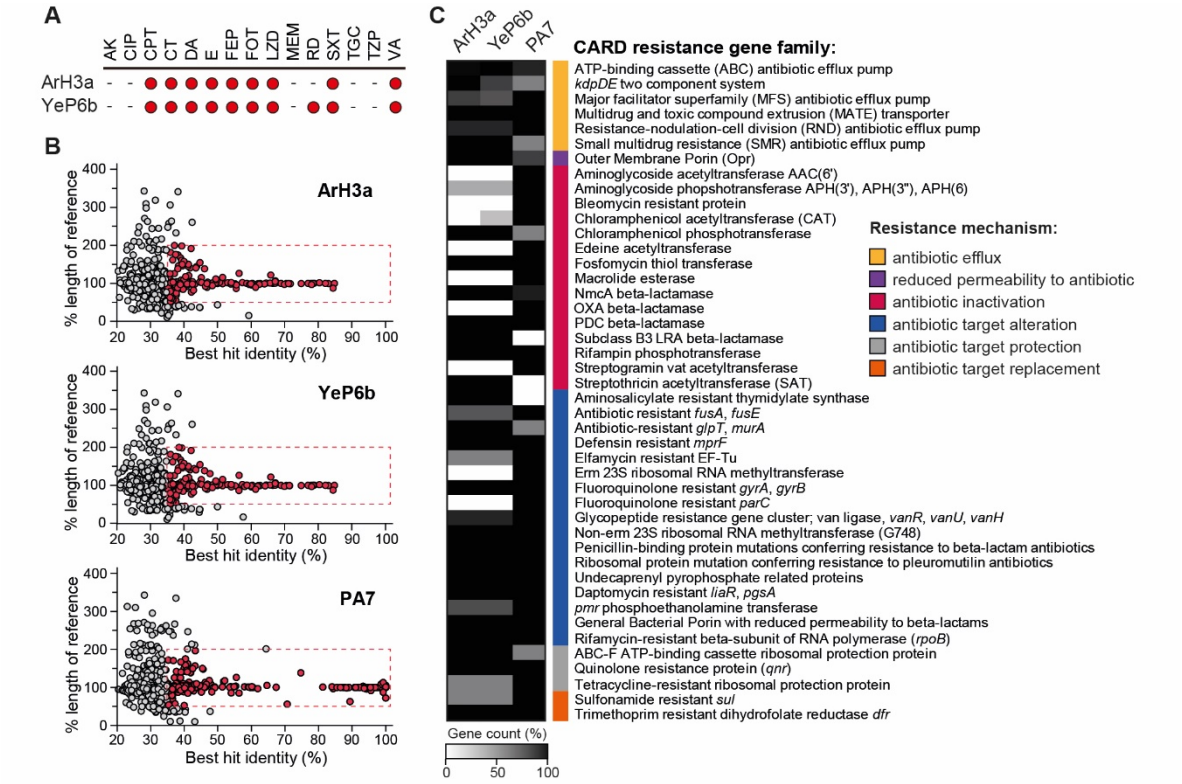

Figure 2. Experimental and genomic antibiotic resistance profile of the multirresistant Antarctic isolates *Pseudomonas* spp. ArH3A and YeP6b. (A) Experimental sensitivity to antibiotics determined for each Antarctic isolate. (B) Distribution of the percentage identity and the reference coverage of hits obtained upon searching for ARGs in the indicated genomes using RGI-CARD. The hits showed in red were selected to construct the heatmap shown in (C). (C) Heatmap showing the abundance of ARG gene families from different classes among Antarctic *Pseudomonas* ArH3a and Yep6b, and from the multiresistant clinical isolate *Pseudomonas aeruginosa* PA7. The gene count was calculated by dividing he number of all the detected ARGs falling in the mentioned category in each genome, divided by the highest number of ARGs.

Antibiotic Resistance Database (CARD) (Jia et al., 2017) to identify and classify putative antibiotic resistance genes (ARGs). For comparison, we included in the analysis the genome of the virulent clinical isolate *Pseudomonas aeruginosa* PA7, previously studied as a model multi-resistant strain (Roy et al., 2010; Singh et al., 2020). RGI allows ARG identification under three paradigms: perfect, strict, or loose. The Perfect and Strict algorithms detect perfect matches to the curated reference sequences or previously unknown variants of known AMR genes and are best suited for clinical surveillance of known bacterial pathogens (Jia et al., 2017).

We found no perfect hits and only four strict hits for both YeP6b and ArH3a (Supplementary Table 5), including a homolog of the *P. aeruginosa* transcriptional activator SoxR linked to increased expression of RND efflux pumps (Palma et al., 2005; Sporer et al., 2018). Also, two genes similar to the AdeF RND efflux pump and one homolog to the AbaQ MFS efflux pump, both of them described in *Acinetobacter baumannii* and reported to confer resistance to tetracycline, fluoroquinolones and other antibiotics (Coyne et al., 2010; Pérez-Varela et al., 2018). RGI analysis indicated five perfect and forty-one strict hits for PA7, including genes encoding four aminoglycoside phosphotransferases and one acetyltransferase, the Cmx chloramphenicol exporter, the OXA-50 beta-lactamase, and several efflux pumps (i.e. MexAB-OprM, MexCD-OprJ, and MexXY) (Supplementary Table 5). The lower number of ARGs found in the multi-resistant Antarctic isolates suggests that functionally similar genes sharing limited identity with known ARGs or genes encoding different molecular mechanisms would explain their resistance phenotype.

To get a more comprehensive view of possible genes involved in Antarctic isolate multi-resistance, we included in the analysis the loose hits, algorithm recommended for new resistance genes discovery in bacteria not closely related to clinical isolates, although with increased expected false positives (Jia et al., 2017). We found 700 and 701 additional putative ARGs in YeP6b and ArH3a, respectively, and 671 in PA7. As a way to identify and filter out possible false positives, we plotted for each hit the percentage identity and length to the reference ARG, noticing a high dispersion of the coverage length below 35% identity (Figure 2B). In contrast, we found a clear tendency to similar length at increasing identity values. Thus, to focus on the genes more likely sharing functions with known resistance genes, we decided to include the hits showing ≥35%+ identity and 50-200% length of the reference. This way, 168, 165, and 228 hits remained for ArH3a, YeP6b, and PA7, respectively. We grouped all the putative ARGs by gene family and resistance mechanism, then calculating each category's relative gene count in each strain (Figure 2C). Antarctic isolates showed an overall lower number of predicted ARGs compared to P. aeruginosa PA7, lacking genes encoding aminoglycoside acetyltransferase AAC(6'), Macrolide esterases, OXA beta-lactamases, Streptogramin vat acetyltransferases, and Fluoroquinolone-resistant parC. However, the Antarctic isolates showed a higher number of genes encoding ATP-binding cassette (ABC) antibiotic efflux pumps and Small multidrug resistance (SMR) antibiotic efflux pumps, compared to PA7. Also, ArH3a and YeP6b strains shared with PA7 a putative carbapenem-hydrolyzing Class-A NmcA beta-lactamase originally described in *Enterobacter cloacae* (Naas & Nordmann, 1994), and a putative class-C PDC beta-lactamase, which is normally found in *Pseudomonas aeruginosa*, conferring resistance to monobactam, carbapenems, and cephalosporins (Rodríguez-Martínez et al., 2009). Remarkably, among the putative ARGs found in the Antarctic isolates but absent in PA7, we found an homolog of the LRA-3 beta-lactamase, previously isolated from soil samples in Alaska (Allen et al., 2009). Functional metagenomics approaches showed that this beta lactamase confer *E. coli* resistance to cephalosporins and penam antibiotics.

All these results indicate that Antarctic soil culturable microbiota is a source of genes that, although encoding proteins sharing limited identity with known resistance genes, would retain the capability to confer resistance to various antibiotics. This microbiota would also host genes encoding novel resistance mechanisms that could be further studied to anticipate possible new resistance mechanisms arising in the future among bacterial pathogens.

Microbial diversity present in the sampled Antarctic soils

To evaluate and compare the microbial diversity present in humanized and non-intervened areas, we performed 16S rRNA amplicon sequencing analysis using the platform QIIME2. Relative abundance calculations indicated that Proteobacteria, Bacteroidota, Acidobacteriota, and Verrucomicrobiota corresponded to the most abundant Phyla in most of the soils (Figure 3A), while among the most abundant genera were *Candidatus Udaeobacter*, RB41, *Polaromonas*, and *Ferruginibacter* (Figure 3B). corresponded to the predominant Phyla in most of the sampled sites. Remarkably, a recent study of microbial diversity from ice-free soils in the Arctic also reported a high proportion of Candidatus Udaeobacter (Verrucomicrobiota) and RB41 (Acidobacteriota), followed by uncultured Gemmatimonadota and *Polaromonas* (Venkatachalam et al., 2021). A marked decrease in the two taxa mentioned first was observed in humanized soils.
Total (alpha) microbial diversity, estimated by both Shannon and Pielou Indexes, were high for all the samples, showing no significant differences when comparing the average value of humanized vs. non-intervened sites (Figure 3C). On the other hand, beta diversity calculations based on Bray-Curtis dissimilarity and Robust Aitchison compositional analysis revealed differences in the composition of the microbial communities at each site, although a significant clustering of non-intervened sites apart from humanized areas was observed (P=0.004 and 0.006, for Bray-Curtis and Aitchison, respectively) (Figure 3D).

Detection of antibiotic resistance genes among Antarctic soil microbial communities.

To evaluate the presence of ARGs among the soil microbial communities, we performed shotgun metagenomic sequencing. First, illumina sequencing was performed for soil extracted from the four non-intervened and three humanized areas. Sequence diversity analysis using Mash supported the differential clustering of both kinds of areas (Figure 3E and 3F). illumina reads were used to perform a mapping-based strategy to identify and calculate the relative abundance of ARGs using deepARG and custom scripts. This approach has the advantage of being independent on the results of metagenomic assembly, which could be very different for distinct soil samples. However, due to its reduced length, short read mapping does not allow to identify complete ARGs. Overall, the metagenomic reads from the sum of all the samples revealed the presence of hundreds of different ARGs, which fall into more than 17 classes according to the drug to which they confer resistance (Figure 4A). As observed for the multi-resistant isolates, we found a high proportion and diversity of genes encoding multidrug efflux pumps and transporters among the Antarctic metagenomes, followed by genes putatively conferring resistance to Macrolides, Streptogramin and Lincosamide, as well as to tetracycline, Bacitracin, and glycopeptides. The Arctic LRA-1 beta-lactamase was also found in our Antarctic metagenomes. In order to get a first approach of the microbial taxa harboring the ARGs that generated positive hits, we extracted all the reads mapping to any ARG and performed taxonomic assignment using Kaiju (Menzel et al., 2016). This analysis indicated that most of the reads mapping to ARGs come from *Pseudomonas*, *Streptomyces*, *Gemmatimonas*, and *Paenibacillus*, among others. This supports that the ARGs found in the soil metagenomes came from bacteria

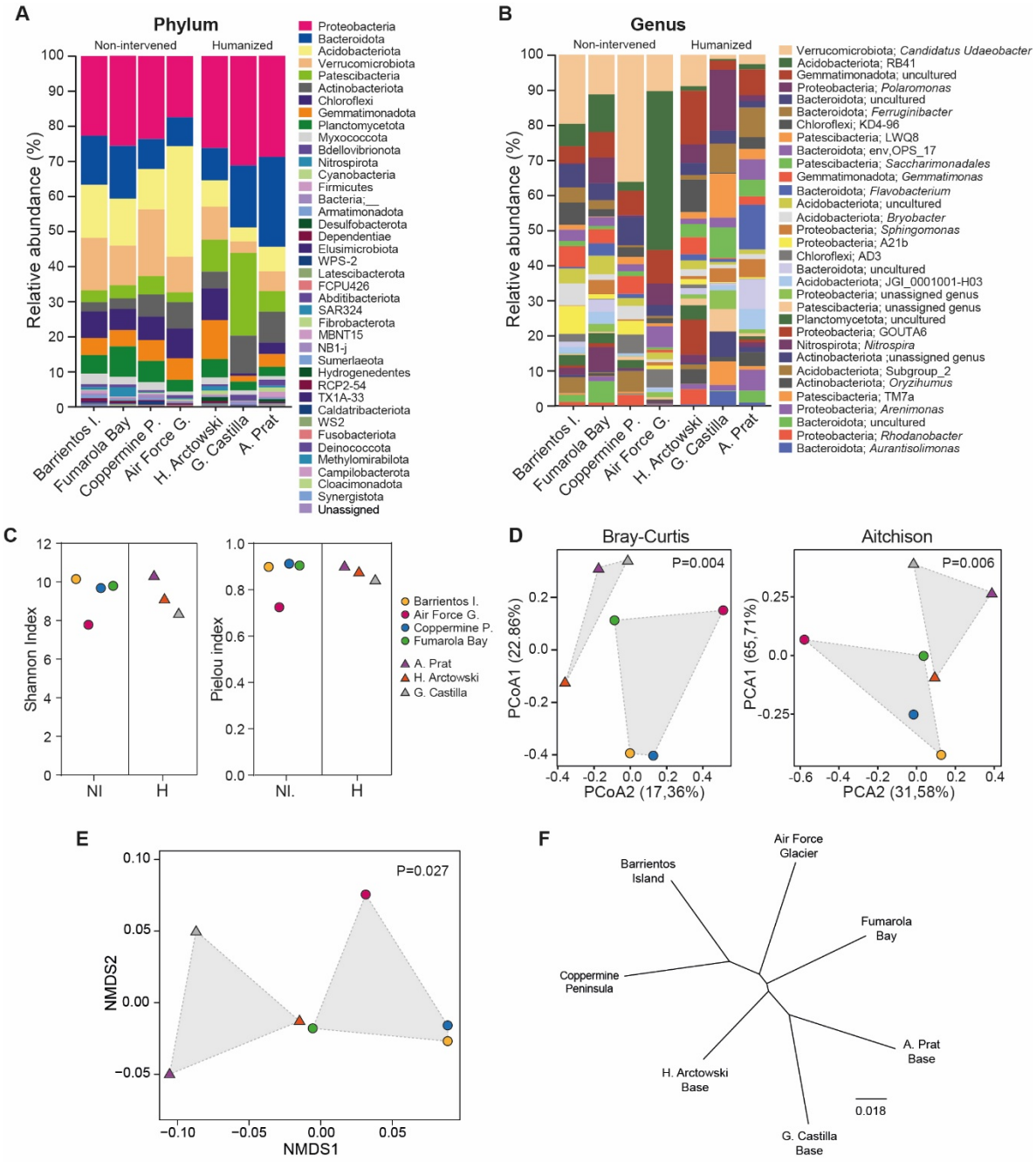

Figure 3. Microbial and DNA sequence diversity in the sampled Antarctic soils. (A-D) 16S rRNA-based microbial diversity analysis. Relative abundance of bacterial phyla (A) and genera (B) among the sampled sites. (C) Alpha diversity estimation based on Shannon and Pielou index. (D) Beta diversity estimation based on Bray-curtis dissimilarity index and the Robust Aitchison analysis. (E-F) Mash analysis of sequence diversity among the shotgun metagenomes of the analyzed soil samples. (E) Non-metric multidimensional scaling based on the Mash distances. Mashtree inferred from the Mash distances calculated for the set of metagenomes. The P-value supporting sample clustering as represented by gray polygons was calculated performing a PERMANOVA analysis.

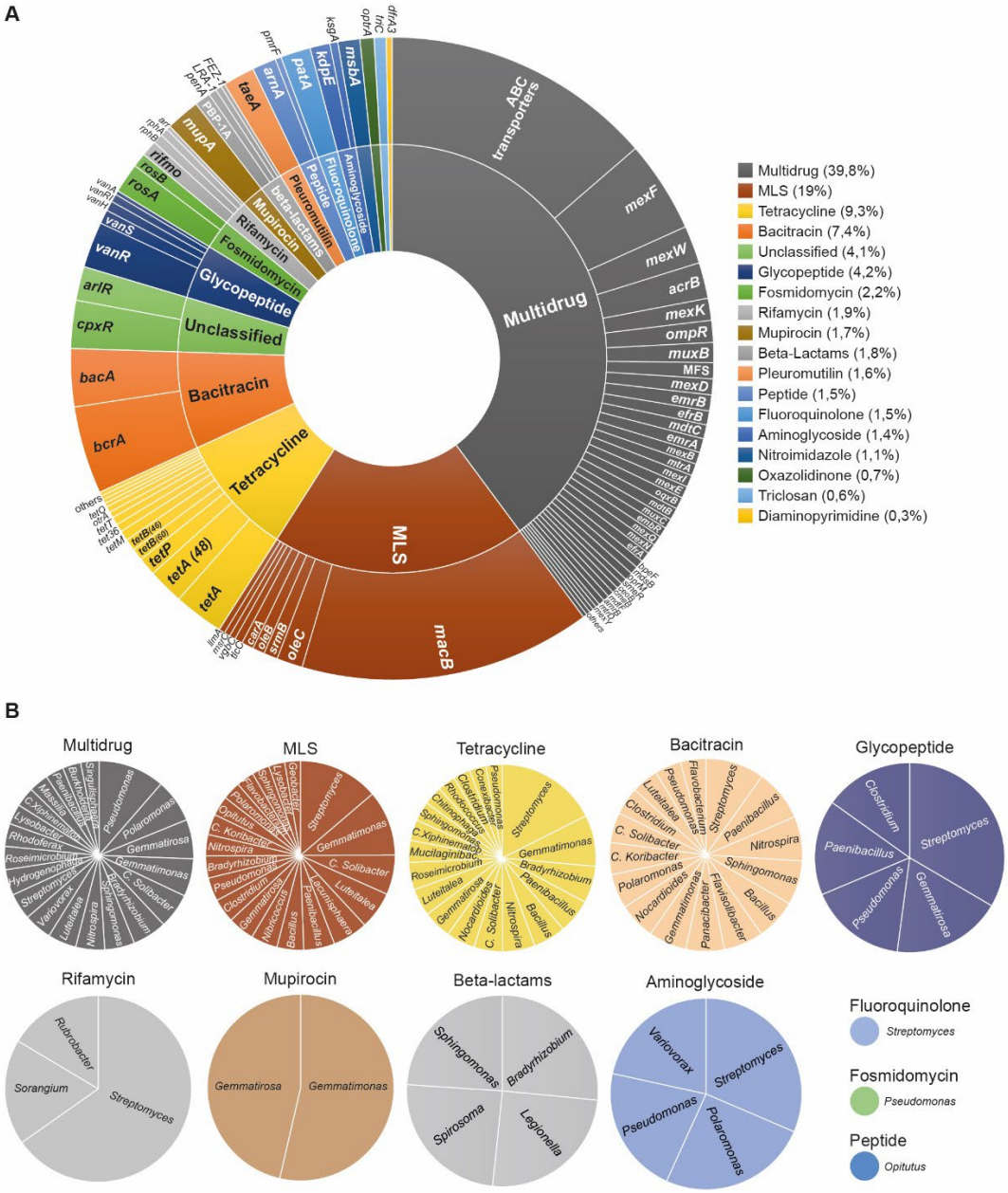

Figure 4. Overview of the ARG hits found in among Antarctic soil metagenomes. (A) Sunburst plot showing the proportion of hits found for each resistance gene family and drug class. (B) Taxonomic asignment of the reads mapped to the sum of ARGs detected for each drug class.

naturally inhabiting Antarctic soils, rather than from non-indigenous bacteria introduced from other ecosystems.

To evaluate and compare the relative abundance of ARGs in different soils, we conducted a normalization strategy based on dividing the number of reads mapping to a defined ARG in a defined site, by the total number of reads mapping to a *rpoB* gene database in the same site. The relative calculated abundances were used to construct a heatmap including the 100 most abundant ARGs (Figure 5A). No evident patterns were observed when comparing the abundances of specific ARGs in humanized vs. non-intervened sites, although an overall higher abundance or ARGs was noticed in Coppermine Peninsula and Fumarola Bay. To search for possible ARGs able to discriminate both kinds of soils, we used the tool extrARG. 23 discriminatory ARGs were identified, covering different ARG classes and mechanisms (Figure 5B). The discriminatory power of these ARGs was assessed through NMDS analysis, obtaining a clear significant differential grouping of the samples (Figure 5C). Among the ARGs found to be characteristic of humanized zones are *dfrK*, arr-5, *axyX*, *pmrE*, and JOHN-1. Conversely, ARGs characteristic of non-intervened zones are *cpxA*, AAC(6')-I, *vgbC, mexM*, and *mel.*

Detection of complete antibiotic resistance genes in assembled Antarctic soil metagenomes and examination of their context

To get further evidence of the presence of ARGs among the Antarctic soil metagenomes and to gain more information regarding these genes and their genomic context, we performed Nanopore sequencing of three of the seven metagenomes analyzed in the previous section (G. de Castilla Base, A. Prat Base, and Air Force Glacier). Long Nanopore reads, when combined with short illumina reads, allow to construct hybrid assemblies, which normally results in less fragmented assemblies (longer contigs). Upon constructing the hybrid assemblies, the generated contigs were searched for the presence of ARGs using the platform nanoARG (Arango-Argoty et al., 2019). Positive hits (with a cutoff of 40% identity and 1E-10 e-value, defined based on the observed distribution of identity and e-value), classified, counted, and the proportion of each gene and category over the total number of ARGs identified was determined and plotted as a treemap (Figure 6A). A similar array of ARGs than those found in short reads-based analysis was observed for assembled metagenomic data. Taxonomic assignment of the identified contigs harboring ARGs using Centrifuge (Kim et al., 2016) supported *Polaromonas*, *Pseudomonas*, *Streptomyces*, *Variovorax*, *Bhurkolderia, and Gemmatimonas* as the main host taxa for the identified ARGs.

To evaluate if some of the identified ARGs are located inside mobile genetic elements and thus could be disseminated through horizontal gene transfer, we used the tools PlasFlow (Krawczyk et al., 2018) and Viralverify (Antipov et al., 2020), to identify putative plasmids among the assembled metagenomic contigs predicted to harbor one or more ARGs. From 24,742 contigs harboring ARGs, 915 were predicted as plasmids by Viralverify, 3238 by PlasFlow, and 177 by both tools (Figure 7A). To minimize false positives, we focused on these 177. Putative plasmid contigs examination led to the identification of a 15 kbp contig harboring a OXA-like beta-lactamase which share roughly 65% identity with OXA carbapenemases detected in pathogenic bacteria, harboring the conserved beta-lactamase domain and the critical residues required for its characteristic activity (Figure 7B). Moreover, upon comparing this putative plasmid-borne Antarctic beta-lactamase with the NCBI database using BLAST, we noticed that it shares 98% identity with a Class-D beta-lactamase from *Polaromonas* sp. (Accession QJS06455.1). This sequence was found to be part of a circular replicon annotated as the plasmid pA29BJ2 from *Polaromonas* sp.

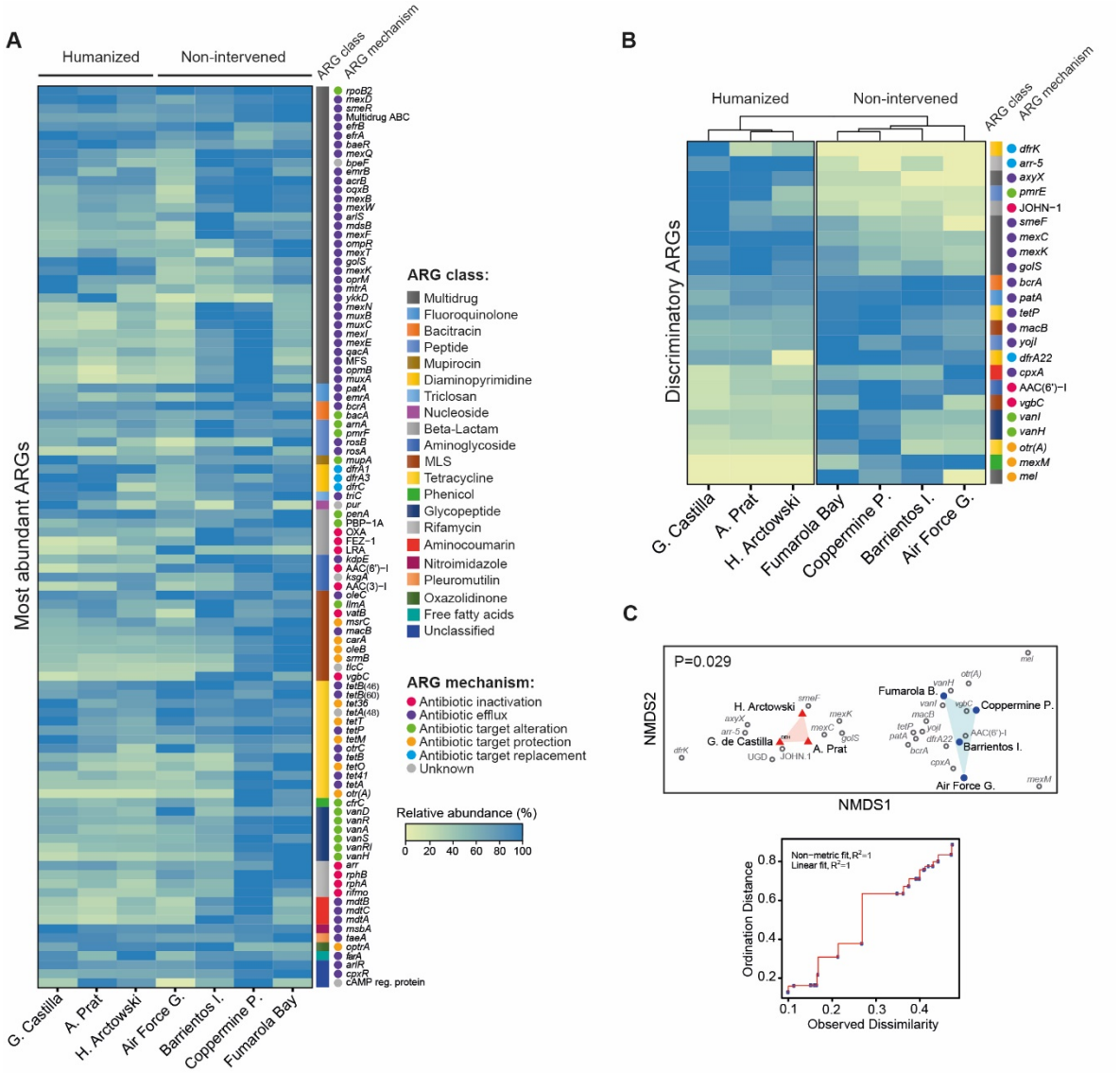

Figure 5. Relative abundance of predicted antibiotic resistance genes among soil microbiota inhabiting humanized and non-intervened Antarctic areas. (A) Heatmap showing the normalized abundance of different ARG families in different Antarctic sites, calculated using a short-reads mapping based strategy. Normalization of the abundance comparing different sites was performed based on reads mapping to *rpoB*. Also, for each ARG the abundance was normalized by the highest value among the sites (for each column). For each ARG, the resistance gene family and drug class is indicated (based on CARD categorization). (B) Heatmap showing the relative abundance of discriminant ARGs identified with the extrARG tool. (C) Non-metric multidimensional scaling analysis based on the abundance dissimilarity of the discriminatory ARGs identified among the different sites (shown in gray). The P-value was calculated from PERMANOVA analysis, supporting separate clustering of humanized and non-intervened sites.

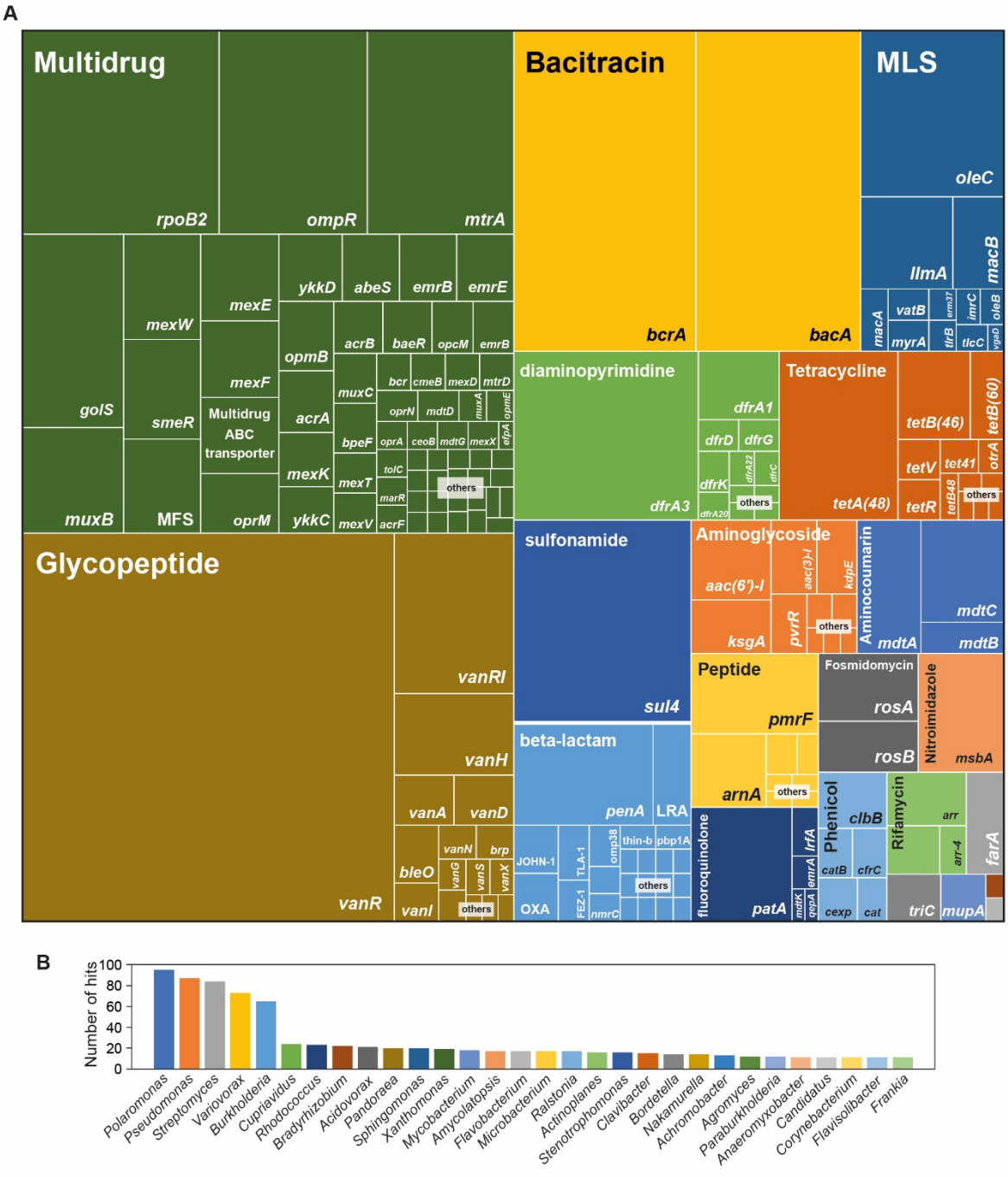

Figure 6. Antibiotic resistance gene families found in assembled Antarctic soil metagenomes. (A) Tree-map showing the proportion of hits found for each resistance gene family and drug class. (B) Number of contigs carrying one or more ARG that were assigned to the listed bacterial genera.

ANT_J29B, a strain isolated from a soil sample from Jardine Peak, King George Island (Antarctica). Taken together all this evidence support the presence of Antarctic ARGs inside mobile genetic elements such as plasmids.

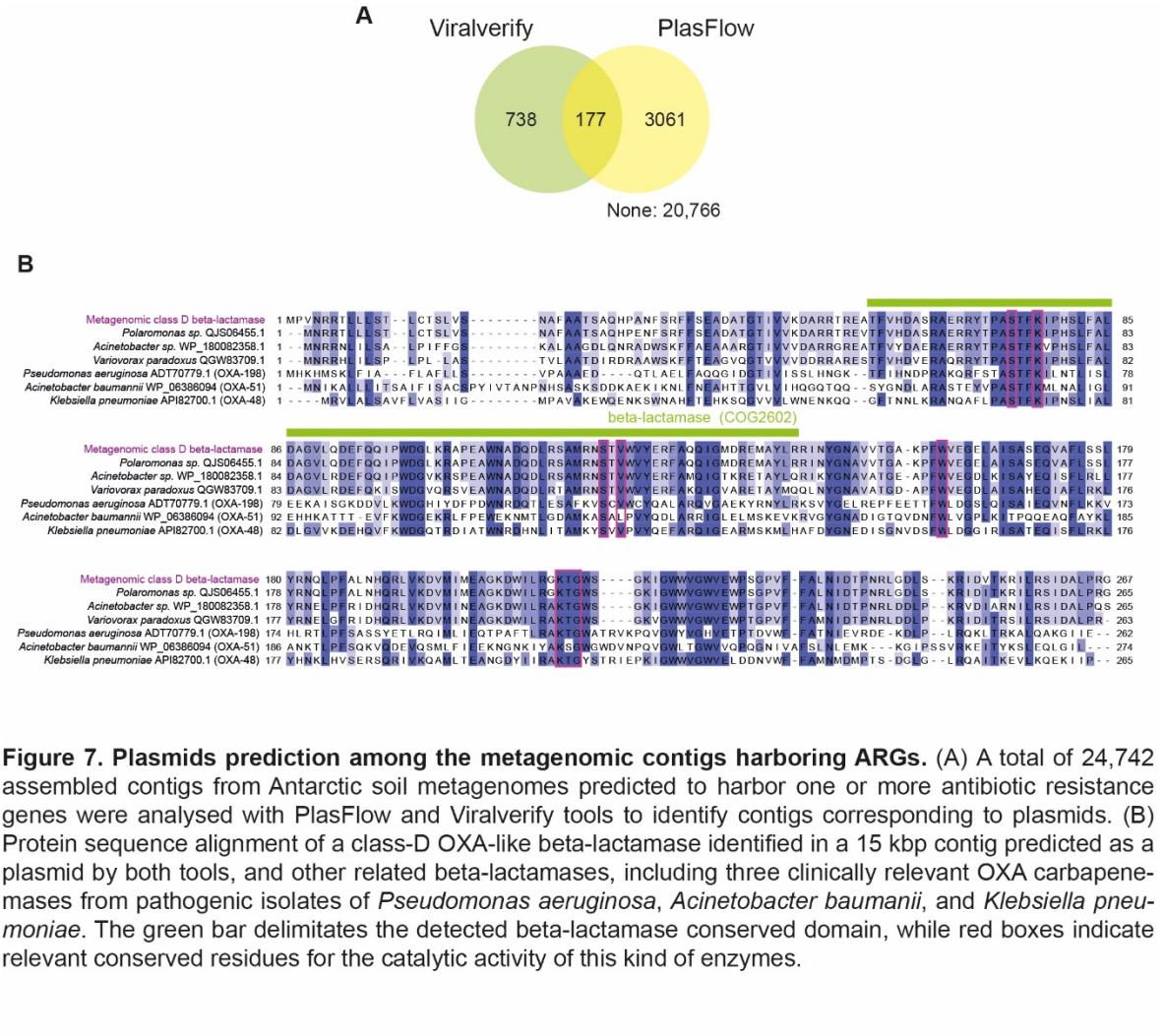

Figure 7. Plasmids prediction among the metagenomic contigs harboring ARGs. (A) A total of 24,742 assembled contigs from Antarctic soil metagenomes predicted to harbor one or more antibiotic resistance genes were analysed with PlasFlow and Viralverify tools to identify contigs corresponding to plasmids. (B) Protein sequence alignment of a class-D OXA-like beta-lactamase identified in a 15 kbp contig predicted as a plasmid by both tools, and other related beta-lactamases, including three clinically relevant OXA carbapenemases from pathogenic isolates of *Pseudomonas aeruginosa*, *Acinetobacter baumanii*, and *Klebsiella pneumoniae*. The green bar delimitates the detected beta-lactamase conserved domain, while red boxes indicate relevant conserved residues for the catalytic activity of this kind of enzymes.

REFERENCES

Allen, H. K., Moe, L. A., Rodbumrer, J., Gaarder, A., & Handelsman, J. (2009). Functional metagenomics reveals diverse B-lactamases in a remote Alaskan soil. *ISME Journal*, *3*(2), 243–251. https://doi.org/10.1038/ismej.2008.86

Alneberg, J., Bjarnason, B. S., De Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., Lahti, L., Loman, N. J., Andersson, A. F., & Quince, C. (2014). Binning metagenomic contigs by coverage and composition. *Nature Methods*, *11*(11), 1144–1146. https://doi.org/10.1038/nmeth.3103

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

Antipov, D., Raiko, M., Lapidus, A., & Pevzner, P. A. (2020). MetaviralSPAdes: assembly of viruses from metagenomic data. *Bioinformatics*, *36*(14), 4126–4129. https://doi.org/10.1093/bioinformatics/btaa490

Arango-Argoty, G., Dai, D., Pruden, A., Vikesland, P., Heath, L. S., & Zhang, L. (2019). NanoARG: A web service for detecting and contextualizing antimicrobial resistance genes from nanopore-derived metagenomes. *Microbiome*, *7*(1), 1–18. https://doi.org/10.1186/s40168-019-0703-9

Arango-Argoty, G., Garner, E., Pruden, A., Heath, L. S., Vikesland, P., & Zhang, L. (2018). DeepARG: A deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*, *6*(1), 1–15. https://doi.org/10.1186/s40168-018-0401-z

Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., & Wishart, D. S. (2016). PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Research*, *44*(W1), W16–W21. https://doi.org/10.1093/nar/gkw387

Blair, J. M. A., Webber, M. A., Baylay, A. J., Ogbolu, D. O., & Piddock, L. J. V. (2015). Molecular mechanisms of antibiotic resistance. In *Nature Reviews Microbiology* (Vol. 13, Issue 1). https://doi.org/10.1038/nrmicro3380

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

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., … Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology*, *37*(8), 852–857. https://doi.org/10.1038/s41587-019-0209-9

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R., & Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, *5*. https://doi.org/10.1038/srep08365

Chaumeil, P.-A., Mussig, A. J., Hugenholtz, P., & Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics*, *36*(November 2019), 1925–1927. https://doi.org/10.1093/bioinformatics/btz848

Coyne, S., Rosenfeld, N., Lambert, T., Courvalin, P., & Périchon, B. (2010). Overexpression of resistance-nodulation-cell division pump AdeFGH confers multidrug resistance in Acinetobacter baumannii. *Antimicrobial Agents and*

Chemotherapy, *54*(10), 4389–4393. https://doi.org/10.1128/AAC.00155-10

Forsberg, K. J., Reyes, A., Wang, B., Selleck, E. M., Sommer, M. O. A., & Dantas, G. (2012). The Shared Antibiotic Resistome of Soil Bacteria and Human Pathogens. *Science*, *337*(August), 1107–1111. https://doi.org/10.1126/science.1220761

Friedman, N. D., Temkin, E., & Carmeli, Y. (2016). The negative impact of antibiotic resistance. *Clinical Microbiology and Infection*, *22*(5), 416–422. https://doi.org/10.1016/j.cmi.2015.12.002

Goethem, M. W. Van, Pierneef, R., Bezuidt, O. K. I., Peer, Y. Van De, Cowan, D. A., & Makhalanyane, T. P. (2018). *A reservoir of ' historical ' antibiotic resistance genes in remote pristine Antarctic soils*. 1–12.

Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., Von Mering, C., & Bork, P. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution*, *34*(8), 2115–2122. https://doi.org/10.1093/molbev/msx148

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., Von Mering, C., & Bork, P. (2019). EggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Research*, *47*(D1), D309–D314. https://doi.org/10.1093/nar/gky1085

Hutchings, M., Truman, A., & Wilkinson, B. (2019). Antibiotics: past, present and future. *Current Opinion in Microbiology*, *51*(Figure 1), 72–80. https://doi.org/10.1016/j.mib.2019.10.008

Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., Lago, B. A., Dave, B. M., Pereira, S., Sharma, A. N., Doshi, S., Courtot, M., Lo, R., Williams, L. E., Frye, J. G., Elsayegh, T., Sardar, D., Westman, E. L., Pawlowski, A. C., … McArthur, A. G. (2017). CARD 2017: Expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*, *45*(D1), D566–D573. https://doi.org/10.1093/nar/gkw1004

Jolley, K. A., Bliss, C. M., Bennett, J. S., Bratcher, H. B., Brehony, C., Colles, F. M., Wimalarathna, H., Harrison, O. B., Sheppard, S. K., Cody, A. J., & Maiden, M. C. J. (2012). Ribosomal multilocus sequence typing: Universal characterization of bacteria from domain to strain. *Microbiology*, *158*(4), 1005–1015. https://doi.org/10.1099/mic.0.055459-0

Jolley, K. A., & Maiden, M. C. J. (2010). BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics*, *11*(1), 595. https://doi.org/10.1186/1471-2105-11-595

Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., & Wang, Z. (2019). MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*, *2019*(7), 1–13. https://doi.org/10.7717/peerj.7359

Katz, L., Griswold, T., Morrison, S., Caravas, J., Zhang, S., Bakker, H., Deng, X., & Carleton, H. (2019). Mashtree: a rapid comparison of whole genome sequence files. *Journal of Open Source Software*, *4*(44), 1762. https://doi.org/10.21105/joss.01762

Kim, D., Song, L., Breitwieser, F. P., & Salzberg, S. L. (2016). Centrifuge: Rapid and sensitive classification of metagenomic sequences. *Genome Research*, *26*(12), 1721–1729. https://doi.org/10.1101/gr.210641.116

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive κ-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

Krawczyk, P. S., Lipinski, L., & Dziembowski, A. (2018). PlasFlow: predicting plasmid sequences in metagenomic data using genome signatures. *Nucleic Acids Research*,

46(6), e35. https://doi.org/10.1093/nar/gkx1321

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

Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 1–4. https://doi.org/10.1093/nar/gkab301

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

McCann, C. M., Christgen, B., Roberts, J. A., Su, J. Q., Arnold, K. E., Gray, N. D., Zhu, Y. G., & Graham, D. W. (2019). Understanding drivers of antibiotic resistance genes in High Arctic soil ecosystems. *Environment International*, *125*(November 2018), 497–504. https://doi.org/10.1016/j.envint.2019.01.034

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications*, *7*. https://doi.org/10.1038/ncomms11257

Mikheenko, A., Saveliev, V., & Gurevich, A. (2016). MetaQUAST: Evaluation of metagenome assemblies. *Bioinformatics*, *32*(7), 1088–1090. https://doi.org/10.1093/bioinformatics/btv697

Naas, T., & Nordmann, P. (1994). Analysis of a carbapenem-hydrolyzing class A β-lactamase from Enterobacter cloacae and of its LysR-type regulatory protein. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(16), 7693–7697. https://doi.org/10.1073/pnas.91.16.7693

Nurk, S., Meleshko, D., Korobeynikov, A., & Pevzner, P. A. (2017). MetaSPAdes: A new versatile metagenomic assembler. *Genome Research*, *27*(5), 824–834. https://doi.org/10.1101/gr.213959.116

Palma, M., Zurita, J., Ferreras, J. A., Worgall, S., Larone, D. H., Shi, L., Campagne, F., & Quadri, L. E. N. (2005). Pseudomonas aeruginosa SoxR does not conform to the archetypal paradigm for SoxR-dependent regulation of the bacterial oxidative stress adaptive response. *Infection and Immunity*, *73*(5), 2958–2966. https://doi.org/10.1128/IAI.73.5.2958-2966.2005

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

Pérez-Varela, M., Corral, J., Aranda, J., & Barbé, J. (2018). Functional Characterization of AbaQ, a Novel Efflux Pump Mediating Quinolone Resistance in Acinetobacter baumannii. *Antimicrobial Agents and Chemotherapy*, *62*(9), 1–4. https://doi.org/10.1128/AAC.00906-18

Rodríguez-Martínez, J. M., Poirel, L., & Nordmann, P. (2009). Extended-spectrum cephalosporinases in Pseudomonas aeruginosa. *Antimicrobial Agents and Chemotherapy*, *53*(5), 1766–1771. https://doi.org/10.1128/AAC.01410-08

Roy, P. H., Tetu, S. G., Larouche, A., Elbourne, L., Tremblay, S., Ren, Q., Dodson, R., Harkins, D., Shay, R., Watkins, K., Mahamoud, Y., & Paulsen, I. T. (2010). Complete genome sequence of the multiresistant taxonomic outlier Pseudomonas aeruginosa PA7. *PLoS ONE*, *5*(1), 1–10. https://doi.org/10.1371/journal.pone.0008842

Seemann, T. (2014). Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, *30*(14), 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Singh, M., Sykes, E. M. E., Li, Y., & Kumar, A. (2020). MexXY RND pump of <i>Pseudomonas aeruginosa</i> PA7 effluxes bi-anionic β-lactams carbenicillin and sulbenicillin when it partners with the outer membrane factor OprA but not with OprM. *Microbiology (Reading, England)*, *166*(11), 1095—1106.

https://doi.org/10.1099/mic.0.000971

Smith, R. A., M'ikanatha, N. M., & Read, A. F. (2015). Antibiotic Resistance: A Primer and Call to Action. *Health Communication*, *30*(3), 309–314. https://doi.org/10.1080/10410236.2014.943634

Sporer, A. J., Beierschmitt, C., Bendebury, A., Zink, K. E., Price-Whelan, A., Buzzeo, M. C., Sanchez, L. M., & Dietrich, L. E. P. (2018). Pseudomonas aeruginosa PumA acts on an endogenous phenazine to promote self-resistance. *Microbiology (United Kingdom)*, *164*(5), 790–800. https://doi.org/10.1099/mic.0.000657

Tan, L., Li, L., Ashbolt, N., Wang, X., Cui, Y., Zhu, X., Xu, Y., Yang, Y., Mao, D., & Luo, Y. (2018). Arctic antibiotic resistance gene contamination, a result of anthropogenic activities and natural origin. *Science of the Total Environment*, *621*(October), 1176–1184. https://doi.org/10.1016/j.scitotenv.2017.10.110

Uritskiy, G. V., DiRuggiero, J., & Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*, *6*(1), 1–13. https://doi.org/10.1186/s40168-018-0541-1

Vásquez-Ponce, F., Higuera-Llantén, S., Pavlov, M. S., Marshall, S. H., & Olivares-Pacheco, J. (2018). Phylogenetic MLSA and phenotypic analysis identification of three probable novel Pseudomonas species isolated on King George Island, South Shetland, Antarctica. *Brazilian Journal of Microbiology*, *49*(4), 695–702. https://doi.org/10.1016/j.bjm.2018.02.005

Venkatachalam, S., Kannan, V. M., Saritha, V. N., Loganathachetti, D. S., Mohan, M., & Krishnan, K. P. (2021). Bacterial diversity and community structure along the glacier foreland of Midtre Lovénbreen, Svalbard, Arctic. *Ecological Indicators*, *126*, 107704. https://doi.org/10.1016/j.ecolind.2021.107704

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). https://doi.org/10.1371/journal.pcbi.1005595

Wu, Y. W., Simmons, B. A., & Singer, S. W. (2016). MaxBin 2.0: An automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*, *32*(4), 605–607. https://doi.org/10.1093/bioinformatics/btv638