Front. Environ. Sci. Eng. 2019, 13(3): 44
https://doi.org/10.1007/s11783-019-1128-1

RESEARCH ARTICLE

Detection of presumed genes encoding beta-lactamases by sequence based screening of metagenomes derived from Antarctic microbial mats

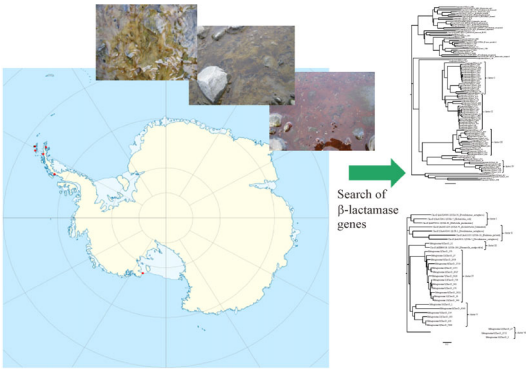
Gastón Azziz (✉)[1,2], Matías Giménez[1], Héctor Romero[3], Patricia M. Valdespino-Castillo[4], Luisa I. Falcón[5,6], Lucas A. M. Ruberto[7,8], Walter P. Mac Cormack[7,8], Silvia Batista[1]

1 Molecular Microbiology Unit, Clemente Estable Biological Research Insitute, Montevideo 11600, Uruguay
2 Microbiology Laboratory, Faculty of Agronomy, UdelaR, Montevideo 12900, Uruguay
3 Genome Organization and Evolution Laboratory, Ecology and Evolution Department, Faculty of Sciences, UdelaR, Montevideo 11400, Uruguay
4 Molecular Biophysics and Integrated Bioimaging, BSISB Imaging Program, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA
5 Bacteial Ecology Laboratory, Ecology Institute, National Autonomous University of Mexico, CDMX 04510, Mexico
6 UNAM, Yucatan Technology and Science Park, Merida 97302, Mexico
7 Argentine Antarctic Institute, Buenos Aires 1650, Argentina
8 Biotechnology Unit, Faculty of Pharmacy and Biochemistry, Nanobiotec Institute UBA-CONICET, Buenos Aires 1113, Argentina

HIGHLIGHTS

- Beta-lactamase genes were found in all samples from distant places in Antarctica.
- Class C beta-lactamase coding genes were the most frequently found.
- Diversity of sequences exceeds that of the beta-lactamases from clinical environment.

GRAPHIC ABSTRACT



ARTICLE INFO

Article history:
Received 7 January 2019
Revised 10 May 2019
Accepted 15 May 2019
Available online 20 June 2019

Keywords:
Beta-lactamases
Antibiotic resistance coding genes
Metagenomes
Antarctic microbial mats

ABSTRACT

Analysis of environmental samples for bacterial antibiotic resistance genes may have different objectives and analysis strategies. In some cases, the purpose was to study diversity and evolution of genes that could be grouped within a mechanism of antibiotic resistance. Different protocols have been designed for detection and confirmation that a functional gene was found. In this study, we present a sequence-based screening of candidate genes encoding beta-lactamases in 14 metagenomes of Antarctic microbial mats. The samples were obtained from different sites, representing diverse biogeographic regions of maritime and continental Antarctica. A protocol was designed based on generation of Hidden Markov Models from the four beta-lactamase classes by Ambler classification, using sequences from the Comprehensive Antibiotic Resistance Database (CARD). The models were used as queries for metagenome analysis and recovered contigs were subsequently annotated using RAST. According to our analysis, 14 metagenomes analyzed contain A, B and C beta-lactamase genes. Class D genes, however, were identified in 11 metagenomes. The most abundant was class C (46.8%), followed by classes B (35.5%), A (14.2%) and D (3.5%). A considerable number of sequences formed clusters which included, in some cases, contigs from different metagenomes. These assemblies are clearly separated from reference clusters, previously identified using CARD beta-lactamase sequences. While bacterial antibiotic resistance is a major challenge of public health worldwide, our results suggest that environmental diversity of beta-lactamase genes is higher than that currently reported, although this should be complemented with gene function analysis.

✉ Corresponding author
E-mail: gazziz@gmail.com

Special Issue—Environmental Antibiotics and Antibiotic Resistance (Responsible Editors: Xin Yu, Hui Li & Virender K. Sharma)

1 Introduction

Since their discovery, antibiotics have been widely used as therapeutic agents to treat bacterial infections. Penicillin G, a beta-lactam antibiotic, was the first antibiotic used for clinical therapy. It was discovered in 1928 by Alexander Fleming and the first use of penicillin as therapeutic agent dates back to 1940 (Bennett and Chung, 2001). Since the first years of clinical use, it became apparent that bacteria were able to develop resistance to antibiotics. The strategy was then to replace antibiotics to which resistance had emerged with new ones, which were continuously discovered or designed (Davies and Davies, 2010). Today, establishment and dispersion of bacterial resistance has become one of the main challenges faced by those responsible for the design of public health policies worldwide (Högberg et al., 2010).

Beta-lactams are the most widely used group of antibiotics. These compounds are bactericidal, and their mechanism of action involves inhibition of bacterial peptidoglycan synthesis. Few mechanisms of resistance to beta-lactams have been described in bacteria. For instance, target proteins of beta-lactams, known as penicillin binding protein (PBPs), may be modified in which their affinity for beta-lactams is diminished. Also, in some Gram-negative bacteria, downregulation in expression of specific outer membrane proteins results in a reduced uptake of the antibiotic (Babic et al., 2006). However, the most important mechanism of resistance is the production of beta-lactamases. These enzymes inactivate the antibiotic by hydrolysis of the beta-lactam ring. Understanding the diversity and distribution of beta-lactamases is key to design effective public health strategies.

Four classes of beta-lactamases are recognized according to Ambler classification, which is based on amino-acid sequences (Hall and Barlow, 2005). Three classes, i.e., A, C and D, are identified as serine beta-lactamases because a serine residue is positioned in their active sites (Hall and Barlow, 2004). Class B beta-lactamases, also termed metallo beta-lactamases, require Zn^{2+} for activity and comprise a clearly different group from the serine beta-lactamases (Garau et al., 2004). **E**xtended **s**pectrum **b**eta-**l**actamases (ESBLs), defined as those able to hydrolyze penicillins, aztreonam, and cephalosporins up to the third generation (Naas et al., 2008), can be found in classes A and D. Various beta-lactamases of class C have a substrate range similar to those of ESBLs, but are less sensitive to inhibitors like clavulanic acid (Jacoby, 2009). Class B beta-lactamases, on the other hand, are able to hydrolyze almost all beta-lactams in use, and their ability to confer carbapenems resistance is particularly alarming (Garau et al., 2004).

Genes coding for members of all classes have been found in transferable plasmids (Coudron et al., 2000; Pérez-Pérez and Hanson, 2002; de Been et al., 2014), and some have been found to be associated with class 1 integrons (Bonnet, 2004). This emphasizes the relevance of these enzymes for spread of resistant traits to diverse bacteria including pathogens.

Presently, over a thousand beta-lactamase genes sequences can be retrieved from the Comprehensive Antibiotic Resistance Database (CARD). Beta-lactamases are among the most diverse and widespread groups of **a**ntibiotic **r**esistance **d**etermining **g**enes (ARDGs). Most variants were obtained from clinical isolates (Coudron et al., 2000; Pérez-Pérez and Hanson, 2002; Jeong et al., 2004; de Been et al., 2014; Shaikh et al., 2015). However, it has been shown that non-pathogenic bacteria present in environments with reduced anthropogenic influence could be considered as natural reservoirs for beta-lactamase coding genes (Allen et al., 2009). Some of these genes may be different, both structurally and functionally from those obtained and evolved in clinical environments.

Studying these genes in a pristine environment offers the opportunity to minimize the anthropogenic factors involved in the complex evolutionary history of antibiotic resistance. The Antarctic regions studied are certainly heterogeneous, offering different study frames to our exploration. The Antarctic Polar Frontal Zone is classified into three biogeographically distinct regions: continental Antarctica, sub-Antarctica and maritime Antarctica. Continental Antarctica is most of the continent. It has an extremely cold and dry climate and is exposed to strong winds and intense UV radiation. Some sites of this region, like McMurdo Dry Valleys, have been ice-free for several years. Maritime Antarctica includes north-western areas of the Antarctic Peninsula and surrounding islands, including South Shetland Islands. This region has a less rigorous climate, with more humidity and higher average temperatures compared with continental Antarctica. King George Island, in this region, is home to numerous marine mammals and birds, and one of the most visited sites on the continent, with ten research stations in operation.

The objective of this study was to survey 14 microbial mats metagenomes from different sites along a latitudinal gradient in Antarctica, specifically for the presence of beta-lactamase coding genes (BLCGs). Microbial mats are common features of Antarctic environments, representing native and diverse microbial communities. Samples were collected from sites exposed to diverse environmental conditions and with different levels of human and animal influence. The spectrum of antibiotic resistance genes present in these metagenomes may indicate a greater diversity than that currently described.

2 Materials and methods

2.1 Samples collection

Microbial mats samples were collected from different

Antarctic sites representing a latitudinal gradient. Table 1 summarizes the sites and geographical coordinates.

2.2 Shotgun metagenomics sequence processing

Metagenomic libraries were prepared with the Nextera DNA Flex library prep kit (Illumina, San Diego, CA, USA) where fragments of total DNA (1 μg) were inserted into vectors and sequenced with whole genome sequencing technology (HiSeq2 × 150), at the Yale Keck Center for Genomic Sciences. A mean of 7.8±2 Gb of data were obtained for each metagenome, for a total of 109 Gbp of sequenced DNA. Raw reads were quality filtered and *de novo* assembled using IDBA-UD, with k-mer lengths between 120 and 150 bp (Peng et al., 2012). N50 values, from 2500 to 9000 bp are presented in Table S1 in Supplementary material. The set of contigs generated, ca. 878 Mbp, was used as input for searching beta-lactamase genes.

2.3 Identification of beta-lactamase coding genes (BLCGs)

Nucleotide and translated amino acid sequences of known BLCGs were obtained from CARD database. A Hidden Markov Model (HMM) based on amino acid sequences was generated for each class. These models were used as queries for the search of BLCGs in each metagenome.

Selected contigs containing a presumptive BLCG, longer than 2000 bp, were considered for further analysis. These contigs were then submitted to RAST server (Rapid Annotation using Sub-system Technology) (Aziz et al., 2008) for annotation. Contigs in which RAST server annotated a beta-lactamase were then considered for the following step. Lengths of the longest and the shortest known beta-lactamases were considered as limits for putative genes. Hits whose lengths were beyond these limits were not evaluated.

2.4 Construction of phylogenetic trees for beta-lactamases

Sequences were aligned using Seaview software (v. 4.6.3) (PRABI-Doua, Pôle Rhône-Alpes de Bioinformatique Site Doua, France). The alignment was constructed using Muscle algorithm. Phylogenetic trees were also constructed with Seaview software using the PhyML algorithm.

One phylogenetic tree for each class was constructed including previously described sequences for beta-lactamases from CARD database. This was done in order to identify clusters and to select representative sequences to be used as references for the construction of a final tree for each class. Sequences selected as reference for class A tree were variants of CARB, CfxA, CTX-M, GES, IMI, KPC, LEN, LRA, NPS, OKP-B, OXY, PER, SHV, TEM, TLE and VEB beta-lactamases. For class B tree, 6 sequences were selected and included variants of cphA, IND, LRA, MUS, NDM and VIM beta-lactamases. Eight sequences were included for class C with variants of ACC, ACT, CMY, DHA, FOX, LRA, OCH and PDC sequences. For class D tree 8 reference sequences were selected, 7 of which belonged to different clusters of the OXA variants and the remaining sequence was a LCR type. Reference trees are not shown.

Four multifasta files were built for each class, including environmental beta-lactamase sequences obtained from the 14 metagenomes and selected reference sequences. These files were used to construct alignments and their respective trees.

Table 1 Geographic location of sampling sites

Sample	Geographic reference	Latitude	Longitude
Sample 1	King George Island (Fildes Peninsula)	62°09′31″ S	58°56′31″ W
Sample 2	King George Island (Fildes Peninsula)	62°09′59″ S	58°58′33″ W
Sample 3	King George Island (Fildes Peninsula)	62°12′14″ S	58°57′16″ W
Sample 4	King George Island (Fildes Peninsula)	62°10′00″ S	58°58′34″ W
Sample 5	King George Island (Potter Peninsula)	62°14′35″ S	58°40′39″ W
Sample 6	King George Island (Potter Peninsula)	62°14′34″ S	58°40′26″ W
Sample 7	Antarctic Peninsula (Trinity Peninsula)	63°28′13″ S	57°12′30″ W
Sample 8	Antarctic Peninsula (Danco Coast)	64°09′22″ S	60°57′30″ W
Sample 9	Antarctic Peninsula (Fallieres Coast)	68°07′45″ S	67°06′20″ W
Sample 10	McMurdo Dry Valleys	78°01′24″ S	163°55′03″ E
Sample 11	McMurdo Dry Valleys	78°01′23″ S	163°54′56″ E
Sample 12	McMurdo Dry Valleys	78°01′23″ S	163°54′07″ E
Sample 13	McMurdo Dry Valleys	78°01′30″ S	164°06′02″ E
Sample 14	McMurdo Dry Valleys	77°39′40″ S	163°05′31″ E

2.5 Calculation of Shannon diversity index

Sequences from each class of beta-lactamase were clustered using the CD-HIT program (Fu et al., 2012). Sequences that grouped in the same cluster were treated as members of the same taxonomic unit for index calculation purposes. Grouping by CD-HIT was done using word size of 3, and a sequence identity threshold value of 0.5. Indexes were calculated for each class of beta-lactamases retrieved from our metagenomes and from the CARD database, separately.

3 Results and discussion

A total of 1341 BLCG sequences were obtained from the CARD database. HMM were constructed with 660, 163, 226 and 292 amino acid sequences retrieved from classes A, B, C and D, respectively. Figure S1 (supplementary material) shows the origin of different sequences used. Most of these sequences belong to genera that include human pathogenic bacteria, showing the driving force in discovery and description of beta-lactamases. The main driving force for study of these sequences has mainly been involved in the study of human pathogens and clinical environments.

According to our analysis, the 14 metagenomes analyzed contain classes A, B and C beta-lactamase genes. Class D genes were only identified in 11 of these metagenomes (Tables 2 and 3). The most abundant was class C with 306 confirmed hits (46.8%), followed by class B with 232 hits (35.5%) and class A with 93 hits (14.2%). Only 23 class D beta-lactamases were found (3.5%) (Table 3). As shown in Tables 2 and 3, before filtering the sequences using the criteria mentioned in methods section, the number of hits were 260, 765, 708 and 188 for classes A, B, C and D, respectively. Therefore, a total of 1921 beta-lactamase primary hits were identified initially, however, only 654 hits remained after filtering representing 34% of the total. It is certainly possible that some of these primary hits, which were not further considered, would be beta-lactamase coding genes. The sequence-based approach used in this study relies on HMM generated with beta-lactamase sequences obtained from the CARD database. These sequences are usually found in clinical environments (Boolchandani et al., 2019). This would indicate that the number and diversity of ARDGs in our samples could be higher than values shown as final results of this work. The number of class D beta-lactamases after filtering had the highest reduction, representing only 12.2% of the set of primary hits of this class.

We found an average of 0.75, 1.95, 2.62 and 0.22 class A, B, C and D genes per 1000 contigs analyzed, respectively. Contigs carrying class B and D genes were more abundant in metagenome 9 (Antarctic Peninsula, Fallieres Coast). In this community, we found 4.65 ‰ and 6.02 ‰ of contigs contain class B and D beta-lactamase genes, respectively. The highest proportion of beta-lactamase genes per 1000 obtained for class A was 1.40 ‰ in metagenome 11 (Table 2) and for class D was 0.53 ‰ in metagenome 2 (Table 3).

In a previous study, a meta-analysis of metagenomes from diverse environments, found that about 83% of beta-lactamase hits belonged to class A, followed by 10% class

Table 2 Number and per mille of beta-lactamase hits found for classes A and B in each metagenome

Sample[a)]	Class A			Class B		
	Primary Hits (*n*)	Confirmed Hits (*n*)	Confirmed Hits (‰)	Primary Hits (*n*)	Confirmed Hits (*n*)	Confirmed Hits (‰)
Metagenome 1 (14344)	4	2	0.139	17	6	0.418
Metagenome 2 (3760)	10	3	0.798	33	5	1.330
Metagenome 3 (3724)	8	1	0.269	14	6	1.611
Metagenome 4 (4706)	7	4	0.850	28	8	1.700
Metagenome 5 (15035)	42	13	0.865	78	21	1.397
Metagenome 6 (11962)	32	8	0.669	66	15	1.254
Metagenome 7 (9291)	21	8	0.861	67	16	1.722
Metagenome 8 (10127)	39	12	1.185	72	19	1.876
Metagenome 9 (7315)	14	5	0.684	64	34	4.648
Metagenome 10 (6973)	20	7	1.004	53	15	2.151
Metagenome 11 (9269)	22	13	1.403	68	24	2.589
Metagenome 12 (8972)	7	2	0.223	76	15	1.672
Metagenome 13 (7187)	13	5	0.696	58	16	2.226
Metagenome 14 (11580)	21	10	0.864	71	32	2.763

Notes: a) Number in parenthesis in the first column indicates the total number of contigs in each metagenome. About Class A, the total number of Primary Hits is 260; the total number of Confirmed Hits is 93; about Class B, the total number of Primary Hits is 765; the total number of Confirmed Hits is 232.

Table 3 Number and per mille of beta-lactamase hits found for classes C and D in each metagenome

Sample[a)]	Class C			Class D		
	Primary Hits (*n*)	Confirmed Hits (*n*)	Confirmed Hits (‰)	Primary Hits (*n*)	Confirmed Hits (*n*)	Confirmed Hits (‰)
Metagenome 1 (14344)	19	12	0.837	0	0	0
Metagenome 2 (3760)	27	10	2.660	9	2	0.532
Metagenome 3 (3724)	17	7	1.880	4	1	0.269
Metagenome 4 (4706)	21	10	2.125	8	0	0
Metagenome 5 (15035)	58	19	1.264	22	1	0.067
Metagenome 6 (11962)	75	26	2.174	28	1	0.084
Metagenome 7 (9291)	36	14	1.507	21	2	0.215
Metagenome 8 (10127)	70	26	2.567	26	5	0.494
Metagenome 9 (7315)	86	44	6.015	10	3	0.410
Metagenome 10 (6973)	51	23	3.298	14	2	0.287
Metagenome 11 (9269)	64	32	3.452	14	3	0.324
Metagenome 12 (8972)	62	19	2.118	9	2	0.223
Metagenome 13 (7187)	43	25	3.479	8	1	0.139
Metagenome 14 (11580)	79	39	3.368	15	0	0

Notes: a) Numbers in parenthesis in the first column indicates the total number of contigs in each metagenome. About Class C, the total number of Primary Hits is 708; the total number of Confirmed Hits is 306; about Class D, the total number of Primary Hits is 188; the total number of Confirmed Hits is 23.

C genes (Nesme et al., 2014). Comparing our results with those obtained previously using samples from other environments, we only detected one coincidence, related to the low abundance of contigs containing class D beta-lactamase genes.

The 232 hits of class B beta-lactamases found in the 14 metagenomes imply that 35.5% of total hits belong to metallo beta-lactamases. This proportion is higher compared with representation of this class in CARD database sequences (12%). This result should be considered with caution, since we are comparing gene abundance detected in Antarctic communities with the abundance from a global database. This in turn reflects the abundance of genes detected in diverse environments, which are mainly of clinical origin. The existence of an extensive pool of class B beta-lactamases could be alarming and should be considered in the development of guidelines for antibiotics use. Most of known metallo beta-lactamases are broad-spectrum and not responsive to inhibitors such as clavulanic acid (Lisa et al., 2017). However, this might not be an immediate problem in this kind of environment, although the presence of these "proto-resistance genes" must be taken into account to avoid their selection and evolution to resistance genes that could result in serious clinical problems.

The abundance of class C beta-lactamase genes (46.8%) found in metagenomes is higher compared with the proportion of class C genes recovered from CARD database (16.9%). The opposite situation is found for class A genes, with abundances of 14.2% and 49.2% from the Antarctic microbial mat metagenome analysis and CARD database, respectively. However, it can be argued that the number of presently characterized beta-lactamases is influenced by evolutionary radiation promoted by selective pressures that result from use of antibiotics. In this way, since class C are primarily cephalosporinases (Jacoby, 2009), the specific selection pressure began later compared with that applied to class A beta-lactamases. However, many class C coding genes have been shown to be cryptic or unexpressed and thus no increased resistance should result from their incorporation to a genome (Hall and Barlow, 2004). Therefore a higher number of class C beta-lactamases does not necessarily mean higher resistance potential.

Most known class D beta-lactamases are specialized in hydrolyzing a subset of penicillins, oxacillins (June et al., 2014) and others can hydrolyze carbapenems (Al Bayssari et al., 2015). Due to these substrate specificities, it is not surprising that class D is the least abundant in the Antarctic environment. However, this group of beta-lactamases has expanded rapidly since introduction of carbapenems (Evans and Amyes, 2014).

Analysis of the phylogenetic tree of class A beta-lactamases allows for identification of 4 clusters of candidate genes (Fig. 1). The rest of the sequences are distributed across the tree. Clusters have 20, 22, 16 and 7 sequences from different metagenomes. Reference sequences are not located in these clusters, although 12 grouped together in a broader group that includes three candidate beta-lactamases, two from metagenome 10 and one from metagenome 11. Interestingly, some class A beta-lactamases obtained were found to be closely related despite being found in metagenomes obtained from geographically distant samples. That is the case, e.g., of

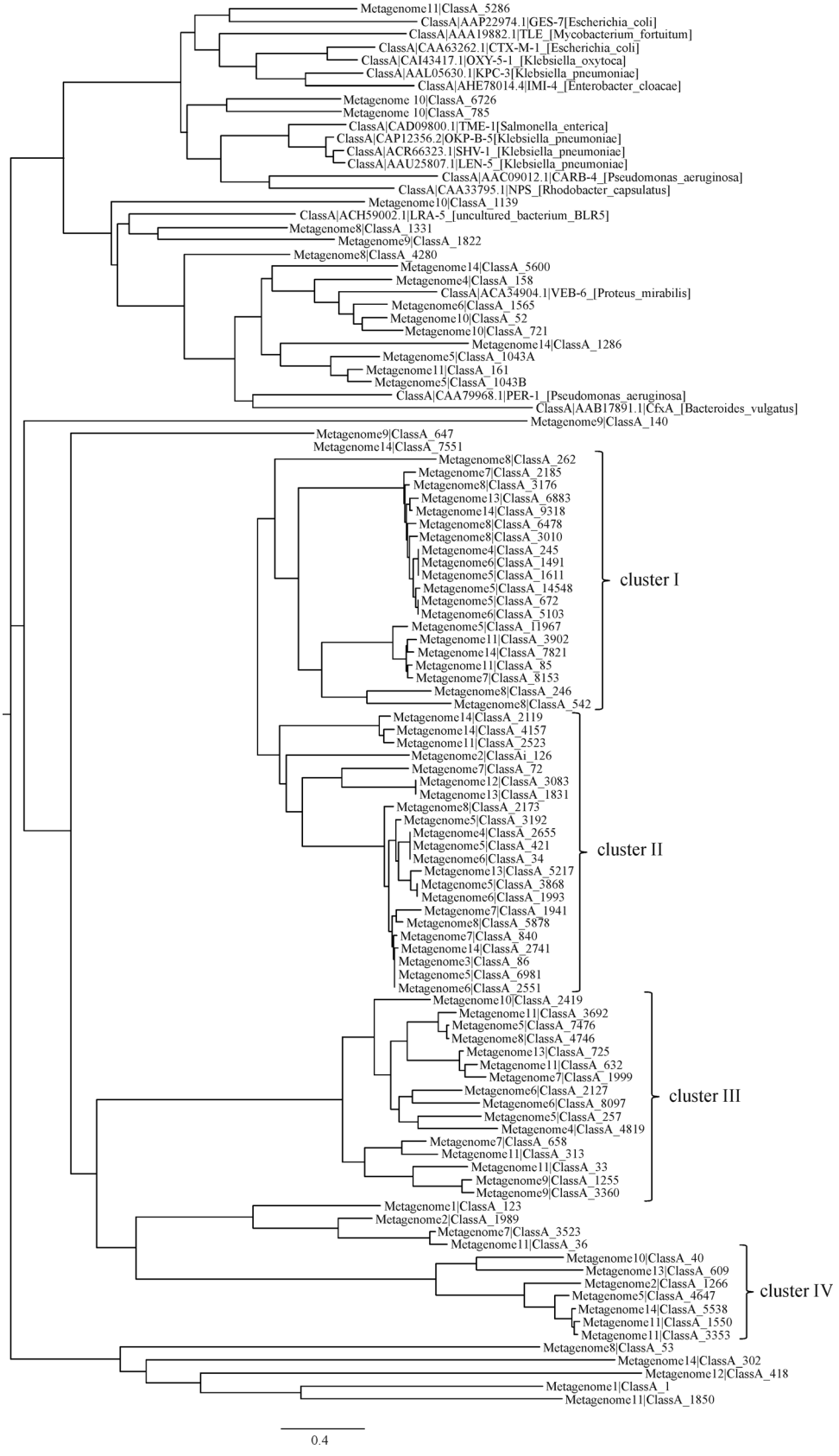


Fig. 1 Phylogenetic tree constructed with deduced amino acid sequences of class A beta-lactamases identified in Antarctic microbial mat metagenomes. Representative reference sequences were also included in the tree. The tree was constructed using Neighbor Joining algorithm.

the hit found in contig 3692 of metagenome 11, one in contig 7476 of metagenome 5 and the hit in contig 4746 of metagenome 8, all located in cluster III. There are examples of this relatedness in all clusters identified, especially in cluster I, including a sub-cluster of 12 beta-lactamases. All these sequences are closely related and they were obtained from 7 different metagenomes. Furthermore, in that cluster there is one beta-lactamase that was found in metagenomes 4, 5 and 6. This would imply that most class A beta-lactamases from the Antarctic environment differ markedly from presently known beta-lactamases. Moreover, some of them are closely related between themselves despite their disparate origin. This should be taken into account when metagenome data are analyzed. As expected, databases have a bias toward genes that appear in commonly studied environments. These biases affect the strategies dependent on previously described sequences, like primer design for PCR detection. Antarctic environmental sequences are not well represented in databases and this could be why some previous studies may have underestimated the abundance of beta-lactamases (Segawa et al., 2013; Van Goethem et al., 2018).

Phylogenetic tree of class B beta-lactamases shows that query sequences are clustered in 9 major groups and in other minor ones (Fig. 2; see Fig. S2 for full tree). Five of the reference sequences are grouped in cluster VIII together with 11 candidate beta-lactamases from our samples, originating from 8 different metagenomes. Reference sequences of types VIM and NDM are grouped together, as well as type sequences IND and MUS. However, the first pair of these sequences clustered with a

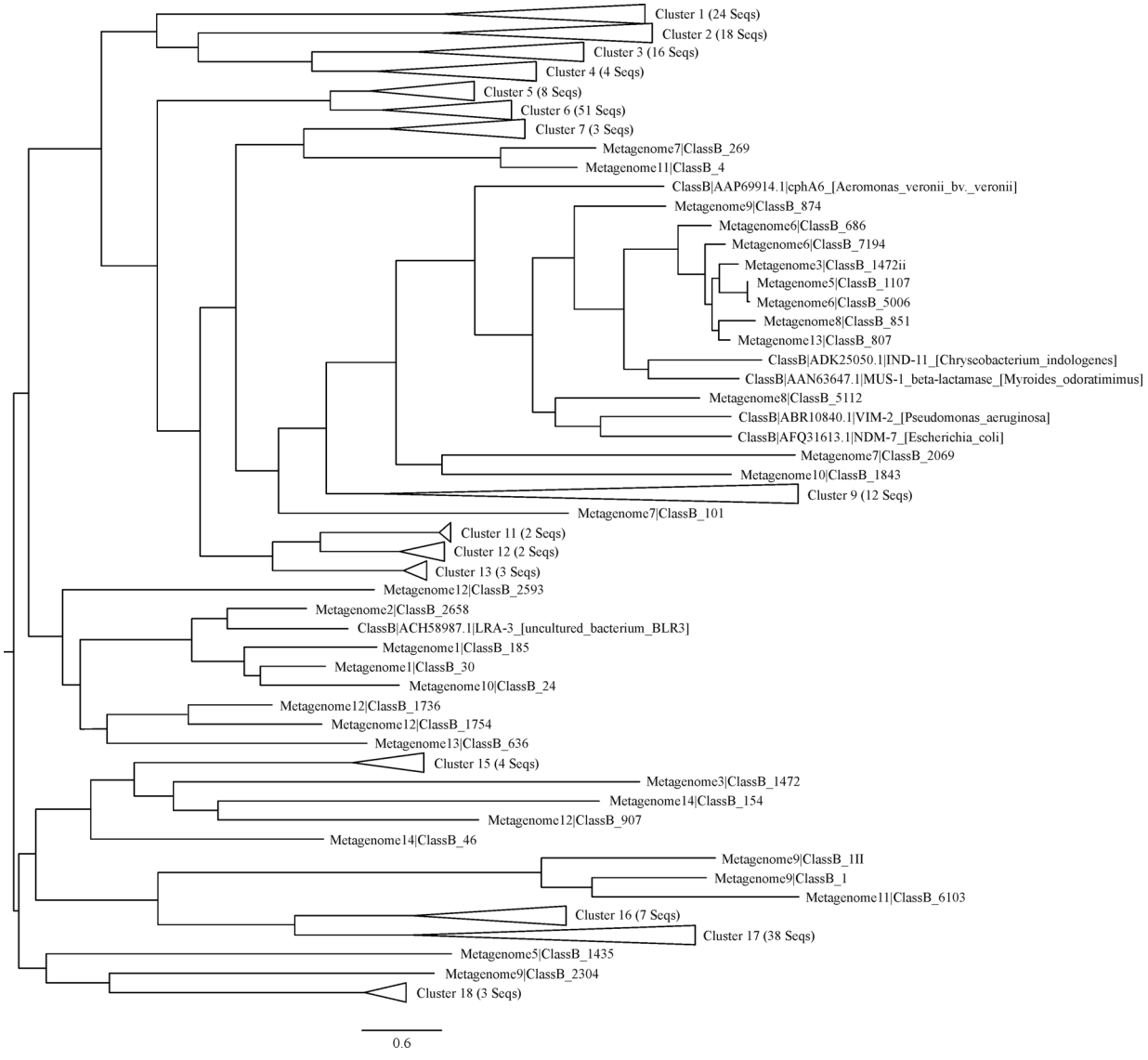


Fig. 2 Phylogenetic tree constructed with deduced amino acid sequences of class B beta-lactamases identified in Antarctic microbial mat metagenomes. Representative reference sequences were also included in the tree. The tree was constructed using Neighbor Joining algorithm.

metagenome 8 beta-lactamase, while the second pair grouped with 8 other beta-lactamases. This indicates that among environmental metallo beta-lactamases, there is a subset that is closely related to those described in clinical environments. This underscores the importance of studying the natural pool of ARDGs and may, at least partially, explain the rapid emergence and dissemination of certain variants such as New Delhi metallo beta-lactamase (Bush et al., 2011). Reference sequence type LRA, which is a group of beta-lactamases derived from uncultured bacteria, was located in a minor cluster and is related to a beta-lactamase from metagenome 2. Cluster VI is the largest and contains 51 beta-lactamases, some of which are highly related. This cluster contains sequences from all metagenomes sampled except from metagenomes 1, 7 and 10.

The phylogenetic tree constructed with class C beta-lactamases shows 5 differentiated clusters with two sequences outside of these clusters (Fig. 3; see Fig. S3 for full tree). Sequences outside clusters belong to metagenomes 9 and 14. The 8 reference sequences from cluster IV, which contains only reference sequences. Cluster V is the mostly related to reference cluster (IV), and contains 22 sequences from 8 metagenomes (metagenome 2, 6, 8, 9, 10, 11, 12 and 14). The remaining sequences seem to be distantly related to reference beta-lactamases. Cluster III is the largest, with a total of 233 sequences from all metagenomes. These results show that there exists a great reservoir of uncharacterized class C beta-lactamases, most of which are more related to each other than to previously described beta-lactamases. These results support the idea that natural environments and non-pathogenic bacteria harbor a myriad of uncharacterized ARDGs (Allen et al., 2010). Uncharacterized beta-lactamases, particularly those distantly related to those already characterized, may differ substantially in substrates affinities or kinetics. As observed for class A, some class C beta-lactamases identical sequences (indicated with arrows in supplementary Fig. S3), were obtained from different metagenomes. Metagenomes 5 and 6 share six different beta-lactamases. One particular sequence was found in metagenomes 4, 5, 6 and 7. Thus, this beta-lactamase may be common in these environments.

Six clusters can be identified in the class D phylogenetic tree (Fig. 4). Two clusters contain only reference sequences. Cluster III includes only two sequences, reference gene type OXA-209 and the sequence found in contig 21 of metagenome 9. This is the only candidate class D sequence found in the 14 metagenomes closely related to a reference sequence. OXA-209 was originally described in *Riemerella anatipestifer,* a bird pathogen

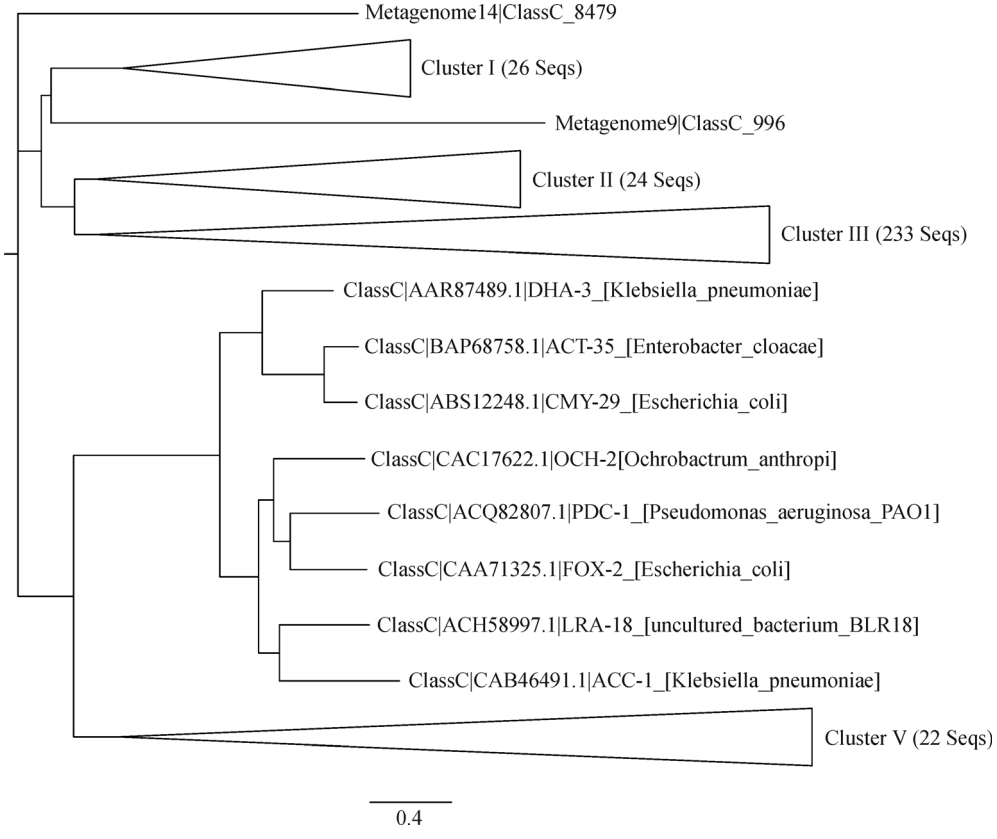


Fig. 3 Phylogenetic tree constructed with deduced amino acid sequences of class C beta-lactamases identified in Antarctic microbial mat metagenomes. Representative reference sequences were also included in the tree. The tree was constructed using Neighbor Joining algorithm.

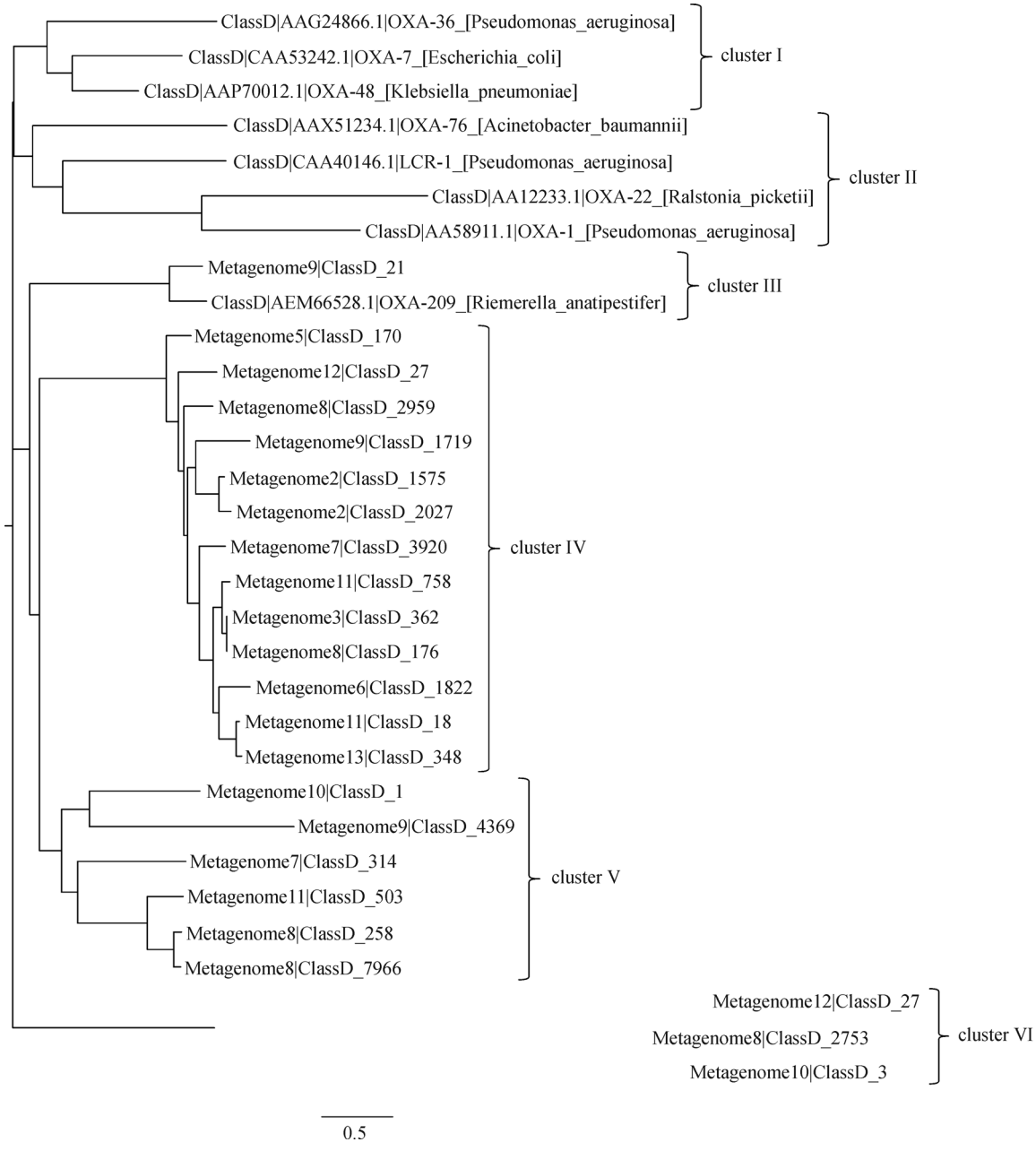


Fig. 4 Phylogenetic tree constructed with deduced amino acid sequences of class D beta-lactamases identified in Antarctic microbial mat metagenomes. Representative reference sequences were also included in the tree. The tree was constructed using Neighbor Joining algorithm.

(Chen et al., 2012). The largest assembly is cluster IV, which includes 13 beta-lactamases. In this cluster, it is also observed that the beta-lactamase sequence found in contig 362 of metagenome 3 was also found in metagenome 8. Therefore, the observations made for the other classes of beta-lactamases also apply to class D despite being the less abundant class in our metagenomes.

Comparing Shannon diversity index of database beta-lactamases and Antarctic sequences, it was clear that the microbial mats from Antarctica hold a higher diversity (Table 4). Only for class D beta-lactamases the diversity index of database sequences was higher than that of beta-lactamases recovered from Antarctic metagenomes. Interestingly, the 93 sequences of class A beta-lactamases found in metagenomes were more diverse than the 660 from the CARD database.

Previous studies provided evidence on the evolution and emergence of antibiotic resistance genes from environmental bacteria. Genes encoding enzymes that modify antibiotics (e.g. beta-lactamase genes) have generally evolved to confer other activities (Martínez, 2008; Wright, 2010). However, due to substrate similarity they can also

Table 4 Shannon diversity indexes for each class of beta-lactamases for sequences retrieved from the Antarctic metagenomes and from CARD database

Class	Shannon index of metagenome sequences	Shannon index of database sequences
Class A[a)]	2.94 (n = 93)	1.35 (n = 660)
Class B[a)]	3.32 (n = 232)	2.30 (n = 163)
Class C[a)]	4.85 (n = 306)	0.72 (n = 226)
Class D[a)]	1.54 (n = 23)	1.66 (n = 292)

Notes: a) The n indicates the number of sequences considered for index calculation.

confer a low antimicrobial resistance activity. The high diversity of beta-lactamase genes discovered in pristine Antarctic microbial mats sheds evidence to their natural distribution and occurrence, as means of bacteria-bacteria interactions. When these genes are selected, due to antibiotic presence, they can accumulate mutations that eventually enhance this activity. Another important aspect that can lead to formation of clinically-relevant ARGDs is their incorporation into mobile genetic elements. Such events may change the biochemical regulation and context in which these genes are expressed. The positive selection for these traits is important for acceleration of such evolutionary events and thus leads to drug resistant phenotypes.

The methods used for ARDGs' detection and their applications to study metagenomes are matters of intense discussion and research efforts. It is now recognized that a high number of genes identified in non-culturable and environmental bacteria are not well represented in available databases (Allen et al. 2009). To avoid bias toward clinically relevant genes and to describe novel ARDGs different strategies are being developed (Wallace et al. 2017; Berglund et al. 2019).

One particular limitation of sequence-based studies is the possibility that some selected sequences do not confer the predicted resistance (Quince et al., 2017). However, there are records of metagenome-predicted genes whose synthesis, cloning and expression in *Escherichia coli* strains resulted in resistant phenotypes (Ruppé et al., 2019). Even though only a functional analysis assures the activity of a sequence-predicted ARDG, it is well recognized as valid the use of HMMs to recover sequences with low identity scores but with a similar functionality (Gibson et al. 2015). To reduce the potential selection of false positives, we incorporated a highly restrictive approach to select candidate BLCGs. In this study, the search and filtering pipeline has been restrictive enough to discard most non-beta-lactamase genes.

The Protocol on Environmental Protection to the Antarctic Treaty (Madrid Protocol) is part of the Antarctic Treaty System. The objective of this Protocol is protection of the Antarctic environment and dependent and associated ecosystems. Then, due to the environmental protection of this continent, a low selection toward resistant phenotypes is expected. However, there are certain studies that present evidence of sewage pollution near research stations, as the main contributors to the release of antibiotics (Hughes and Thompson, 2004). It is important to have strong control on this due to the highly diverse natural reservoir of ARDGs present in Antarctica.

4 Conclusions

We present evidence that the Antarctic environment can be an important reservoir of beta-lactamase genes. Our results evidence a wider diversity of candidate beta-lactamase genes than the expected based on existing databases. Sequence identity and coverage of the genes found are similar to most previously described beta-lactamase genes which can result in clinical problems. Assessing the diversity and distribution of bacterial antibiotic resistance strategies will have increasing clinical relevance to design effective public health strategies.

Acknowledgements Authors acknowledge the following agencies that were involved in sampling: NSF-USA (Dry Valleys), DNA-Argentina (Peninsula), IAU-Uruguay (Maritime). We also would like to thank Osiris Gaona for technical support and Paul Gill for discussion and technical reading of the manuscript. AMEXCID-Mexico and AUCI-Uruguay supported the study through an international cooperation project (PNUD URU/113). The work was also supported by IAU and PEDECIBA-Biología (Programa de Ciencias Básicas).

Electronic Supplementary Material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11783-019-1128-1 and is accessible for authorized users.

References

Al Bayssari C, Olaitan A O, Dabboussi F, Hamze M, Rolain J M (2015). Emergence of OXA-48-producing *Escherichia coli* clone ST38 in fowl. Antimicrobial Agents and Chemotherapy, 59(1): 745–746

Allen H K, Donato J, Wang H H, Cloud-Hansen K A, Davies J, Handelsman J (2010). Call of the wild: Antibiotic resistance genes in natural environments. Nature Reviews. Microbiology, 8(4): 251–259

Allen H K, Moe L A, Rodbumrer J, Gaarder A, Handelsman J (2009). Functional metagenomics reveals diverse β-lactamases in a remote Alaskan soil. The ISME Journal, 3(2): 243–251

Aziz R K, Bartels D, Best A A, DeJongh M, Disz T, Edwards R A, Formsma K, Gerdes S, Glass E M, Kubal M, Meyer F, Olsen G J,

Olson R, Osterman A L, Overbeek R A, McNeil L K, Paarmann D, Paczian T, Parrello B, Pusch G D, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O (2008). The RAST Server: rapid annotations using subsystems technology. BMC Genomics, 9(1): 75–90

Babic M, Hujer A M, Bonomo R A (2006). What's new in antibiotic resistance? Focus on beta-lactamases. Drug Resistance Updates: Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy, 9(3): 142–156

Bennett J W, Chung K T (2001). Alexander Fleming and the discovery of penicillin. Advances in Applied Microbiology, 49: 163–184

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): 52–66

Bonnet R (2004). Growing group of extended-spectrum β-lactamases: The CTX-M enzymes. Antimicrobial Agents and Chemotherapy, 48 (1): 1–14

Boolchandani M, D'Souza A W, Dantas G (2019). Sequencing-based methods and resources to study antimicrobial resistance. Nature Reviews. Genetics,

Bush K, Courvalin P, Dantas G, Davies J, Eisenstein B, Huovinen P, Jacoby G A, Kishony R, Kreiswirth B N, Kutter E, Lerner S A, Levy S, Lewis K, Lomovskaya O, Miller J H, Mobashery S, Piddock L J, Projan S, Thomas C M, Tomasz A, Tulkens P M, Walsh T R, Watson J D, Witkowski J, Witte W, Wright G, Yeh P, Zgurskaya H I (2011). Tackling antibiotic resistance. Nature Reviews. Microbiology, 9(12): 894–896

Chen Y P, Lee S H, Chou C H, Tsai H J (2012). Detection of florfenicol resistance genes in *Riemerella anatipestifer* isolated from ducks and geese. Veterinary Microbiology, 154(3–4): 325–331

Coudron P E, Moland E S, Thomson K S (2000). Occurrence and detection of AmpC beta-lactamases among Escherichia coli, Klebsiella pneumoniae, and Proteus mirabilis isolates at a veterans medical center. Journal of Clinical Microbiology, 38(5): 1791–1796

Davies J, Davies D (2010). Origins and evolution of antibiotic resistance. Microbiology and Molecular Biology Reviews : MMBR, 74(3): 417–433

de Been M, Lanza V F, de Toro M, Scharringa J, Dohmen W, Du Y, Hu J, Lei Y, Li N, Tooming-Klunderud A, Heederik D J, Fluit A C, Bonten M J, Willems R J, de la Cruz F, van Schaik W (2014). Dissemination of cephalosporin resistance genes between *Escherichia coli* strains from farm animals and humans by specific plasmid lineages. PLOS Genetics, 10(12): e1004776–e1004793

Evans B A, Amyes S G (2014). OXA β-lactamases. Clinical Microbiology Reviews, 27(2): 241–263

Fu L, Niu B, Zhu Z, Wu S, Li W (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics (Oxford, England), 28(23): 3150–3152

Garau G, García-Sáez I, Bebrone C, Anne C, Mercuri P, Galleni M, Frère J M, Dideberg O (2004). Update of the standard numbering scheme for class B β-lactamases. Antimicrobial Agents and Chemotherapy, 48(7): 2347–2349

Gibson M K, Forsberg K J, Dantas G (2015). Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. The ISME Journal, 9(1): 207–216

Hall B G, Barlow M (2004). Evolution of the serine β-lactamases: Past, present and future. Drug Resistance Updates: Reviews and Commentaries in Antimicrobial and Anticancer Chemotherapy, 7 (2): 111–123

Hall B G, Barlow M (2005). Revised Ambler classification of β-lactamases. The Journal of Antimicrobial Chemotherapy, 55(6): 1050–1051

Högberg L D, Heddini A, Cars O (2010). The global need for effective antibiotics: Challenges and recent advances. Trends in Pharmacological Sciences, 31(11): 509–515

Hughes K A, Thompson A (2004). Distribution of sewage pollution around a maritime Antarctic research station indicated by faecal coliforms, *Clostridium perfringens* and faecal sterol markers. Environmental Pollution (Barking, Essex: 1987), 127(3): 315–321

Jacoby G A (2009). AmpC β-lactamases. Clinical Microbiology Reviews, 22(1): 161–182

Jeong S H, Bae I K, Lee J H, Sohn S G, Kang G H, Jeon G J, Kim Y H, Jeong B C, Lee S H (2004). Molecular characterization of extended-spectrum beta-lactamases produced by clinical isolates of *Klebsiella pneumoniae* and *Escherichia coli* from a Korean nationwide survey. Journal of Clinical Microbiology, 42(7): 2902–2906

June C M, Vallier B C, Bonomo R A, Leonard D A, Powers R A (2014). Structural origins of oxacillinase specificity in class D β-lactamases. Antimicrobial Agents and Chemotherapy, 58(1): 333–341

Lisa M N, Palacios A R, Aitha M, González M M, Moreno D M, Crowder M W, Bonomo R A, Spencer J, Tierney D L, Llarrull L I, Vila A J (2017). A general reaction mechanism for carbapenem hydrolysis by mononuclear and binuclear metallo-β-lactamases. Nature Communications, 8(1): 538–549

Martínez J L (2008). Antibiotics and antibiotic resistance genes in natural environments. Science, 321(5887): 365–367

Naas T, Poirel L, Nordmann P (2008). Minor extended-spectrum β-lactamases. Clinical Microbiology and Infection, 14(Suppl 1): 42–52

Nesme J, Cécillon S, Delmont T O, Monier J M, Vogel T M, Simonet P (2014). Large-scale metagenomic-based study of antibiotic resistance in the environment. Current Biology: CB, 24(10): 1096–1100

Peng Y, Leung H C, Yiu S M, Chin F Y (2012). IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinformatics (Oxford, England), 28(11): 1420–1428

Pérez-Pérez F J, Hanson N D (2002). Detection of plasmid-mediated AmpC β-lactamase genes in clinical isolates by using multiplex PCR. Journal of Clinical Microbiology, 40(6): 2153–2162

Quince C, Walker A W, Simpson J T, Loman N J, Segata N (2017). Shotgun metagenomics, from sampling to analysis. Nature Biotechnology, 35(9): 833–844

Ruppé E, Ghozlane A, Tap J, Pons N, Alvarez A S, Maziers N, Cuesta T, Hernando-Amado S, Clares I, Martínez J L, Coque T M, Baquero F, Lanza V F, Máiz L, Goulenok T, de Lastours V, Amor N, Fantin B, Wieder I, Andremont A, van Schaik W, Rogers M, Zhang X, Willems R J L, de Brevern A G, Batto J M, Blottière H M, Léonard P, Léjard V, Letur A, Levenez F, Weiszer K, Haimet F, Doré J, Kennedy S P, Ehrlich S D (2019). Prediction of the intestinal resistome by a three-dimensional structure-based method. Nature Microbiology, 4(1): 112–123

Segawa T, Takeuchi N, Rivera A, Yamada A, Yoshimura Y, Barcaza G,

Shinbori K, Motoyama H, Kohshima S, Ushida K (2013). Distribution of antibiotic resistance genes in glacier environments. Environmental Microbiology Reports, 5(1): 127–134

Shaikh S, Fatima J, Shakil S, Rizvi S M D, Kamal M A (2015). Antibiotic resistance and extended spectrum beta-lactamases: Types, epidemiology and treatment. Saudi Journal of Biological Sciences, 22(1): 90–101

Van Goethem M W, Pierneef R, Bezuidt O K I, Van De Peer Y, Cowan D A, Makhalanyane T P (2018). A reservoir of 'historical' antibiotic resistance genes in remote pristine Antarctic soils. Microbiome, 6(1): 40–52

Wallace J C, Port J A, Smith M N, Faustman E M (2017). FARME DB: A functional antibiotic resistance element database. Database (Oxford), 2017: baw165–7

Wright G D (2010). Antibiotic resistance in the environment: A link to the clinic? Current Opinion in Microbiology, 13(5): 589–594