



# Biosynthetic gene clusters with biotechnological applications in novel Antarctic isolates from *Actinomycetota*

Pablo Bruna<sup>1,2</sup> · Kattia Núñez-Montero<sup>3,4</sup> · María José Contreras<sup>5</sup> · Karla Leal<sup>5</sup> · Matías García<sup>1,2,6</sup> · Michel Abanto<sup>2</sup>  · Leticia Barrientos<sup>3</sup>

Received: 28 December 2023 / Revised: 11 April 2024 / Accepted: 19 April 2024  
© The Author(s) 2024

**Abstract** *Actinomycetota* have been widely described as valuable sources for the acquisition of secondary metabolites. Most microbial metabolites are produced via metabolic pathways encoded by biosynthetic gene clusters (BGCs). Although many secondary metabolites are not essential for the survival of bacteria, they play an important role in their adaptation and interactions within microbial communities. This is how bacteria isolated from extreme environments such as Antarctica could facilitate the discovery of new BGCs with biotechnological potential. This study aimed to isolate rare *Actinomycetota* strains from Antarctic soil and sediment samples and identify their metabolic potential based on genome mining and exploration of biosynthetic gene clusters. To this end, the strains were sequenced using Illumina and Oxford Nanopore Technologies platforms. The assemblies were annotated and subjected to phylogenetic analysis. Finally, the BGCs present in each genome were identified using the antiSMASH tool, and the biosynthetic diversity of the *Micrococcaceae* family was evaluated. Taxonomic annotation revealed that seven strains were new and two were previously reported in the NCBI database. Additionally, BGCs encoding type III polyketide synthases (T3PKS), beta-lactones, siderophores, and non-ribosomal peptide synthetases (NRPS) have been identified, among others. In addition, the sequence similarity network showed a predominant type of BGCs in the family *Micrococcaceae*, and some genera were distinctly grouped. The BGCs identified in the isolated strains could be associated with applications such as antimicrobials, anticancer agents, and plant growth promoters, among others, positioning them as excellent candidates for future biotechnological applications and innovations.

## Key points

- Novel Antarctic rare *Actinomycetota* strains were isolated from soil and sediments
- Genome-based taxonomic affiliation revealed seven potentially novel species
- Genome mining showed metabolic potential for novel natural products

**Keywords** Actinomycetota · Biosynthetic gene clusters · Antarctic · Secondary metabolites · *Micrococcaceae*

✉ Michel Abanto  
mfabanto@gmail.com

✉ Leticia Barrientos  
leticia.barrientos@uautonoma.cl

<sup>1</sup> Programa de Doctorado en Ciencias mención Biología Celular y Molecular Aplicada, Universidad de La Frontera, Temuco, Chile

<sup>2</sup> Núcleo Científico y Tecnológico en Biorecursos (BIOREN), Universidad de La Frontera, Avenida Francisco Salazar, 01145 Temuco, Chile

<sup>3</sup> Facultad de Ciencias de la Salud, Instituto de Ciencias Aplicadas, Universidad Autónoma de Chile, Avenida Alemania 1090, Temuco, Chile

<sup>4</sup> Centro de Investigación en Biotecnología, Departamento de Biología, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica

<sup>5</sup> Facultad de Ingeniería, Instituto de Ciencias Aplicadas, Universidad Autónoma de Chile, Avenida Alemania 1090, Temuco, Chile

<sup>6</sup> Biocontrol Research Laboratory, Facultad de Ciencias Agropecuarias y Medioambiente, Universidad de La Frontera, Temuco, Chile

## Introduction

*Actinomycetota* have attracted increasing scientific interest in the vast field of microorganisms owing to their importance in the production of biologically and pharmaceutically relevant natural products (Matsumoto and Takahashi 2017; Albarano et al. 2020). Natural products derived from microorganisms have a broad spectrum of biological activities, making them promising candidates for the development of new drugs. However, the discovery of new bioactive compounds poses a great challenge to the scientific community due to the need to identify new sources of bioactive compounds with antimicrobial potential (Kumar et al. 2020; Núñez-Montero et al. 2019). Within the phylum *Actinomycetota*, *Micrococcaceae* plays a key role in the synthesis of bioactive compounds with a wide range of therapeutic properties. Exploration of these microorganisms and their natural products has opened new perspectives to address public health challenges (Borker et al. 2021; Núñez-Montero et al. 2019), such as the increasing antimicrobial resistance observed in pathogenic species.

Antarctica and its extreme environment have emerged as promising sources for the discovery of extraordinary taxa (Benaud et al. 2021) and the exploration of novel natural products (Silva et al. 2020). The extreme environmental conditions present, such as low temperatures, high radiation, and nutrient scarcity, have driven the development of unique adaptations in Antarctic microorganisms, contributing to richer biodiversity and high potential for bioactive compound discovery (Gummerlich et al. 2020). *Actinomycetota*, in particular, has been the subject of great attention in Antarctica because of their ability to produce secondary metabolites with pharmacological and biotechnological properties (Oyedoh et al. 2023) (Reis-Mansur et al. 2019). In recent years, Antarctic bacteria have been found to produce plant growth promoters, pigments, bioremediation agents, bioactive compounds, antibiofilm agents, nanoparticles, and enzymes for the food industry (Styczynski et al. 2022; Ramasamy et al. 2023).

On the other hand, advancements in DNA sequencing technology have revolutionized the study of microorganisms and their biosynthetic capabilities (Singh et al. 2021). Next-generation sequencing and related technologies have significantly accelerated and improved access to genetic information on microorganisms (Scherlach and Hertweck 2021), aiding the swift and accurate identification of biosynthetic gene clusters responsible for the production of natural products. These bioinformatics tools play a crucial role in the discovery (Tizabi et al. 2022) and characterization of new bioactive compounds (Albarano et al. 2020; Benaud et al. 2022) as well as in understanding the underlying molecular mechanisms of their biosynthesis.

While there is no certainty regarding the expression or silencing of the gene clusters identified through genomics, this initial exploration enabled us to detect their presence and structure. Mechanisms for the activation and overexpression of these genes could be proposed based on genome mining to obtain novel bioactive molecules. Therefore, genomic exploration of emerging microorganisms and their biosynthetic gene clusters lays the foundation for advanced metabolomics studies and characterization of novel biomolecules (Javed et al. 2021). In this context, the objective of our study was to isolate rare *Actinomycetota* strains from Antarctic soil and sediment samples and identify their metabolic potential based on genome mining and exploration of biosynthetic gene clusters.

## Material and methods

### Isolation and sample treatment of actinobacterial strains

In this study, soil and sediment samples were collected from Antarctica during the Chilean Antarctic Expedition conducted between 2019 and 2020. For isolation of rare *Actinomycetota* strains (less frequently isolated species), the samples were treated with high temperature, detergent solutions and plated with antibiotics supplementation to avoid the growth of most frequent species in the soil. To do this, 1 g of each sample was exposed to 100 °C for 60 min, followed by a 10<sup>-1</sup> dilution in each of the following chemical treatments: 1.5% phenol, 0.05% SDS, or 1% chloramine-T, incubating for 30 min at 30 °C. A 0.85% solution was used for the control samples. Dilutions of 10<sup>-2</sup> and 10<sup>-3</sup> of each pre-treated sample were cultured on soil agar (100 g/L Antarctic soil, 18 g/L agar) and oatmeal agar (60 g/L oatmeal, 18 g/L agar) supplemented with nalidixic acid (25 µg/mL). The cultures were incubated for 4 weeks at 15 °C. The resulting colonies were isolated and subsequently plated until purification in ISP-2 agar medium (4 g/L yeast extract, 4 g/L glucose, 10 g/L malt extract, and 20 g/L agar). Pure cultures were preserved by freezing at -80 °C in glycerol (20% v/v) until further use.

### Culture of actinobacterial strains

In this work, nine isolated strains deposited at the Colección Chilena de Cultivos Tipo – CCCT (Universidad de La Frontera, Chile) under accession codes CCCT 24.01–CCCT 24.09 were utilized. These strains were inoculated on ISP-2 agar plates, nutrient agar (Liofilchem, Italy), and R2A agar (Merck Millipore, Germany) and incubated at 15 °C for 2 weeks until suitable microbial growth was observed.

## DNA extraction from actinobacterial strains

From the above cultures, DNA extraction was performed using DNeasy UltraClean Microbial Kit (QIAGEN, Germany) following the manufacturer's instructions. Subsequently, DNA was quantified using a Qubit dsDNA HS Assay Kit (Invitrogen, USA), considering a minimum range of concentrations equal to or greater than 50 ng/μL. Finally, the integrity of the obtained DNA was visualized on an agarose gel in TAE 1 × 0.9% m/v buffer.

## Library preparation and whole genome sequencing

The samples were sequenced using Illumina and Oxford Nanopore Technologies (ONT). The Illumina library was prepared with 2 × 150 bp paired-end fragments on an Illumina NovaSeq platform. Quality control of the reads was assessed using FastQC v0.11.9 (Andrews 2010). Subsequently, adapters were cut from the reads and quality filtered using Fastp v0.20.0 tool (Chen et al. 2018) with the following parameters: `--detect_adapter_for_pe -f 12 -F 12`. For genome sequencing using long reads, the Rapid Sequencing Kit SQK-RBK004 (ONT) was used for library preparation, and sequencing was performed using an R9.4 flow cell (FLO-MIN106D) on a MinION Mk1C machine (ONT) using MinKNOW v4.3.7 software. Basecalling was performed using Guppy v5.0.12, in the fast mode. Quality control of the reads was assessed using the Nanoplot v1.40.0 tool (De Coster et al. 2018). Subsequently, the adapters were cut with Porechop v0.2.4, and sequences with a quality equal to or greater than 10 were filtered out using Nanofilt v2.8.0 (De Coster et al. 2018).

## Hybrid genome assembly

To achieve *de novo* hybrid assembly, short and long reads that passed quality control and filtering were assembled using Unicycler v0.4.8 tool (Wick et al. 2017). Subsequently, the hybrid genomes were polished with Medaka v1.2.3 (<https://github.com/nanoporetech/medaka>) using ONT reads. Then, the genomes were again polished with Polypolish v0.5.0 (Wick and Holt 2022) and poLCA v4.0.5 (Zimin and Salzberg 2020) using short reads. Finally, quality and contamination determination of the polished hybrid genomes were performed using Quast v5.0.2 and CheckM v1.1.3 (Gurevich et al. 2013; Parks et al. 2015).

## Genomic and taxonomic annotation

Annotation of genomic sequences was achieved using Prokka v1.14 (Seemann 2014). The taxonomic annotation was performed by uploading the hybridized assembled genomes to the Type Strain Genome Server (TYGS)

available under <https://tygs.dsmz.de> (Meier-Kolthoff and Göker 2019). Subsequently, the average nucleotide identity (ANI) was calculated using the FastANI v1.32 tool (Jain et al. 2018), and digital DNA-DNA hybridization (dDDH) was performed using the web tool Genome-to-Genome Distance Calculator (GGDC) (Meier-Kolthoff et al. 2022). In addition, a genome-based phylogenetic tree was inferred using the autoMLST: Automated Multi-Locus Species Tree pipeline (<https://automlslst.ziemertlab.com/>), where the “denovo mode” was used and the IQ-TREE Ultrafast Bootstrap analysis (1000 replicates) option was selected (Alanjary et al. 2019). The constructed tree was visualized and annotated using the Interactive Tree Of Life (iTOL) available at <https://itol.embl.de/> (Letunic and Bork 2021).

## Identification and analysis of biosynthetic gene clusters (BGCs)

The biosynthetic gene clusters in the nine strains were identified using antiSMASH v6.1.1 (Blin et al., 2021) with the following parameters: `--taxon bacteria --fullhmmer --cmibig --cb-knownclusters --rre --hmm-detection-strictness strict --gene-finding-tool prodigal`. Finally, to summarize the number of BGCs identified in the samples, a matrix plot was created using RAWGraphs (Mauri et al. 2017).

## Biosynthetic gene clusters network of the family *Micrococcaceae*

Genome assemblies belonging to the family *Micrococcaceae* were recovered from the Biosample database of the National Center for Biotechnology Information (NCBI), which includes the environmental packages of soil (50), plant-associated (43), water (11), sediments (7), and wastewater/sludge (4). Subsequently, the BGCs were identified using the same methodology described in this manuscript. To determine biosynthetic biodiversity, sequence similarity networks were generated using the “biosynthetic gene similarity clustering and prospecting engine” BiG-SCAPE v1.1.2 (Navarro-Muñoz et al. 2020) with the PFAM database (v35.0) (Mistry et al. 2021). A distance matrix was created by calculating the distance between every pair of BGCs within the dataset. This distance matrix incorporates three metrics: (i) Jaccard index, representing the percentage of shared domain types; (ii) domain sequence similarity, reflecting the similarity between aligned domain sequences; and (iii) adjacency index, which measures the similarity of domain pair types. In addition, the MiBIG database v3.1 (Minimum Information about a Biosynthetic Gene cluster) (Terlouw et al. 2023) was used. The .gbk files of the BGCs obtained from antiSMASH were used as the inputs. The following settings were used to run the analysis: `--mode auto --mix --no_classify --mibig --cutoffs 0.3 0.4`. Finally,

the networks obtained were annotated and visualized using Cytoscape v3.6.0 (Shannon et al. 2003).

## Results

### Antarctic isolates genome sequencing and taxonomic affiliation

We successfully isolated nine *Actinomycetota* strains from different soil and sediment samples from the Antarctic, following different sample treatments (Table 1). The hybrid assembly of those nine genomes, isolated from samples collected in the Antarctic territory, exhibited completeness exceeding 98% and contamination of less than 1%. The number of contigs varied between 1 and 5 for all assemblies, demonstrating a high quality of contiguity. Specifically, all assemblies had an L50 value of 1 and a notably high N50 value. The strain Sec 6.3 exhibited the lower contiguity. Furthermore, a large number of reads were obtained for each assembly, contributing to the genomic coverage and continuity (Supplementary Table S1). Additionally, all assemblies obtained had a guanine-cytosine content (GC) ranging from

61.5 to 72.9%, with predicted coding sequences numbering between 2348 and 5007, as expected for *Actinomycetota* phylum (Table 2).

The genomics-based tree revealed that the strains belonged to the families *Micrococcaceae* and *Dermatophilaceae* within the phylum *Actinomycetota*, as per the information sourced from the Genome Taxonomy Database (available at <https://gtdb.ecogenomic.org/>) (Figure 1). Specifically, within the family *Micrococcaceae*, strains Sec 5.8 and Sec 7.4 were closely grouped, and Sec 5.7 formed a distinct phylogenetic lineage within the genus *Arthrobacter*. On the other hand, strain Se 16.17 was grouped with *Pae-narthrobacter nicotinovorans* 231Sha2.1M6, while Sec 5.1 formed a distinct phylogenetic lineage within the genus *Pae-narthrobacter*. Strain Sec 6.3 was placed in a separate phylogenetic lineage within the genus *Micrococcus*. Additionally, strain Sec 5.9 was clustered with the species *Janibacter terrae*, and strains Sec 6.4 and Soc 4.6 formed a distinct phylogenetic lineage within the *Dermatophilaceae* family.

Comparing the whole genomic sequences using ANI and dDDH values, we found that almost all the strains exhibited the closest ANI values to species obtained through the NCBI database via autoMLST. However, these values were

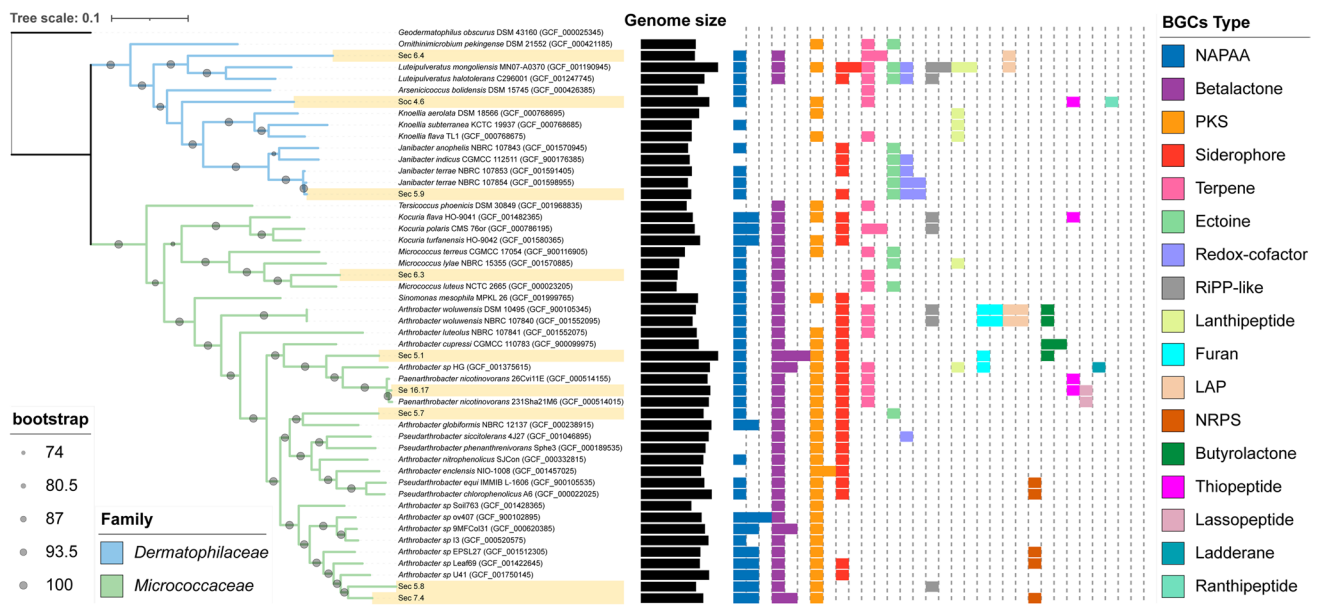
**Table 1** Origin and isolation description of nine bacterial strains obtained from soil and sediment samples from the Antarctic

Isolate strain	Geographical location	Sample source	Sample type	Pretreatment	Isolation medium
Soc 4.6	Antarctica: Deception Island	Deception soil	Soil	No	Oatmeal agar
Sec 5.1	Antarctica: Deception Island	Deception lagoon	Sediment	Phenol	Oatmeal agar
Sec 5.7	Antarctica: Deception Island	Deception lagoon	Sediment	No	Soil agar
Sec 5.8	Antarctica: Deception Island	Deception lagoon	Sediment	No	Soil agar
Sec 5.9	Antarctica: Deception Island	Deception lagoon	Sediment	SDS	Oatmeal agar
Sec 6.3	Antarctica: Doumer Island	Yelcho soil	Soil	Phenol	Soil agar
Sec 6.4	Antarctica: Doumer Island	Yelcho soil	Soil	Chloramine-T	Oatmeal agar
Sec 7.4	Antarctica: Doumer Island	Yelcho channel sediment	Sediment	No	Oatmeal agar
Se 16.17	Antarctica: Rey Jorge (King George) Island, Fildes Peninsula	Lake sediment	Sediment	Chloramine-T	Oatmeal agar

**Table 2** Quality assessments and annotation of hybrid genome assemblies from nine Antarctic isolates

Strain	Genome length	Contigs	N50	L50	GC (%)	Completeness (%)	Contamination (%)	CDS	rRNA	tRNA	tmRNA
Soc 4.6	4,806,652	3	4,746,819	1	71.57	99.46	0	4407	6	52	1
Sec 5.1	5,426,886	1	5,426,886	1	61.49	99.12	0.39	5007	18	56	1
Sec 5.7	4,402,594	2	4,307,194	1	63.92	99.71	0	4215	12	55	1
Sec 5.8	3,870,955	1	3,870,955	1	66.14	99.71	0	3534	15	53	2
Sec 5.9	3,530,731	2	3,492,674	1	71.81	99.82	0.14	3493	6	50	1
Sec 6.3	2,568,479	4	1,294,273	1	72.88	98.21	0.23	2348	6	50	1
Sec 6.4	3,792,347	3	3,718,359	1	65.54	100	0.36	3579	6	49	1
Sec 7.4	4,384,395	5	3,981,032	1	66.81	99.71	0.49	3973	12	60	1
Se 16.17	4,883,589	4	4,495,321	1	62.67	99.71	0.73	4440	18	56	1





**Fig. 1** Multi-locus phylogenetic species tree indicating the distances between isolated strains (highlighted in yellow) and closely related species. The tree was constructed using the web tool autoMLST with the denovo mode and 1000 bootstrap replicates. Circles on branches indicate bootstrap values of > 74%. The Genome Taxonomy Database was used to obtain taxonomic information. *Geodermatophilus*

*obscurus* DSM 43160 (GCF\_000025345) was used as an outgroup. Bar charts in the middle represent the genome size of the strains in the tree. The barplot on the right side shows the number of BGCs of the species in the tree, where each color belongs to a type of BGCs according to the legend

relatively low, except for strain Se 16.17, which shared a high similarity with *P. nicotinovorans* 231Sha2.1M6, showcasing ANI and dDDH values of 98.8% and 86.6%, respectively, and allowing its affiliation as a *P. nicotinovorans* strain. Another exception was strain Sec 5.9, showing 98.8% ANI and 90.1% dDDH, classifying it as the species *Janibacter terrae* (Table 3). Based on our results, seven out of nine strains might be representative of novel species of the genera *Lapillicoccus* sp. (Soc 4.6), *Pseudarthrobacter* sp. (Sec 5.7, Sec 5.8 and Sec 7.4), *Micrococcus* sp. (Sec 6.3), and *Allobranchiibius* sp. (Sec 6.4).

### Analysis and comparison of biosynthetic gene clusters

The nine genomes isolated from the *Micrococcaceae* and *Dermatophilaceae* families collectively possess 49 BGCs. Remarkably, about 84% of these BGCs exhibit a similarity of less than 50% with the closest clusters in the MiBIG database. The most prevalent BGC types within the strains were beta-lactones, type III polyketide synthases (T3PKS), and non-alpha-poly amino acids (NAPAA), with all of them containing NAPAA. Furthermore, BGCs corresponding to

**Table 3** Average nucleotide identity and digital DNA-DNA hybridization values between the isolated strains and their closely related species obtained by the Type Strain Genome Server (TYGS) and Automated Multi-Locus Species Tree (autoMLST)

Strain	Closest genome (TYGS)	ANI (%)	dDDH (%)	Closest genome (autoMLST)	ANI (%)	dDDH (%)
Soc 4.6	<i>Lapillicoccus jejuensis</i> DSM 18607	81.7	21.4	<i>Phycoccus</i> sp. Root563	79.5	20.4
Sec 5.1	<i>Paenarthrobacter nicotinovorans</i> DSM 420	83.3	23.7	<i>Paenarthrobacter nitroguajacolicus</i> HG	85.3	24.3
Sec 5.7	<i>Pseudarthrobacter psychrotolerans</i> YJ56	81.8	23.6	<i>Arthrobacter</i> sp. KBS0703	82.9	23.2
Sec 5.8	<i>Pseudarthrobacter albicanus</i> NJ-Z5	83.2	25.4	<i>Arthrobacter</i> sp. U41	86.7	29.5
Sec 5.9	<i>Janibacter terrae</i> NBRC 107854	98.8	90.1	<i>Janibacter terrae</i> NBRC 107854	98.8	90.1
Sec 6.3	<i>Micrococcus endophyticus</i> BCRC 16908	88.7	33.6	<i>Micrococcus luteus</i> RIT305	88.9	25.5
Sec 6.4	<i>Allobranchiibius huperziae</i> DSM 29531	80.0	20.7	<i>Dermacoccus</i> sp. PE3	74.9	19.6
Sec 7.4	<i>Pseudarthrobacter albicanus</i> NJ-Z5	84.8	27.1	<i>Arthrobacter</i> sp. U41	89.5	35.5
Se 16.17	<i>Paenarthrobacter nitroguajacolicus</i> HG	85.2	27.2	<i>Paenarthrobacter nicotinovorans</i> 231Sha2.1M6	98.8	86.6

siderophores, terpenes, thiopeptide, and non-ribosomal peptide synthetase (NRPS), among others, were identified in smaller quantities (Supplementary Figure S1).

Among the BGCs exhibiting a similarity identity greater than 50% with known clusters in the database, we found siderophores with 75% and 100% similarity to the desferrioxamine E cluster; a type III polyketide synthase with 100% similarity to the alkylresorcinol; a cluster with 75% similarity to the ectoine; and a terpene from strain Sec 6.3, showing 66% similarity to the carotenoid cluster (Supplementary Figure S2). The strains belonging to the *Micrococcaceae* family shared NAPAA, beta-lactone, and T3PKS clusters. Then, we wanted to further investigate whether this feature could be associated with the taxonomic family or could be part of genetic material related to strains inhabiting the extreme environment.

### Biosynthetic gene cluster network analysis in the *Micrococcaceae* family

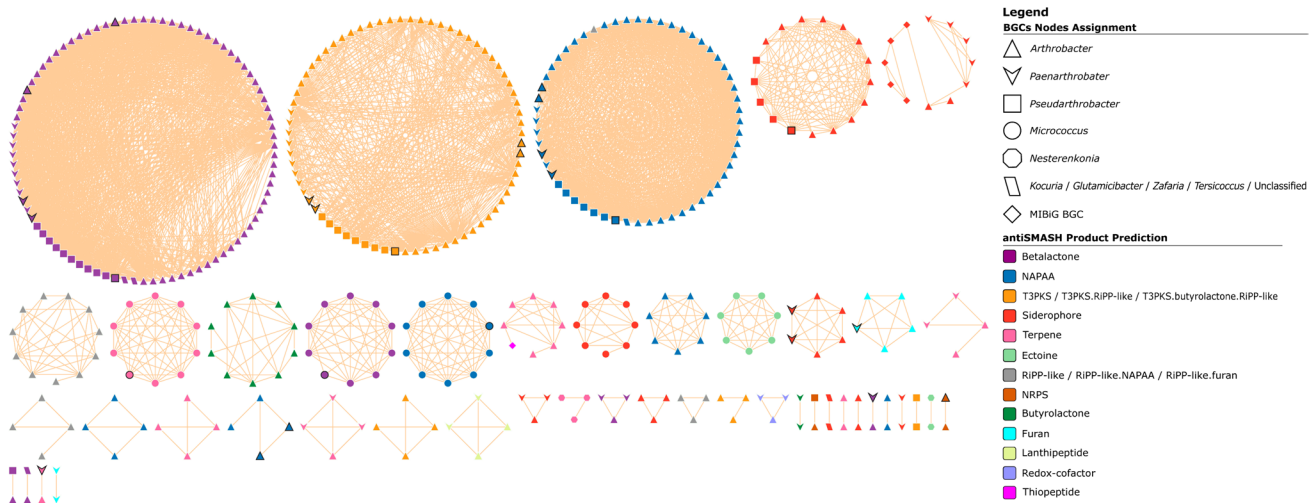
To assess the biosynthetic biodiversity of the *Micrococcaceae* family, we employed the biosynthetic gene similarity clustering and prospecting engine (BiG-SCAPE) to generate sequence similarity networks. The analysis involved 597 BGCs identified from 121 genomic assemblies within the *Micrococcaceae* family (Fig. 2 and Supplementary Table S2). Among these, six assemblies were from strains isolated in Antarctica, while the rest were retrieved from the Biosample database of the NCBI. These genomes were classified under different environmental categories, including soil (50), plant-associated (43), water (11), sediments

(7), and wastewater/sludge (4). The similarity cutoff used for the clustering was set at 0.3.

The sequence similarity network of the BGCs unveiled six predominant types: beta-lactone, NAPAA, T3PKS, siderophores, terpenes, and ectoine. Moreover, the networks containing BGCs from the genera *Micrococcus* and *Nesterenkonia* were distinctly grouped, unlike other genera that shared their BGCs within the same network. Within the identified clusters, two were noteworthy for their annotations from six MiBIG reference BGCs with known functions. Five of these were associated with siderophores, encompassing natural products such as desferrioxamine B (MiBIG BGC0000940 and BGC0000941), desferrioxamine E (MiBIG BGC0001478), and legonoxamine A (MiBIG BGC0002305). Conversely, one was linked to the thiopeptide category but grouped in the same cluster as terpenes, representing a compound called “TP-1161” (MiBIG BGC0000615). This limited representation of networks containing previously characterized BGCs underscores the extensive BGC diversity within the *Micrococcaceae* family. We attempted to generate networks based on the source of isolation and continent, but no discernible relationships were identified related to the origin of the isolated (Supplementary Figure S3 and S4).

## Discussion

The Antarctic stands as the coldest, driest, windiest, and most challenging-to-access continent on Earth. Due to its extreme conditions, it has been identified as a potential



**Fig. 2** Sequence similarity network of 597 BGCs with a cutoff of 0.3 belonging to the *Micrococcaceae* family generated by BiG-SCAPE and visualized with Cytoscape v3.9.1. Each node represents an individual BGC, colored according to the antiSMASH product prediction. The nodes with black borders represent the BGCs identified in

our Antarctic strains. Singletons, which are unique BGCs with no connections, are not displayed. NAPAA, non-alpha-polyamino acids; T3PKS, type iii polyketide synthases; RIPP, ribosomally synthesized and post-translationally modified peptides; NRPS, non-ribosomal peptide synthetase

source for the discovery of novel natural products from bacteria (Núñez-Montero and Barrientos 2018). We isolated and analyzed nine strains from soil and sediment samples collected across diverse locations within the Antarctic territory. These strains were obtained through selective isolation methods to increase the probability of obtaining rare strains and to facilitate the discovery of natural products. Heating in the pretreatment of the samples, the use of chemicals such as phenol and sodium dodecylsulfate and antimicrobial agents, such as nalidixic acid, was used to reduce the growth of undesirable bacteria and increase the selectivity of rare *Actinomycetota* strains. This is possible because there are several strains of *Actinomycetota* resistant to a broad spectrum of antibiotics, together with the addition of chemical compounds toxic to bacteria, which allows for a decrease in the number of bacteria sensitive to these agents, thus enabling the isolation of new strains for the search of new natural products (Subramani and Aalbersberg 2013).

Genomic assembly of the isolated strains was facilitated using a hybrid approach, integrating short reads through Illumina sequencing and long reads using Oxford Nanopore Technology (ONT). Combining these sequencing technologies enhances genome assembly, addressing the inherent limitations of Illumina, which struggles with errors in sequencing repetitive structures and relies on short read sizes, typically less than 500 base pairs. These challenges are mitigated with ONT, generating large reads, often exceeding 10,000 bp. This allows for the assembly of complex genomes, resulting in fewer contigs and enhanced precision and contiguity (De Maio et al. 2019). This is evidenced by genomic annotation, wherein all assemblies display a high N50 value and a low L50 value, except for strain Sec 6.3, which demonstrates a slightly lower N50 value. The L50 and N50 values represent the number of contigs and contig lengths, respectively, crossing the 50% mark of the assembly. Hence, higher N50 values and lower L50 values signify a more contiguous assembly (Jayakumar and Sakakibara 2019). Additionally, we achieved a substantial depth of coverage in the genomic assemblies utilizing both technologies, as outlined in the supplementary tables. Consequently, this precise analysis allows for a thorough examination of the biosynthetic gene clusters and various genomic characteristics of the obtained strains.

Taxonomic annotation revealed that strains Sec 5.9 and Se 16.17 exhibited high dDDH and ANI values, with their closest genomes being *Janibacter terrae* NBRC 107854 and *Paenarthrobacter nicotinovorans* 231Sha2.1M6, respectively. The generally accepted cutoff values for identifying new species within a genus are 95–96% for ANI and AAI and 70% for dDDH (Richter and Rosselló-Móra 2009; Chun et al. 2018). Consequently, these strains are identified as belonging to the mentioned species. Strains affiliated with the *J. terrae* species were previously

isolated from environmental water and soil samples, showcasing the capability to degrade environmental pollutants (Lang et al. 2003). On the other hand, *P. nicotinovorans* 231Sha2.1M6 was isolated from *Arabidopsis thaliana* soil samples. As for the remaining seven strains, both their dDDH and ANI values fell below the defined cutoff values, designating them as new species within the *Micrococcaceae* and *Dermatophilaceae* families.

In the analysis of the identified BGCs within the isolated strains, a variable number, ranging from three to eight BGCs, was observed. Among them, three main types stood out: NAPAA, beta-lactone, and T3PKS. Beta-lactone-associated BGCs have demonstrated applications as anticancer agents, such as Marizomib (salinosporamide A) and antimicrobial and antiobesity agents, like Xenical (Robinson et al. 2019; Wang et al. 2021a, b). T3PKS, responsible for biosynthesizing a plethora of natural products, encompass antibiotics, immunosuppressants, and cancer chemotherapy (Nivina et al. 2019). NAPAA, such as  $\epsilon$ -poly-L-lysine, exhibit bacteriostatic and biodegradable properties, leading to their use as food preservatives and in the pharmaceutical industry (Wang et al. 2021a, b). Another noteworthy category of BGCs with potential public health applications includes RiPPs, which feature a domain known as the RiPP recognition element (RRE-containing). This domain is prevalent in various RiPP clusters (Kloosterman et al. 2020) and has been instrumental in addressing therapeutic challenges, such as combating cystic fibrosis and acting as antimicrobials (Hetrick and Donk 2017).

Furthermore, concerning BGCs with a similarity greater than 50% to known BGCs in the MiBIG database, we identified alkylresorcinol in the Soc 4.6 strain. Alkylresorcinol is a lipophilic molecule with a polyphenol structure known for its antimicrobial, anticancer, antilipidemic, antioxidant, and other properties (Zabolotneva et al. 2022). Additionally, we found BGCs linked to siderophores, particularly desferrioxamine E in the Sec 5.1, 5.7, 5.9, and Se 16.17 strains, which play a crucial role in microbial growth (Yamanaka et al. 2005). On a different note, we also identified BGCs associated with more physiological functions, such as terpenes and ectoine. Terpenes are linked to electron transport, light uptake, photoprotection, and signaling (Caulier et al. 2019), while ectoine is known for its protective properties on enzymes, DNA, cell membranes, and cells against various types of stress, including osmotic, cold, and heat stress (Zhang et al. 2009). Hence, our results showed that the possibly known molecules are related to primary microbial metabolism and or common functions across. On the other hand, our strains showed a large number of BGCs with low or non-similarity to other clusters, and belonging to types of specialized metabolites previously related to biological activities. Hence, our data highlights the metabolic potential

for the production of novel natural products from novel strains isolated from the Antarctic continent.

The network analysis of BGCs allows us to visualize a phylogenetic correlation between certain genera of the family *Micrococcaceae*, especially observed in genera such as *Micrococcus* and *Nesterenkonia*. This observation has been observed in studies of other species belonging to the phylum *Actinomycetota*, such as the genera *Amycolatopsis*, *Salinispora*, and *Streptomyces*, among others, where these correlations were detected using different bioinformatic approaches, such as the identification of gene cluster families, operational biosynthetic units, and BGCs (Doroghazi and Metcalf 2013; Ziemert et al. 2014; Adamek et al. 2018). In addition, it has been detected in the phylum Firmicutes, with the genus *Bacillus*, where clade- and species-specific BGCs have been found (Steinke et al. 2021). These BGCs-related families might be associated with core metabolism of the genus. Furthermore, among the three most abundant groups of BGCs within this family, T3PKS have been previously reported for the genus *Arthrobacter* (Doroghazi and Metcalf 2013), but not the beta-lactone BGCs, which are grouped in one network and currently do not have a close reference and could be part of the core genome of these strains because of their high level of conservation.

In addition, no relationship was found between the source of isolation and the BGCs. This might be because there are many types of environments and different conditions of temperature, humidity, and pressure, among others, which generates a very diverse grouping in the network. On the other hand, there are many BGCs to be studied within this family, since only two groups are similar to any known natural product, being siderophores and thiopeptides. The latter obtained from an *Actinomycetota*, *Nocardiopsis* sp., has potent antibiotic properties (Engelhardt et al. 2010). Therefore, it is important to further investigate BGCs belonging to this family as several beneficial properties have been demonstrated, such as the ability to act as plant growth promoters in various plant species, as seen in members belonging to the genera *Arthrobacter* and *Pseudarthrobacter* (Chhetri et al. 2022; Jiang et al. 2022; Ham et al. 2022; Platamone et al. 2023). In addition, species of the genera *Arthrobacter* and *Micrococcus* have been used for the bioremediation of organic compounds and heavy metals (Behera and Das 2023). For the species belonging to the genus *Micrococcus*, extracts with pharmaceutical properties, such as antimicrobial, antifungal, and antioxidant properties, have been obtained, which are described in more detail in the review by Tizabi and Hill (2023).

In summary, the species isolated and sequenced in this study exhibit BGCs with significant potential across various domains including clinical applications,

pharmaceuticals, the food industry, and agriculture. Moreover, we observed a substantial gap in our understanding of the functionality of different BGC groups within the *Micrococcaceae* family, which holds significant implications for the future. Our research also provides valuable genomics data from untapped microorganisms from rare *Actinomycetota* strains from extreme environments.

Given that this analysis was conducted *in silico*, we were unable to confirm whether these BGCs are actively expressed or silenced. Future investigations integrating transcriptomics and metabolomics are essential to determine their expression and functionality. This comprehensive approach will pave the way for these BGCs to emerge as promising candidates for future applications and biotechnological advancements.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00253-024-13154-x>.

**Acknowledgements** The authors wish to express their deep gratitude to Agencia Nacional de Investigación y Desarrollo (ANID) Chile by FONDECYT Regular project 1210563.

**Author contribution** PB, MG, KL, and MJC contributed to the laboratory work and analysis of the data and drafted the paper. LB was awarded with funding from Agencia Nacional de Investigación y Desarrollo (ANID) Chile. KN and LB offered critical suggestions for designing experimental assays and data analysis. KN carried out the isolation of the strains as part of her Ph.D. research under the supervision of LB. LB and MA designed the work, supervised the study, and critically revised the manuscript. All authors read and approved the manuscript.

**Funding** This work was supported by the Agencia Nacional de Investigación y Desarrollo (ANID) Chile, by FONDECYT Regular project 1210563; Grant Doctorado Nacional 2023-21230832 for PB and 2023-21231681 for MG.

**Data availability** The microbial reads and genome assemblies generated in this study are available in the National Center for Biotechnology Information (NCBI) under the BioProject ID: PRJNA1063326.

## Declarations

**Ethics approval** No human participants or animals were used in this study.

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



## References

- Adamek M, Alanjary M, Sales-Ortells H, Goodfellow M, Bull AT, Winkler A, Wibberg D, Kalinowski J, Ziemert N (2018) Comparative genomics reveals phylogenetic distribution patterns of secondary metabolites in *Amycolatopsis* species. *BMC genomics* 19(1):426. <https://doi.org/10.1186/s12864-018-4809-4>
- Alanjary M, Steinke K, Ziemert N (2019) AutoMLST: an automated web server for generating multi-locus species trees highlighting natural product potential. *Nucleic Acids Res* 47(W1):W276–W282. <https://doi.org/10.1093/nar/gkz282>
- Albarano L, Esposito R, Ruocco N, Costantini M (2020) Genome mining as new challenge in natural products discovery. *Mar Drugs* 18(4):199. <https://doi.org/10.3390/MD18040199>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed 13 Apr 2023
- Behera S, Das S (2023) Potential and prospects of *Actinobacteria* in the bioremediation of environmental pollutants: cellular mechanisms and genetic regulations. *Microbiol Res* 273:127399. <https://doi.org/10.1016/j.micres.2023.127399>
- Benaud N, Edwards RJ, Amos TG, D'Agostino PM, Gutiérrez-Chávez C, Montgomery K, Nicetic I, Ferrari BC (2021) Antarctic desert soil bacteria exhibit high novel natural product potential, evaluated through long-read genome sequencing and comparative genomics. *Environ Microbiol* 23(7):3646–3664. <https://doi.org/10.1111/1462-2920.15300>
- Benaud N, Chelliah DS, Wong SY, Ferrari BC (2022) Soil substrate culturing approaches recover diverse members of *Actinomyces* from desert soils of Herring Island, East Antarctica. *Extremophiles* 26(2). <https://doi.org/10.1007/S00792-022-01271-2>
- Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T (2021) antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res* 49(W1):W29–W35. <https://doi.org/10.1093/nar/gkab335>
- Borker SS, Thakur A, Kumar S, Kumari S, Kumar R, Kumar S (2021) Comparative genomics and physiological investigation supported safety, cold adaptation, efficient hydrolytic and plant growth-promoting potential of psychrotrophic *Glutamicibacter arilaitensis* LJH19, isolated from night-soil compost. *BMC Genomics* 22(1):307. <https://doi.org/10.1186/S12864-021-07632-Z>
- Caulier S, Nannan C, Gillis A, Licciardi F, Bragard C, Mahillon J (2019) Overview of the antimicrobial compounds produced by members of the *Bacillus subtilis* group. *Front. Microbiol* 10:302. <https://doi.org/10.3389/fmicb.2019.00302>
- Chen S, Zhou Y, Chen Y, Gu J (2018) fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* (Oxford, England) 34(17):i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chhetri G, Kim I, Kang M, So Y, Kim J, Seo T (2022) An isolated *Arthrobacter* sp. enhances rice (*Oryza sativa* L.) plant growth. *Microorganisms* 10(6):1187. <https://doi.org/10.3390/microorganisms10061187>
- Chun J, Oren A, Ventosa A, Christensen H, Arahall DR, da Costa MS, Rooney AP, Yi H, Xu XW, De Meyer S, Trujillo ME (2018) Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68(1):461–466. <https://doi.org/10.1099/ijssem.0.002516>
- De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C (2018) NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* (Oxford, England) 34(15):2666–2669. <https://doi.org/10.1093/bioinformatics/bty149>
- De Maio N, Shaw LP, Hubbard A, George S, Sanderson ND, Swann J, Wick R, AbuOun M, Stubberfield E, Hoosdally SJ, Crook DW, Peto T, Sheppard AE, Bailey MJ, Read DS, Anjum MF, Langmead AS, Stoesser N, On Behalf Of The Rehab Consortium (2019) Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. *Microb Genomics* 5(9):e000294. <https://doi.org/10.1099/mgen.0.000294>
- Doroghazi JR, Metcalf WW (2013) Comparative genomics of *actinomyces* with a focus on natural product biosynthetic genes. *BMC genomics* 14:611. <https://doi.org/10.1186/1471-2164-14-611>
- Engelhardt K, Degnes KF, Kemmler M, Bredholt H, Fjaervik E, Klinkenberg G, Sletta H, Ellingsen TE, Zotchev SB (2010) Production of a new thiopeptide antibiotic, TP-1161, by a marine *Nocardiopepsis* species. *Appl Environ Microbiol* 76(15):4969–4976. <https://doi.org/10.1128/AEM.00741-10>
- Gummerlich N, Rebets Y, Paulus C, Zapp J, Luzhetskyy A (2020) Targeted genome mining—from compound discovery to biosynthetic pathway elucidation. *Microorganisms* 8(12): 2034 <https://doi.org/10.3390/MICROORGANISMS8122034>
- Gurevich A, Saveliev V, Vyahhi N, Tesler G (2013) QUASt: quality assessment tool for genome assemblies. *Bioinformatics* (Oxford, England) 29(8):1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
- Ham SH, Yoon AR, Oh HE, Park YG (2022) Plant growth-promoting microorganism *Pseudarthrobacter* sp NIBRBAC000502770 enhances the growth and flavonoid content of *Geum aleppicum*. *Microorganisms* 10(6):1241 <https://doi.org/10.3390/microorganisms10061241>
- Hetrick KJ, van der Donk WA (2017) Ribosomally synthesized and post-translationally modified peptide natural product discovery in the genomic era. *Curr Opin Chem Biol* 38:36–44. <https://doi.org/10.1016/j.cbpa.2017.02.005>
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S (2018) High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9(1):5114. <https://doi.org/10.1038/s41467-018-07641-9>
- Javed Z, Tripathi GD, Mishra M, Dashora K (2021) *Actinomyces* – the microbial machinery for the organic-cycling, plant growth, and sustainable soil health. *Biocatal Agric Biotechnol* 31:101893. <https://doi.org/10.1016/J.BCAB.2020.101893>
- Jayakumar V, Sakakibara Y (2019) Comprehensive evaluation of non-hybrid genome assembly tools for third-generation PacBio long-read sequence data. *Brief Bioinform* 20(3):866–876. <https://doi.org/10.1093/bib/bbx147>
- Jiang Y, Song Y, Jiang C, Li X, Liu T, Wang J, Chen C, Gao J (2022) Identification and characterization of *Arthrobacter nicotinovorans* JI39, a novel plant growth-promoting rhizobacteria strain from panax ginseng. *Front. Plant Sci.* 13:873621. <https://doi.org/10.3389/fpls.2022.873621>
- Kloosterman AM, Shelton KE, van Wezel GP, Medema MH, Mitchell DA (2020) RRE-Finder: a genome-mining tool for class-independent RiPP discovery. *mSystems* 5(5): e00267-20 <https://doi.org/10.1128/mSystems.00267-20>
- Kumar PS, Ling CY, Zhou ZB, Dong YL, Sun CL, Song YX, Wong NK, Ju JH (2020) Chemical diversity of metabolites and antibacterial potential of *Actinomyces* associated with marine invertebrates from intertidal regions of Daya Bay and Nansha Islands. *Microbiology* (Russian Federation) 89(4):483–492. <https://doi.org/10.1134/S0026261720040062>
- Lang E, Kroppenstedt RM, Swiderski J, Schumann P, Ludwig W, Schmid A, Weiss N (2003) Emended description of *Janibacter terrae*, including ten dibenzofuran-degrading strains and *Janibacter brevis* as its later heterotypic synonym. *Int J Syst Evol Microbiol* 53(Pt 6):1999–2005. <https://doi.org/10.1099/ijs.0.02602-0>
- Letunic I, Bork P (2021) Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49(W1):W293–W296. <https://doi.org/10.1093/nar/gkab301>

- Matsumoto A, Takahashi Y (2017) Endophytic *actinomycetes*: promising source of novel bioactive compounds. *J Antibiot* 70(5):514–519. <https://doi.org/10.1038/ja.2017.20>
- Mauri M, Elli T, Caviglia G, Uboldi G, Azzi M (2017) RAWGraphs: a visualisation platform to create open outputs. In Proc. of the 12th Biannual Conf. on Italian SIGCHI Chapter (pp. 28:1–28:5). New York, NY, USA: ACM. <https://doi.org/10.1145/3125571.3125585>
- Meier-Kolthoff JP, Göker M (2019) TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10(1):2182. <https://doi.org/10.1038/s41467-019-10210-3>
- Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M (2022) TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50(D1):D801–D807. <https://doi.org/10.1093/nar/gkab902>
- Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonhammer ELL, Tosatto SCE, Paladin L, Raj S, Richardson LJ, Finn RD, Bateman A (2021) Pfam: the protein families database in 2021. *Nucleic Acids Res* 49(D1):D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Navarro-Muñoz JC, Selem-Mojica N, Mullowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos E, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappelini L, Goering AW, Thomson RJ, Metcalf WW, Kelleher NL, Barona-Gomez F, Medema MH (2020) A computational framework to explore large-scale biosynthetic diversity. *Nat Chem Biol* 16(1):60–68. <https://doi.org/10.1038/s41589-019-0400-9>
- Nivina A, Yuet KP, Hsu J, Khosla C (2019) Evolution and diversity of assembly-line polyketide synthases. *Chem Rev* 119(24):12524–12547. <https://doi.org/10.1021/acs.chemrev.9b00525>
- Núñez-Montero K, Barrientos L (2018) Advances in Antarctic research for antimicrobial discovery: a comprehensive narrative review of bacteria from antarctic environments as potential sources of novel antibiotic compounds against human pathogens and microorganisms of industrial importance. *Antibiotics* (Basel, Switzerland) 7(4):90. <https://doi.org/10.3390/antibiotics7040090>
- Núñez-Montero K, Lamilla C, Abanto M, Maruyama F, Jorquera MA, Santos A, Martínez-Urtaza J, Barrientos L (2019) Antarctic *Streptomyces fildesensis* So13.3 strain as a promising source for antimicrobials discovery. *Sci Rep* 9(1). <https://doi.org/10.1038/S41598-019-43960-7>
- Oyedoh OP, Yang W, Dhanasekaran D, Santoyo G, Glick BR, Babalola OO (2023) Rare rhizo-*Actinomycetes*: a new source of agroactive metabolites. *Biotechnol Adv* 67:108205. <https://doi.org/10.1016/J.BIOTECHADV.2023.108205>
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25(7):1043–1055. <https://doi.org/10.1101/gr.186072.114>
- Platamone G, Procacci S, Maccioni O, Borromeo I, Rossi M, Bacchetta L, Forni C (2023) *Arthrobacter* sp. inoculation improves cactus pear growth, quality of fruits, and nutraceutical properties of cladodes. *Curr Microbiol* 80(8):266. <https://doi.org/10.1007/s00284-023-03368-z>
- Ramasamy KP, Mahawar L, Rajasabapathy R, Rajeshwari K, Miceli C, Pucciarelli S (2023) Comprehensive insights on environmental adaptation strategies in Antarctic bacteria and biotechnological applications of cold adapted molecules. *Front Microbiol* 14:1197797. <https://doi.org/10.3389/fmicb.2023.1197797>
- Reis-Mansuro MCPP, Cardoso-Rurr JS, Silva JVMA, de Souza GR, Mansoldo FRP, Pinheiro Y, Schultz J, Lopez Balottin LB, da Silva AJR, Lage C, dos Santos EP, Rosado AS, Vermelho AB (2019) Carotenoids from UV-resistant Antarctic Microbacterium sp. LEMMJ01. *Sci Rep* 9(1):1–14. <https://doi.org/10.1038/s41598-019-45840-6>
- Richter M, Rosselló-Móra R (2009) Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 106(45):19126–19131. <https://doi.org/10.1073/pnas.0906412106>
- Robinson SL, Christenson JK, Wackett LP (2019) Biosynthesis and chemical diversity of  $\beta$ -lactone natural products. *Nat Prod Rep* 36(3):458–475. <https://doi.org/10.1039/c8np00052b>
- Scherlach K, Hertweck C (2021) Mining and unearthing hidden biosynthetic potential. *Nat Commun* 12(1):1–12. <https://doi.org/10.1038/s41467-021-24133-5>
- Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* (Oxford, England) 30(14):2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13(11):2498–2504. <https://doi.org/10.1101/gr.1239303>
- Silva LJ, Crevelin EJ, Souza DT, Lacerda-Júnior GV, de Oliveira VM, Ruiz ALTG, Rosa LH, Moraes LAB, Melo IS (2020) *Actinobacteria* from Antarctica as a source for anticancer discovery. *Sci Rep* 10(1):13870. <https://doi.org/10.1038/S41598-020-69786-2>
- Singh TA, Passari AK, Jajoo A, Bhasin S, Gupta VK, Hashem A, Alqarawi AA, Abd\_Allah EF (2021) Tapping into actinobacterial genomes for natural product discovery. *Front Microbiol* 12:655620. <https://doi.org/10.3389/FMICB.2021.655620>
- Steinke K, Mohite OS, Weber T, Kovács ÁT (2021) Phylogenetic distribution of secondary metabolites in the *Bacillus subtilis* species complex. *mSystems* 6(2):e00057-21. <https://doi.org/10.1128/mSystems.00057-21>
- Styczynski M, Biegniewski G, Decewicz P, Rewerski B, Debiec-Andrzejewska K, Dziewit L (2022) Application of psychrotolerant antarctic bacteria and their metabolites as efficient plant growth promoting agents. *Front Bioeng Biotechnol* 10:772891. <https://doi.org/10.3389/fbioe.2022.772891>
- Subramani R, Aalbersberg W (2013) Culturable rare *Actinomycetes*: diversity, isolation and marine natural product discovery. *Appl Microbiol Biotechnol* 97(21):9291–9321. <https://doi.org/10.1007/s00253-013-5229-7>
- Terlouw BR, Blin K, Navarro-Muñoz JC, Avalon NE, Chevrette MG, Egbert S, Lee S, Meijer D, Recchia MJJ, Reitz ZL, van Santen JA, Selem-Mojica N, Tørring T, Zaroubi L, Alanjary M, Aleti G, Aguilar C, Al-Salihi SAA, Augustijn HE, Avelar-Rivas JA, Avitia-Domínguez LA, Barona-Gómez F, Bernaldo-Agüero J, Bielinski VA, Biermann F, Booth TJ, Carrion Bravo VJ, Castelo-Branco R, Chagas FO, Cruz-Morales P, Du C, Duncan KR, Gavriilidou A, Gayraud D, Gutiérrez-García K, Haslinger K, Helfrich EJJ, van der Hooft JJJ, Jati AP, Kalkreuter E, Kalyvas N, Kang KB, Kautsar S, Kim W, Kunjapur AM, Li Y-X, Lin G-M, Loureiro C, Louwen JJR, Louwen NLL, Lund G, Parra J, Philmus B, Pourmohsenin B, Pronk LJ, Rego A, Rex DAB, Robinson S, Rosas-Becerra LR, Roxborough ET, Schorn MA, Scobie DJ, Singh KS, Sokolova N, Tang X, Udway D, Vigneshwari A, Vind K, Vromans SPJM, Waschulin V, Williams SE, Winter JM, Witte TE, Xie H, Yang D, Yu J, Zdouc M, Zhong Z, Collemare J, Linington RG, Weber T, Medema MH (2023) MIBiG 30 a community-driven effort to annotate experimentally validated biosynthetic gene clusters. *Nucleic Acids Res* 51(D1):D603–D610. <https://doi.org/10.1093/nar/gkac1049>
- Tizabi D, Bachvaroff T, Hill RT (2022) Comparative analysis of assembly algorithms to optimize biosynthetic gene cluster identification in novel marine *actinomycete* genomes. *Front Mar Sci* 9:914197. <https://doi.org/10.3389/FMARS.2022.914197/BIBTEX>

- Tizabi D, Hill RT (2023) *Micrococcus spp* as a promising source for drug discovery: a review. *J Ind Microbiol Biotechnol* 50(1):kuad017 <https://doi.org/10.1093/jimb/kuad017>
- Wang J, Shi Y, Jiang D (2021a)  $\beta$ -Lactone derivatives and their anticancer activities: a short review. *Curr Top Med Chem* 21(18):1645–1656. <https://doi.org/10.2174/1568026621666210402142150>
- Wang L, Zhang C, Zhang J, Rao Z, Xu X, Mao Z, Chen X (2021b) Epsilon-poly-L-lysine: recent advances in biomanufacturing and applications. *Front Bioeng Biotechnol* 9:748976. <https://doi.org/10.3389/fbioe.2021.748976>
- Wick RR, Holt KE (2022) Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18(1):e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>
- Wick RR, Judd LM, Gorrie CL, Holt KE (2017) Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 13(6):e1005595. <https://doi.org/10.1371/journal.pcbi.1005595>
- Yamanaka K, Oikawa H, Ogawa HO, Hosono K, Shinmachi F, Takano H, Sakuda S, Beppu T, Ueda K (2005) Desferrioxamine E produced by *Streptomyces griseus* stimulates growth and development of *Streptomyces tanashiensis*. *Microbiology (Reading, England)* 151(Pt 9):2899–2905. <https://doi.org/10.1099/mic.0.28139-0>
- Zabolotneva AA, Shatova OP, Sadova AA, Shestopalov AV, Roumiantsev SA (2022) An overview of alkylresorcinols biological properties and effects. *J Nutr Metab* 4667607 <https://doi.org/10.1155/2022/4667607>
- Zhang LH, Lang YJ, Nagata S (2009) Efficient production of ectoine using ectoine-excreting strain. *Extremophiles* 13(4):717–724. <https://doi.org/10.1007/s00792-009-0262-2>
- Ziemert N, Lechner A, Wietz M, Millán-Aguíñaga N, Chavarria KL, Jensen PR (2014) Diversity and evolution of secondary metabolism in the marine *actinomycete* genus *Salinispora*. *Proc Natl Acad Sci USA* 111(12):E1130–E1139. <https://doi.org/10.1073/pnas.1324161111>
- Zimin AV, Salzberg SL (2020) The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. *PLoS Comput Biol* 16(6):e1007981. <https://doi.org/10.1371/journal.pcbi.1007981>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.