CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	Discovery and synthesis of leaderless bacteriocins from the Actinomycetota
Authors	Hourigan, David;de Farias, Felipe Miceli;O'Connor, Paula M.;Hill, Colin;Ross, R. Paul
Publication date	2024-11-24
Original Citation	Hourigan, D., Miceli De Farias, F., O'Connor, P.M., Hill, C. and Ross, R.P. (2024) 'Discovery and synthesis of leaderless bacteriocins from the Actinomycetota', Journal of Bacteriology, 206(11), e00298-24 (15pp). https://doi.org/10.1128/jb.00298-24
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1128/jb.00298-24
Rights	© 2024, Hourigan et al. This is an openaccess article distributed under the terms of the Creative Commons Attribution 4.0 International license. - https://creativecommons.org/licenses/by/4.0/
Download date	2026-10-30 00:23:27
Item downloaded from	https://hdl.handle.net/10468/17821

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Journal of Bacteriology

| Editor's Pick | Antimicrobial Chemotherapy | Full-Length Text

Discovery and synthesis of leaderless bacteriocins from the Actinomycetota

David Hourigan,[1,2] Felipe Miceli de Farias,[1] Paula M. O'Connor,[1,3] Colin Hill,[1,2] R. Paul Ross[1,2,3]

AUTHOR AFFILIATIONS See affiliation list on p. 12.

ABSTRACT Leaderless bacteriocins are a unique class of bacteriocins that possess antimicrobial activity after translation and have few cases of documented resistance. Aureocin A53 and lacticin Q are considered two of the most well-studied leaderless bacteriocins. Here, we used *in silico* genome mining to search for novel aureocin A53-like leaderless bacteriocins in GenBank and MGnify. We identified 757 core peptides across 430 genomes with 75 species found currently without characterized leaderless bacteriocin production. These include putative novel species containing bacteriocin gene clusters (BGCs) from the genera *Streptomyces* (sp. NBC_00237) and *Agrococcus* (sp. SL85). To date, all characterized leaderless bacteriocins have been found within the phylum Bacillota, but this study identified 97 core peptides within the phylum Actinomycetota. Members of this phylum are traditionally associated with the production of antibiotics, such is the case with the genus *Streptomyces*. Actinomycetota is an underexplored phylum in terms of bacteriocin production with no characterized leaderless bacteriocin production to date. The two novel leaderless bacteriocins arcanocin and arachnicin from Actinomycetota members *Arcanobacterium* sp. and *Arachnia* sp., respectively, were chemically synthesized and antimicrobial activity was verified. These peptides were encoded in human gut (PRJNA485056) and oral (PRJEB43277) microbiomes, respectively. This research highlights the biosynthetic potential of Actinomycetota in terms of leaderless bacteriocin production and describes the first antimicrobial peptides encoded in the genera *Arcanobacterium* and *Arachnia*.

IMPORTANCE Bacteriocins are gathering attention as alternatives to current antibiotics given the increasing incidence of antimicrobial resistance. Leaderless bacteriocins are considered a commercially attractive subclass of bacteriocins due to the ability to synthesize active peptide and low levels of documented resistance. Therefore, in this work, we mined publicly available data to determine how widespread and diverse leaderless bacteriocins are within the domain of bacteria. Actinomycetota, known for its antibiotic producers but lacking described and characterized bacteriocins, proved to be a rich source of leaderless bacteriocins—97 in total. Two such peptides, arcanocin and arachnicin, were chemically synthesized and have antimicrobial activity. These bacteriocins may provide a novel source of novel antimicrobials that could aid in the development of future alternative antimicrobials and highlight that the Actinomycetota are an underexplored resource of bacteriocin peptides.

KEYWORDS bacteriocins, Actinomycetota, leaderless bacteriocin, aureocin A53, antimicrobial-peptide

Bacteriocins are ribosomally synthesized antimicrobial peptides/proteins that often have antimicrobial activity against strains similar to the producing organism (1). It is estimated that approximately 50% of all bacteria have the genetic capacity to produce at least one bacteriocin (1). They can be broadly categorized into two main classes: class 1

Editor Michael Y. Galperin, NCBI, NLM, National Institutes of Health, Bethesda, Maryland, USA

Address correspondence to R. Paul Ross, p.ross@ucc.ie.

The authors declare no conflict of interest.

See the funding table on p. 12.

Received 30 July 2024
Accepted 24 September 2024
Published 15 October 2024

Copyright © 2024 Hourigan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

is ribosomally produced and post-translationally modified peptides (RiPPs), and class 2 is unmodified peptides (2). RiPPs undergo a series of modifications and require extracellular transport and regulation of production, which usually results in gene clusters above 10 kb in size. Class 2 has less complex biosynthetic gene clusters (BGCs) and do not undergo post-translational modifications. Recent advancements with *in silico* detection methods of bacteriocin BGCs across expanding databases have led to the detection of thousands of novel potential bacteriocins that could have antimicrobial activity against etiological causes for concern, such as the case with lanthipeptides, thiopeptides, and circular bacteriocins (3–5). Deep sequencing and assembly of metagenome-assembled genomes (MAGs) can uncover the genomic content of thousands of unculturable bacteria and have facilitated the in-depth study of microbiomes and their functional potential (6). This abundance of microbial genomes is frequently subjected to bioinformatic mining to decipher the biosynthetic potential of microbiomes and uncover mechanisms of competition and communication between bacteria (7, 8). However, detecting these potentially bioactive peptides has not coincided with advancements in production or the ability to culture obscure bacteria encoding novel gene clusters.

Leaderless bacteriocins represent a medically and commercially attractive subclass of these peptides due to the ability to genetically engineer, synthesize, and hence mass produce them. Indeed, the lack of a leader sequence to direct peptide export renders these peptides unusual; thus, they are active immediately after translation, meaning the producing organism is susceptible. This can be seen with lacticin Q, which can be intracellularly toxic (9). This peculiar lack of identifiable mechanism of leading the peptide for extracellular transport and known ligands means that leaderless bacteriocins remain somewhat of an enigma. Canonical aureocin A53-like leaderless bacteriocins have a characteristic structure similar to circular bacteriocins in that they possess four to five helices that conformationally form a small globular structure. However, a key differentiating factor is that circular bacteriocins are N–C terminal covalently linked, whereas leaderless bacteriocins are linear (10, 11). Leaderless bacteriocins have a formylated methionine at the N-terminal, which is not crucial for antimicrobial activity (10). To date, over 21 leaderless bacteriocins have been described in the literature, including aureocin A53 which is considered to be the prototypical leaderless bacteriocin that defines the class (12). Non-aureocin-like leaderless bacteriocins have been found that are shorter in length and have fewer tryptophan residues in their core peptides (13, 14). Aureocin A53 is a 51 residue, 6 kDa, highly cationic peptide with five tryptophan residues (12). These tryptophan residues contribute to the stability of the peptide and its mechanism of action and play a role in both oligomerization and protease resistance (15). The BGC in *Staphylococcus aureus* A53 is on the 10.4-kb plasmid pRJ9 (12). Interestingly, the operon produces the peptide without obvious regulatory elements or a protease. However, ABC transporter-like machinery has been described (16). Aureocin A53 has activity against *Listeria monocytogenes*, *S. aureus*, and *Streptococcus agalactiae* and is gamma-hemolytic against sheep blood (17–19). It has a broad spectrum of activity against gram-positive bacteria due to a mechanism of action that involves destabilizing the bacterial membrane. It has potent activity against clinically-relevant pathogens such as vancomycin-resistant enterococci (VRE) [minimum inhibitory concentration (MIC) 0.29 µg/mL] (17). Recently discovered leaderless bacteriocins include bawcin, a 48 amino acid-long peptide produced by *Bacillus wiedmannii* and toyoncin, a 70 amino acid-long peptide produced by *Bacillus toyonensis* XIN-YC13 (20, 21). Leaderless bacteriocins can be broadly categorized based on the number of core peptides they possess in their BGCs, that is, single peptide bacteriocins, two peptide bacteriocins, and multi-peptide bacteriocins (13, 14). However, leaderless bacteriocins can also be further categorized based on amino acid sequence into aureocin A53-like peptides, lsbB-like peptides, and entL-50-like peptides (13, 14).

Advancements in *in silico* detection methods of BGCs in tandem with the impressive increase of available genomes for mining have led to the expansion of modified

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

class 1 bacteriocin BGCs to the thousands (3, 4). These methods have exploited the modification machinery of RiPPs to identify novel core peptides. A recent study has expanded this search to circular bacteriocins and expanded the repertoire of peptides to over 6,000 non-unique peptides (5). A recent study identified the leaderless bacteriocin miticin through genome mining using core peptides as driver sequences (22). This novel leaderless bacteriocin was encoded in a *Streptococcus mitis* genome and has activity against *Streptococcus pyogenes*, *S. aureus*, *Lactococcus lactis*, and *L. monocytogenes*. However, in uncovering how widespread bacteriocin production is across the bacterial kingdom, current literature on characterized bacteriocins is skewed toward the phylum Bacillota. There is a paucity of knowledge from the phylum Actinomycetota related to bacteriocin production considering the role they have played in the production of traditional antibiotics (23, 24). The phylum is high-GC content and encompasses the genera *Bifidobacterium*, *Curtobacterium*, *Arcanobacterium*, *Actinomyces*, and notably, *Streptomyces,* which is considered one of the most bioactive genera in terms of antimicrobial production (25).

To our knowledge, this study is the first in-depth systematic mining study to identify novel aureocin A53-like peptides and characterize how widespread they are among bacteria by searching the non-redundant (nr) database and MAGs through MGnify (26). MGnify is a microbiome data resource with over 2.4 billion nr sequences from MAGs (26). Here, we identified 757 putative peptides in 430 genomes and synthesized two novel peptides, arachnicin from *Arachnia* sp. and arcanocin from *Arcanobacterium* sp. both from the phylum Actinomycetota. We also highlight that the Actinomycetota are an underexplored reservoir of leaderless bacteriocins with 97 peptides discovered within the phylum. The synthesized peptides, arcanocin and arachnicin from the *Arcanobacterium* and *Arachnia* genera, respectively, are also the first leaderless bacteriocins reconstructed from metagenomic data sets and the first described from within the Actinomycetota.

RESULTS

Phylogenetic distribution and biochemical properties of leaderless bacteriocins

Using a custom aureocin A53-like Hidden Markov Model (HMM) to screen the nr and MGnify databases, a total of 757 genes encoding putative leaderless bacteriocin core peptides were predicted across 95 unique species. A total of 430 genomes contained these genes. These were split between the phyla Actinomycetota (89/430; 20.7%), Bacillota (334/430; 77.7%), and Bacillota_A (7/430; 1.6%) (Fig. 1a). There are three putative novel species containing BGCs from the genera *Streptomyces* (sp. NBC_00237), *Agrococcus* (sp. SL85), and *Actinoplanes* (sp. NEAU-A12) as determined by the Genome Taxonomy Database (GTDB). A total of 322 BGCs were found encoding a single core peptide, while BGCs with multiple leaderless bacteriocins in operons were mainly within the genus *Bacillus*, with 52 BGCs with four or more core peptides. Genomes encoding six or more core peptides were almost exclusively in the genus *Bacillus*, apart from a single operon found in *Exiguobacterium indicum* containing seven core peptides (Table S1). In total, there are 75 species not currently known to produce leaderless bacteriocins, suggesting that their production is more widespread than previously thought. Some of these species include strict anaerobic bacteria within the class Clostridia and potential pathogens of etiological concern, such as *Streptococcus suis* and *Bacillus anthracis*. A total of 69 peptides were identified in *S. suis*, 51 in *Staphylococcus pseudintermedius*, 41 in *Bacillus altitudinis*, and 18 in *Arcanobacterium phocae*, which is the species within the Actinomycetota with the most core peptides (Fig. 1b). Streptococcaceae has the most BGCs, dominated by polyphyletic *S. suis* and *S. suis P* species. Of note is the existence of singleton peptides among the genera *Curtobacterium*, *Microbacterium*, and *Frondihabitans*, which are among multiple genera that have no described bacteriocins of any nature to date. Predicted peptides have a mean length of 50, a mean hydrophobicity of 0.16, and a mean charge of 4.63. An interesting note is that core peptides from

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

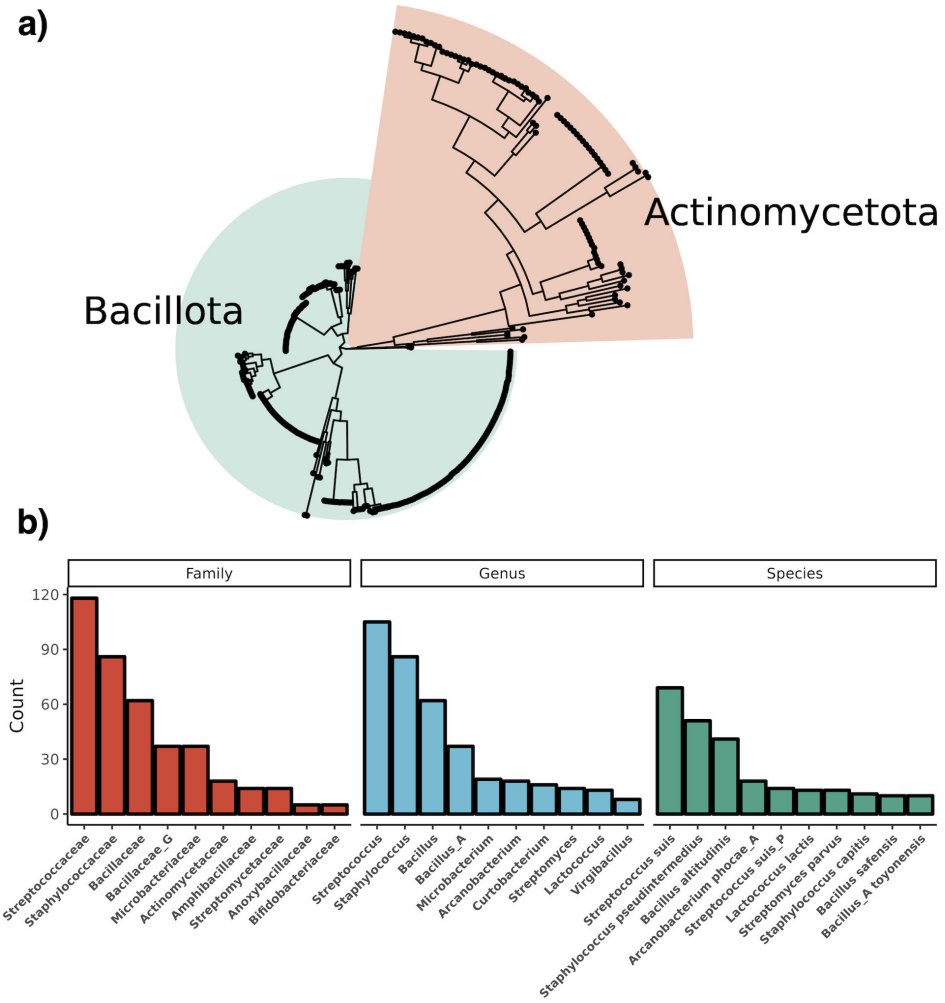


FIG 1 Leaderless bacteriocins are found widespread in the phyla Bacillota and Actinomycetota. (a) A phylogenetic tree of WGS of strains encoding leaderless bacteriocins. The phylogenetic tree was constructed using GTDB-Tk using shared marker genes and rooted to *Exiguobacterium* sp. s191 (GCF_018618955.1). The phylum Actinomycetota makes up 21% of genomes encoding a BGC. (b) A bar chart shows the top 10 species, genera, and families with putative leaderless bacteriocins. The family Streptococcaceae is dominated by *S. suis* genomes, and *A. phocae* is the species with the most core peptides within the phylum Actinomycetota.

the Actinomycetota are, on average, longer, more hydrophobic, and more positively charged as compared to peptides from the Bacillota (Fig. S1). In order to organize the diverse array of peptides, they were clustered using BLAST all-vs-all, which resulted in 31 groups (Fig. 2a). Singletons represent novel leaderless bacteriocins with unique amino acid sequences with <64.29% ID to another peptide, and the network within cluster percent identity is more than 64.29%. The majority of BGCs contained a single core peptide in their operon, and this was the case for both Actinomycetota and Bacillota. The majority of multi-peptide leaderless bacteriocin operons were among the Bacillota, with only eight assemblies having more than one core peptide within the Actinomycetota (Table S1). A total of 46 assemblies had more than four core peptides in their genome with instances of eight core peptides in a single gene cluster, as is the case with *Bacillus mycoides* that encode five proteins with 85.4% ID to thucin A1 (GCF_018739365.1). The operon lies within a putative IS6 family transposon and contains four separate, distinct core peptides with five copies of a single peptide. Another multi-component bacteriocin exists in *E. indicum* with core peptides ranging from 59.6% ID to thucin A2 and 31.25% ID to geobacillin 6 (GCF_018618955). The gene cluster is

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

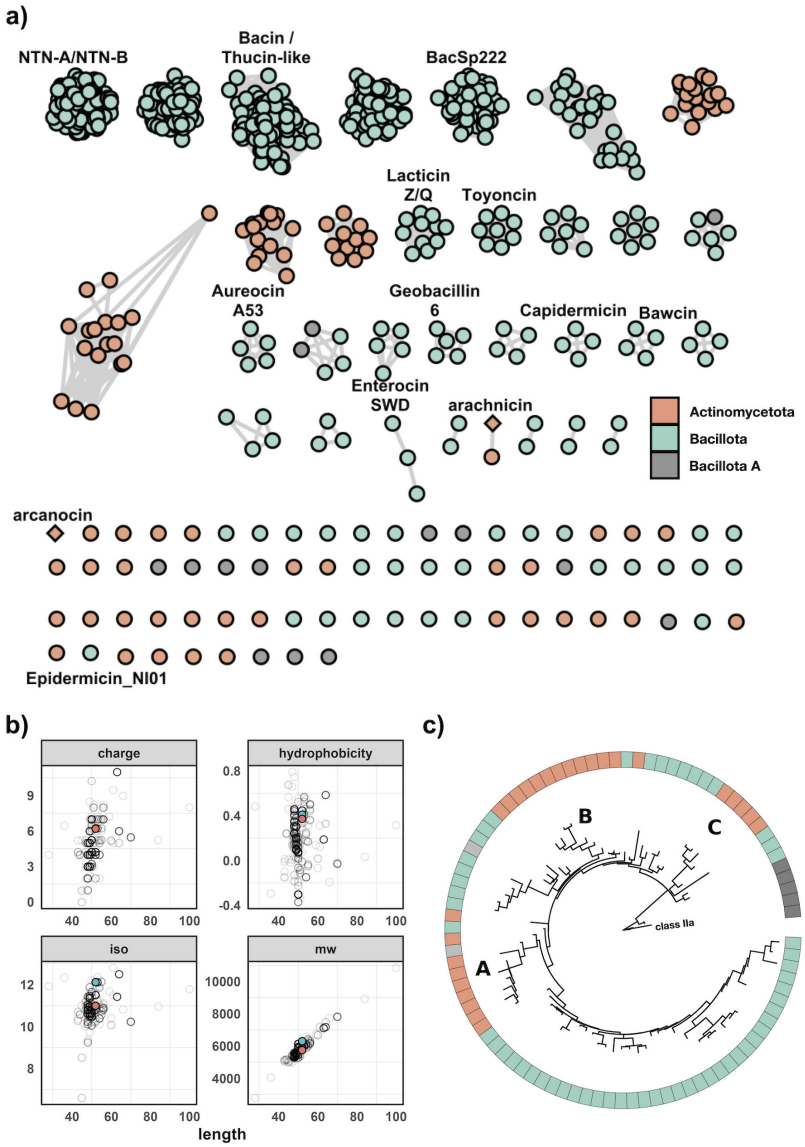


FIG 2 Insights into the diversity and physiochemical properties of Aureocin A53-like bacteriocins. (a) A SSN of leaderless bacteriocins. Orange represents leaderless bacteriocins in the phylum Actinomycetota, and green represents sequences in the phyla Bacillota. A diamond shape represents protein sequences that were synthesized. The singleton diamond is from *Arachnia,* and the non-singleton is from *Arcanobacterium*. (b) The physiological properties of all 757 leaderless bacteriocins, highlighted in green and orange, are peptides from *Arachnia* and *Arcanobacterium*, respectively, which were brought forward for peptide synthesis. Isoelectric point is labelled "iso," molecular weight in daltons is labelled "mw." (c) A reduced phylogenetic tree of all leaderless sequences clustered at 90% ID with a single representative for each cluster. The tree results from 1,000 bootstrap replicates, and each core peptide is colored by the genome's phyla containing the representative sequence in which the cluster was found. The tree is rooted to pediocin PA-1.

on a contig edge and likely contains another core peptide as there is another core peptide in the genome on a separate contig edge. This represents a genus without a described bacteriocin to date, and the gene clusters encode a Pleckstrin homology (PH) domain-containing protein and a helix-turn-helix domain-containing protein, suggesting a functionally-regulated operon. Of the eight leaderless bacteriocin operons with multiple core peptides within the phylum Actinomcyetota, seven are among the genus *Microbacterium* and a single genome from the species *Brevilactibacter sinopodophylli*,

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

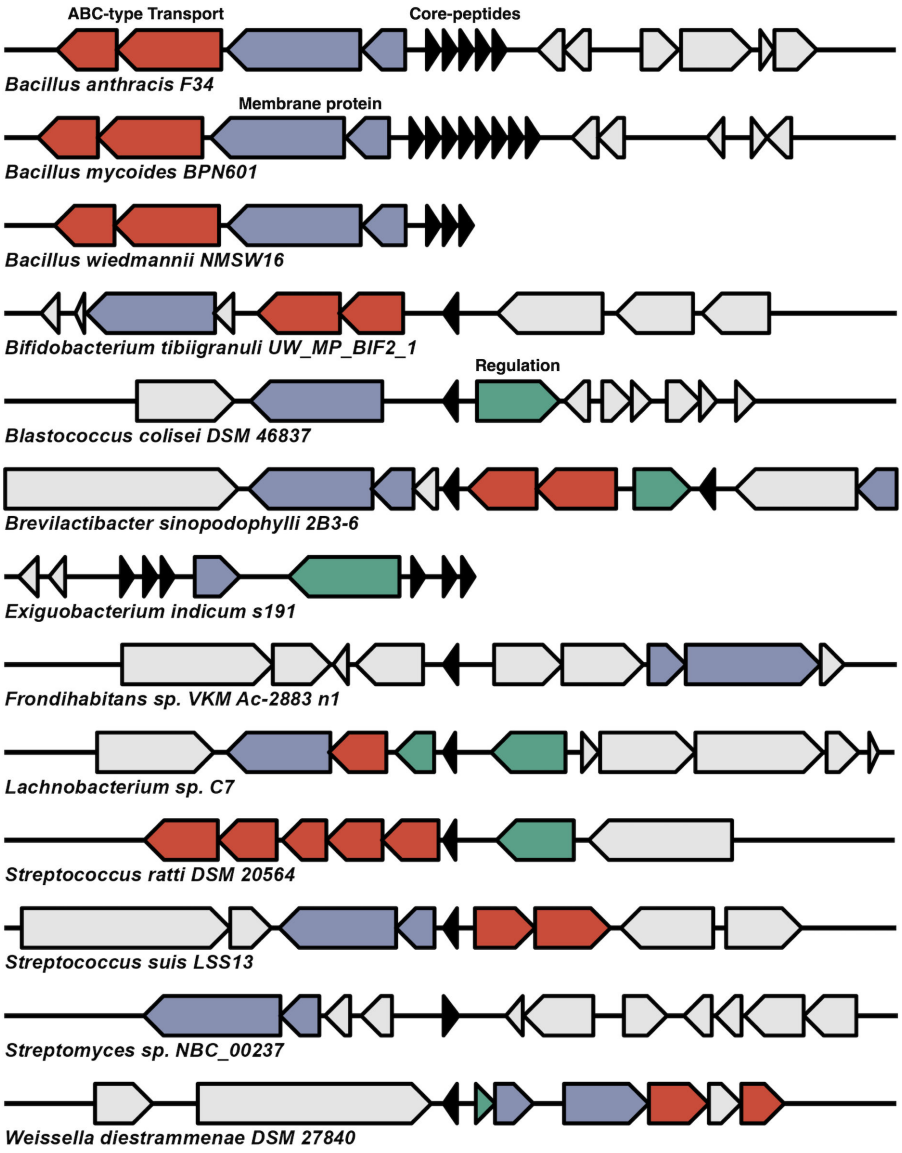


FIG 3 Operons of BGCs from assemblies deemed of particular interest. The plot depicts 13 operons from species of particular interest based on the number of core peptides present or if the operon is from a species without characterized leaderless bacteriocin operons. *Bifidobacterium*, *Blastococcus*, *Brevilactibacter*, *Exiguobacterium*, *Frondhabitans*, and *Lachnobacterium* are all genera without characterized bacteriocin production to date. Red arrows depict proteins involved in transport with an ABC transport-associated domain. Blue arrows encode membrane domain proteins such as Yip1 and PH domain-containing proteins. Green arrows are proteins involved in operon regulation, and black arrows are core peptides.

synonym *Propioniciclava sinopodophylli* with a core peptide 41.7% ID to thucin A1 (Table S1). *Brevilacibacter flavus* is a recently described novel genus with no characterized antimicrobial activity (27). All core peptides were clustered at 90% ID. A member of each cluster was used to construct a phylogenetic tree rooted to the class 2a bacteriocin pediocin PA-1 (Fig. 2c). Three distinct clades of leaderless peptides can be seen within Actinomycetota, which are designated A, B, and C. Clade A represents peptides most similar to NTN-A/B with the representative of the clade being a core peptide from *Streptomyces parvus* having <40% ID to NTN-A. Clade B is most similar to thucin-like peptides, and the representative core peptide from a *Paenarthrobacter* sp. has 43.75% ID to thucin A1. A core peptide from *Microbacterium* sp. with 33.33% ID to geobacillin 6 represents clade C.

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

Leaderless bacteriocin BGCs of interest

B. mycoides has an operon encoding eight core peptides and a shared synteny between *B. anthracis* and *B. wiedmannii*. *B. wiedmannii* is a known producer of the leaderless bacteriocin bawcin, and *B. anthracis* can be a producer of anthrax toxin, but this isolate is a non-toxigenic isolate (GCF_002024565.1). These operons share a high degree of synteny, with all three having multiple membrane-anchored proteins, including a PH domain-containing protein and ABC transport systems that consist of ABC-type lipoprotein export system multi-drug efflux pump subunit AcrA (membrane fusion protein). All three BGCs encode multiple core peptides (Fig. 3). *B. mycoides* encodes four distinct peptides with five gene copies encoding a single peptide. The five same encoded peptides are 85.4% ID to thucin A1 from *Bacillus thuringiensis* P86, and the most dissimilar is a single peptide with 60.4% to thucin A1. The *Bacillus* core peptides are shared between species (Fig. S1). *S. suis* encodes a single core peptide with 40.0% ID to lacticin Q, and the same peptide is encoded in a ST1 *S. suis* isolate from a diseased pig from the United Kingdom (GCF_000944375.1). *Lachnobacterium* sp. C7 (GCF_900113385.1) is a strict gut anaerobic bacterial isolate encoding a single core peptide with 48.8% ID to mutacin BHT-B (Fig. 3). The gene cluster encodes a putatively entire functional operon with an ABC-2 type transport system, an integral membrane protein similar to a PH domain-containing protein with putative function involved in the transport of the bacteriocin and a LacI family transcriptional regulator. *Bifidobacterium tibiigranuli* is a high-quality MAG encoding a core peptide 49.1% ID to lacticin Q from *L. lactis*. The operon encodes a membrane-associated PH domain and a dedicated ABC transport system. *Blastococcus colisei* BMG 822^T is a slow-growing member of the family *Geodermatophilaceae* isolated from an archaeological amphitheatre in Tunisia, and the genus has no characterized bacteriocin production (28). The isolate encodes a core peptide with 46% ID to lacticin Q, a PH domain-containing protein, a transcriptional regulator protein and is located less than 2 kb from a site-specific integrase (Fig. 3.). *Brevilactibacter sinopodophylli* (GCF_947488915.1) is a bacterium isolated from the phycosphere of marine macroalgae. It is within the family *Propionibacteriaceae* and the only member of the genus *Brevilactibacter* with a predicted BGC in this study. The strain encodes two distinct core peptides with 41.7% ID to thucin A1 and 39.6% to lacticin Q. The operon also encodes a PH domain-containing protein, an ABC transport system, and a DNA-binding response regulator, suggesting a functional operon.

Novel aureocin A53-like bacteriocins from *Arachnia* and *Arcanobacterium*

Two peptides not found in the nr database were chosen for synthesis as they were of human microbiome origin, and both have low percentage amino acid identity to known aureocin A53-like peptides, arachnicin from *Arachnia*, and arcanocin from *Arcanobacterium*. Arachnicin shares 38.46% ID to lacticin Q, and arcanocin shares 50.00% ID to lacticin Q. The operons encoding the leaderless bacteriocin BGCs from the genera *Arcanobacterium* (MGYG000000642) and *Aracnhia* (MGYG000298963) were found in gut and oral microbiomes, respectively. *Arcanobacterium* was assembled from a fecal stool sample from rural communities in north-eastern Madagascar (PRJNA485056), and *Aracnhia* was assembled from an oral microbiome assessing the clinical relevance of a dental implant microbiome (PRJEB43277). Both high-quality MAGs have over 96% completeness and less than 2% contamination, and both represent potential novel species within their genera. The operons found in both MAGs are less than 5-kb long but have a low degree of synteny and vary in gene content (Fig. 4a). The operon in *Arcanobacterium* sp. is longer and encodes seven genes putatively involved in the production of arcanocin. These include a PH domain-containing protein (dark blue), ABC transport machinery (red) and a putative bPH_2 domain-containing protein (light blue). The BGC from *Arcanobacterium* also has a longer contiguous operon with a putative PH domain-containing protein and a nucleotide excision repair protein that is not present in the *Arachnia* BGC. The BGC from *Arachnia* sp. has four genes associated with leaderless bacteriocin production, but the putative regulatory protein and transport system are on opposing strands to the core

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

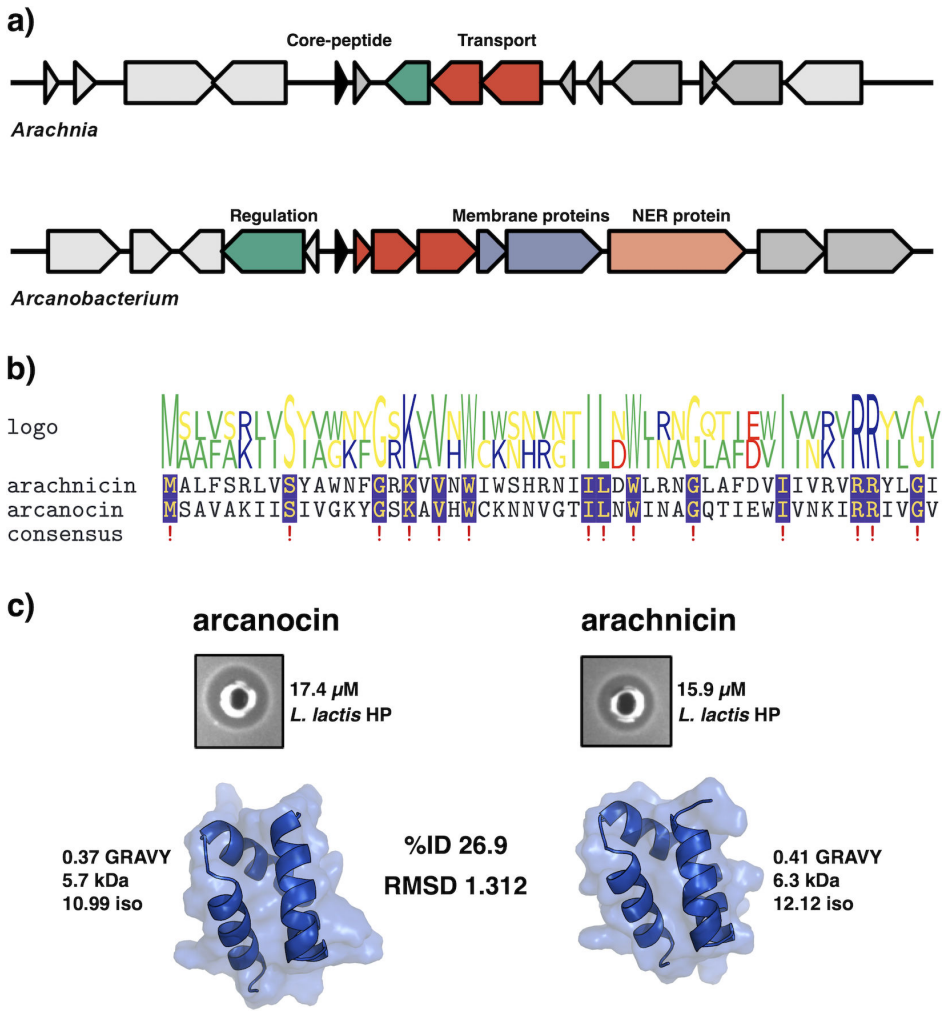


FIG 4 Characterization of two novel leaderless bacteriocin BGCs encoded by *Arcanobacterium* spp. and *Arachnia* spp. with novel core peptides found in the MGnify database. (a) BGCs do not share a high degree of synteny. Black arrows are the core peptides, red arrows are transport machinery, green arrows are regulatory proteins, and blue is a PH domain-containing protein. The blue arrow represents a bPH_2 domain-containing protein. The light orange arrow is a nucleotide excision repair (NER) protein that is a member of the excinuclease ABC subunit B responsible for recognizing damaged DNA. (b) A multiple sequence alignment of the two novel leaderless bacteriocins. MGYG000298963_01692 (arachnicin) is from *Arachnia*, and MGYG000000642_01040 (arcanocin) is from *Arcanobacterium*. Although the peptides differ in amino acid content (26.9%), they share 14 conserved residues, including two tryptophan residues. Both core peptides have three to four tryptophan residues, but only two are conserved in position. (c) The structure of leaderless bacteriocins from *Arachnia* and *Arcanobacterium* share a relatively low sequence identity of 26.9% but conforms to almost identical tertiary structures with an RMSD of 1.312. Arcanocin and arachnicin have antimicrobial activity against *L. lactis* HP at 100 µg/mL.

peptide. It is likely the operon is non-functional due to the organization of the operon, but it is possible that its non-contiguous nature could be a form of regulation (29, 30). Both short open reading frames beside each core peptide are hypothetical proteins. The BGCs encode single core peptides that are each 52 amino acids in length and share 14 residues (26.9% ID). Notably, the core peptides share tryptophan residues at positions 21 and 32 (Fig. 4b). The peptides also have three to four tryptophan residues in total, which probably play a role in their mechanism of action and are crucial for activity in some leaderless bacteriocin cases (15). Arachnicin (MGYG000298963_01692) has 38.5% ID to lacticin Q, and arcanocin (MGYG000000642_01040) has 53.8% ID to lacticin Z. Arcanocin and arachnicin are structurally identical with a root mean square deviation (RMSD) <

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

1.312 (Fig. 4c). The peptides have different physiochemical profiles. Arachnicin is more hydrophobic (0.41 compared to 0.37 GRAVY index), higher in molecular weight (6.3 kDa compared to 5.7 kDa), and has a much higher isoelectric point compared to arcanocin (12.12 compared to 10.99) (Fig. 2b). Both peptides were chemically synthesized. Synthesized arcanocin was soluble in H_2O + 0.1 trifluoroacetic acid (TFA) at 100 µg/mL, whereas synthesized arachnicin was insoluble and had to be resuspended in acetonitrile (ACN) + 0.1% TFA. Both peptides were active against *L. lactis* HP at a concentration of 100 µg/mL (15.9 µM and 17.4 µM for arachnicin and arcanocin, respectively; Fig. 4c).

DISCUSSION

In this study, we identified 757 aureocin A53-like bacteriocins from the nr and MGnify databases, highlighting that they are more widespread than previously thought and are prevalent in the Actinomycetota. *Streptomyces* are one of the most productive bacteria in terms of biosynthetic potential, and recent focus has shifted to other members of the class Actinomycetes, such as *Nocardia* species, where some bacteria can have up to 70 BGCs (31). A total of 97 were identified within the phylum with a mean length of 55 amino acids. To date, no bacteriocins have been described in the genera *Arcanobacterium* and *Aracnhia*, nor has any leaderless bacteriocin-producing organism been characterized within the Actinomycetota. *Arcanobacterium* are facultative anaerobic gram-positive bacteria from the Actinomycetota. The genus currently has 63 species spread across the mammalian microbiota, with isolates from a harbor seal and rhinoceros (32, 33). The genus *Arcanobacterium pyogenes* was reclassified into the genus *Trueperella* in 2011, and this is now creating interest as a skin pathobiont named *Trueperella pyogenes,* which is a common cause of "summer mastitis" and pneumonia in bovine herds and pyometra in canines, respectively (34, 35). One of the most well-studied members of the genus is the rare pathogen *Arcanobacterium haemolyticum*, which is a facultative anaerobic bacterium previously classified as a *Corynebacterium* sp. (36).

The peptides arcanocin and arachnicin were chosen for gene synthesis based on low % ID to characterized leaderless bacteriocins and that they were found in MAGs assembled from the human gut and oral microbiome, respectively. They were also not present in the nr database at the time of synthesis, and they are from genera without characterized bacteriocin production. Arachnicin is most similar to lacticin Q (38.5% ID) and is one of the more hydrophobic peptides with one of the highest isoelectric points in the data set, and the sequence has two aspartate residues with acidic side chains. The peptide was not soluble in H_2O + 0.1% TFA alone, likely because of these properties. Due to this, the peptide was resuspended in ACN + 0.1% TFA and presented antimicrobial activity against *L. lactis* HP at 100 µg/mL. TFA is a widely used chemical solvent used in the purification of peptides and has been used to solubilize garvicin KS and micrococcin P1 (37, 38). It is an effective solvent of hydrophobic peptides (39). However, the solvent is toxic to eukaryotic cells and is considered a contaminant during downstream purification of proteins after reversed-phase high-performance liquid chromatography (40). The removal of TFA as a solvent while retaining antimicrobial activity would be required if arachnicin was to be considered as an antimicrobial agent. TFA has been replaced with HCl during purification of pediocin PA-1 and the end product retained equal activity (39). Similarly, the bacteriocin nisin has 10-fold increased solubility at pH 3.0 compared to pH 7.0, and bio-engineered variants of the peptide have increase solubility which suggests arachnicin could be engineered to improve overall solubility at a higher pH (41). *Arachnia* is a genus from the gram-positive family Propionibacteriaceae that produces the metabolic by-product propionic acid (42, 43). There are only two species within the genus, *Arachnia propionica* and *Arachnia rubra*, and both are members of the oral microbiome (44). Interestingly, this family contains *Cutibacterium acnes*, a commensal bacterium with a clade, clade A1, associated with acne vulgaris and a causative agent of contamination of medical devices (45). The arcanocin peptide shares 26.9% ID and has a highly similar tertiary structure (RMSD 1.312). The peptide has a much lower isoelectric

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

point than arachnicin, was soluble in H_2O + 0.1% TFA, and was effective against *L. lactis* HP at 100 µg/mL.

PH domain-containing proteins are commonly found within BGCs. These proteins are found in many domains of life, play a part in cellular trafficking in eukaryotic cells, and are lipid-binding. Recently, it has been shown that PH domain-containing proteins have a role in the immunity and transport of leaderless bacteriocins (46). Interestingly, the formation of oligomeric ring structures that may resemble membrane-bound transporters supports this evidence, with the PH domain likely playing a role in oligomerization (47). However, PH domain-containing proteins are not present in all leaderless bacteriocin BGCs, suggesting there are multiple mechanisms that can be involved in providing immunity and transport of the core peptide after translation, and these mechanisms could be unique to the sub-classification of leaderless bacteriocin. This can be seen with the operons encoding enterocin DD14 and enterocin Q, where the former encodes PH domain-containing proteins involved in transport, whereas the latter does not (14, 46, 48). Another observation is Yip1 domain-containing proteins within the BGCs, which can also be observed in circular bacteriocin BGCs (5). The Yip1 family of proteins are also involved in cellular trafficking within eukaryotes, and they possess multiple transmembrane domains potentially suggesting a role in immunity or transport (49). Leaderless bacteriocin BGCs do not have a "class defining protein" such as the DUF95 domain membrane protein of circular bacteriocins, nor do they have a highly conserved immunity mechanism. The peptides are also bioactive after translation and are usually encoded with different membrane-associated protein families, suggesting a convergent evolution of immunity systems to protect from post-translational cell suicide. A future method to find dissimilar peptides of novel classes could be to use Yip1 domain-containing proteins and PH domain-containing proteins as driver sequences, as seen with the DUF95/SpoIIM protein in circular bacteriocins (50).

One exploitable trait of leaderless bacteriocins is the ability to capitalize on their "active after translation" trait because the N-formylation of methionine is not crucial for the antimicrobial activity of leaderless bacteriocins, such as aureocin A53 and lacticin Q (10). This trait allows chemical synthesis without needing accessory genes to produce active peptides. This can be exploited by chemically synthesizing peptides found in MAGs, which can dramatically reduce costs associated with production and allow the analysis of bioactive peptides based on DNA sequence alone.

This work identified 757 aureocin A53-like leaderless bacteriocins, 97 of which are in the phylum Actinomycetota, highlighting that they are present beyond the Bacillota. Chemically synthesized arcanocin and arachnicin, two novel leaderless bacteriocins from *Arcanobacterium* and *Arachnia*, respectively, are the first leaderless bacteriocins from the phylum Actinomycetota with demonstrated antibacterial activity. Actinomycetota is a phylum with a higher percentage of hypothetical proteins within their genomes than the Bacillota (51). This highlights the potential of looking beyond the Bacillota for novel bacteriocin production.

MATERIALS AND METHODS

Genome mining

The nr and MGnify databases were downloaded in FASTA format in November 2023. MGnify protein databases included human gut, human oral, human vaginal, bee gut, rumen, pig gut, chicken gut, fish gut, marine, and zebrafish gut. MGnify databases were reduced to include proteins under 100 residues only using Seqkit v2.3.0 and the command "seqkit seq -M 100" (52). A leaderless bacteriocin HMM was created with known aureocin A53-like bacteriocins by aligning sequences with muscle v3.8.31 (53), and a model was created with the sequence alignment using HMMER v3.3.1 with the command "hmmbuild" (54). This model was used to search the MGnify protein database, and these were added to create a second model that was then searched back into both databases. All hits below an *e*-value of 1^{e-5} across the whole protein were brought

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

forward as putative leaderless bacteriocins. Rodeo2 v2.3.3 was used to retrieve contig accessions, and assembly accessions were retrieved using Biopython v1.81 (55).

Phylogenetic tree of organisms

Whole genome assemblies were downloaded using nbci-genome-download v0.3.3 (https://github.com/kblin/ncbi-genome-download). GTDB-Tk v2.3.1 was used with the commands "identify, align, infer" without the GTDB data included to create a *de novo* phylogenetic tree (56). A tree was created using FastTree and visualized using ggtree v3.18. Taxonomy was assigned to whole genomes using GTDB-Tk "classify --skip_ani_screen" using pplacer v1.1.alpha19-0-g807f6f3.

Sequence similarity network

Core peptides were identified from the ".faa" files generated by Bakta by using the previously mentioned leaderless bacteriocin HMM using the command "hmmsearch --tblout." These sequences were then retrieved using the command "grep -hA 1" under the rule get_proteins of the Snakemake workflow and saved as "final_proteins.faa." This file was then uploaded to the Enzyme Function Initiative-The Enzyme Similarity Tool (EFI-EST) (57). A sequence similarity network (SSN) was constructed of the core peptides only using EFI-EST with a blast value of 1^{e-05}. The corresponding network was visualized in Cytoscape v3.10.1 with the "prefuse force-directed" layout, and nodes were colored based on the GTDB phyla.

Phylogenetic tree of peptides

Sequences were clustered at 90% identity using CD-HIT v4.8.1, and sequences were aligned using muscle v3.8.31. RAxML-NG v1.2.0 was used to construct a phylogenetic tree with the options "--model LG, --bs-trees 1000 --bs-metric tbe,fbp" (58). The tree includes the sequences for pediocin PA-1, enterocin A, enterocin CRL35, and sakacin *P* and was rooted to pediocin PA-1 using the R package ape v5.8. The tree was plotted using ggtree v3.18 (59).

Peptide properties

Core peptide properties were predicted using the R package Peptides v2.4.6 (https://cran.r-project.org/web/packages/Peptides/index.html). The properties chosen were hydrophobicity, isoelectric point, molecular weight, and charge. Hydrophobicity was calculated using the GRAVY hydrophobicity using the "KyteDoolittle" scale. The isoelectric point was calculated using the "EMBOSS" pKscale. Properties were plotted using the ggplot2 v3.4.4. R package and R v4.3.0. Protein structures were predicted using Alphafold v2.3.2 with the setting "monomer" (60). A % ID matrix was created using predicted proteins and known leaderless bacteriocins using clustalO v1.2.4 and the parameters "--distmat-out --threads 12 --percent-id --full". This was then used to compare predicted leaderless bacteriocins to known peptides. Only distinct core peptides were brought forward to calculate differences between aureocin A53-like peptides from Actinomycetota and Bacillota. The Kruskal–Wallis test was used to determine the statistical significance between the two groups using the stats v3.6.2 R package.

Synteny plots

Assemblies were annotated using Bakta v1.8.1 with the Bakta database v5.0 (61). Operons were plotted using gggenomes v0.9.9.9000 and R v4.3.

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

Peptide synthesis

Peptides arcanocin and arachnicin were synthesized by ProteoGenix (Schiltigheim, France). Arcanocin was resuspended in H_2O + 0.1% TFA to a final concentration of 100 µg/mL. Arachnicin was resuspended in ACN + 0.1% TFA.

Well diffusion assay

L. lactis HP was grown overnight using GM17 media at 30°C. Agar (0.75% wt/vol) was seeded with 0.2% (vol/vol) culture. The wells were created using a sterile Pasteur pipette. Fifty microliters of crude peptide solution (100 µg/mL) was added to the well and incubated at 30°C overnight. The antimicrobial activity was determined by the presence of an inhibition halo. The results were obtained from two independent replicates.

ACKNOWLEDGMENTS

Co-funded by Science Foundation Ireland (SFI) under Grant Number SFI/ 12/RC/2273_P2 and the European Union (ERC, BACtheWINNER, 101054719). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them. This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the INSPIRE COFUND Marie Skłodowska-Curie grant agreement No. 101034270.

D.H. conceived and carried out genome mining, constructed the figures and wrote the manuscript. F.M.D.F. and P.O.C. carried out wet lab experiments. F.M.D.F. created the database of leaderless bacteriocins. C.H. and P.R. reviewed and contributed to the final manuscript.

AUTHOR AFFILIATIONS

[1]APC Microbiome Ireland, University College Cork, Cork, Ireland
[2]School of Microbiology, University College Cork, Cork, Ireland
[3]Teagasc Food Research Centre, Cork, Ireland

AUTHOR ORCIDs

David Hourigan http://orcid.org/0000-0003-0278-7858
Felipe Miceli de Farias http://orcid.org/0000-0001-5419-9922
Colin Hill http://orcid.org/0000-0002-8527-1445
R. Paul Ross http://orcid.org/0000-0003-4876-8839

FUNDING

Funder	Grant(s)	Author(s)
Science Foundation Ireland (SFI)	SFI/ 12/RC/2273_P2	Colin Hill R. Paul Ross
EC | European Research Council (ERC)	101054719	R. Paul Ross
EC | H2020 | PRIORITY 'Excellent science' | H2020 Marie Skłodowska-Curie Actions (MSCA)	101034270	Felipe Miceli de Farias

AUTHOR CONTRIBUTIONS

David Hourigan, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review and editing | Felipe Miceli de Farias, Conceptualization, Data curation, Writing – review and editing | Paula M. O'Connor, Formal analysis, Methodology, Resources, Writing – review and editing | Colin Hill, Funding acquisition, Project administration, Supervision, Writing – review and

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

editing | R. Paul Ross, Funding acquisition, Project administration, Supervision, Writing – review and editing

DATA AVAILABILITY

Scripts and methods used in this study are stored in https://github.com/DEHourigan/leaderless_project

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental figures (JB00298-24-s0001.pdf). Fig. S1 to S9.
Table S1 (JB00298-24-s0002.csv). This table includes the protein coding sequences for the discovered peptides.
Table S2 (JB00298-24-s0003.xlsx). This table contains supplementary table S1 in excel format and the sequences found in each cluster of the sequence similarity network.
Table S3 (JB00298-24-s0004.csv). This table contains the edge information for the sequence similarity network.
Table S4 (JB00298-24-s0005.csv). This table contains the node information for the sequence similarity network of core peptides.

REFERENCES

1. Klaenhammer TR. 1988. Bacteriocins of lactic acid bacteria. Biochimie 70:337–349. https://doi.org/10.1016/0300-9084(88)90206-4
2. Sugrue I, Ross RP, Hill C. 2024. Bacteriocin diversity, function, discovery and application as antimicrobials. Nat Rev Microbiol 22:556–571. https://doi.org/10.1038/s41579-024-01045-x
3. Walker MC, Eslami SM, Hetrick KJ, Ackenhusen SE, Mitchell DA, van der Donk WA. 2020. Precursor peptide-targeted mining of more than one hundred thousand genomes expands the lanthipeptide natural product family. BMC Genomics 21:387. https://doi.org/10.1186/s12864-020-06785-7
4. Schwalen CJ, Hudson GA, Kille B, Mitchell DA. 2018. Bioinformatic expansion and discovery of thiopeptide antibiotics. J Am Chem Soc 140:9494–9501. https://doi.org/10.1021/jacs.8b03896
5. Xin B, Liu H, Zheng J, Xie C, Gao Y, Dai D, Peng D, Ruan L, Chen H, Sun M. 2020. *In silico* analysis highlights the diversity and novelty of circular bacteriocins in sequenced microbial genomes. mSystems 5:e00047-20. https://doi.org/10.1128/mSystems.00047-20
6. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, Lawley TD, Finn RD. 2019. A new genomic blueprint of the human gut microbiota. Nature New Biol 568:499–504. https://doi.org/10.1038/s41586-019-0965-1
7. Anderson CL, Fernando SC. 2021. Insights into rumen microbial biosynthetic gene cluster diversity through genome-resolved metagenomics. Commun Biol 4:818. https://doi.org/10.1038/s42003-021-02331-7
8. Won MY, Oyama LB, Courtney SJ, Creevey CJ, Huws SA. 2020. Can rumen bacteria communicate to each other? Microbiome 8:23. https://doi.org/10.1186/s40168-020-00796-y
9. Iwatani S, Yoneyama F, Miyashita S, Zendo T, Nakayama J, Sonomoto K. 2012. Identification of the genes involved in the secretion and self-immunity of lacticin Q, an unmodified leaderless bacteriocin from *Lactococcus lactis* QU 5. Microbiology (Reading) 158:2927–2935. https://doi.org/10.1099/mic.0.062943-0
10. Acedo JZ, van Belkum MJ, Lohans CT, Towle KM, Miskolzie M, Vederas JC. 2016. Nuclear magnetic resonance solution structures of lacticin Q and aureocin A53 reveal a structural motif conserved among leaderless bacteriocins with broad-spectrum activity. Biochemistry 55:733–742. https://doi.org/10.1021/acs.biochem.5b01306
11. Sánchez-Barrena MJ, Martínez-Ripoll M, Gálvez A, Valdivia E, Maqueda M, Cruz V, Albert A. 2003. Structure of bacteriocin AS-48: from soluble state to membrane bound state. J Mol Biol 334:541–549. https://doi.org/10.1016/j.jmb.2003.09.060
12. Netz DJA, Pohl R, Beck-Sickinger AG, Selmer T, Pierik AJ, Bastos M do C de F, Sahl H-G. 2002. Biochemical characterisation and genetic analysis of aureocin A53, a new, atypical bacteriocin from *Staphylococcus aureus*. J Mol Biol 319:745–756. https://doi.org/10.1016/S0022-2836(02)00368-6
13. Perez RH, Zendo T, Sonomoto K. 2018. Circular and leaderless bacteriocins: biosynthesis, mode of action, applications, and prospects. Front Microbiol 9:2085. https://doi.org/10.3389/fmicb.2018.02085
14. Ladjouzi R, Dussert E, Teiar R, Belguesmia Y, Drider D. 2023. A review on enterocin DD14, the leaderless two-peptide bacteriocin with multiple biological functions and unusual transport pathway. Antibiotics (Basel) 12:1188. https://doi.org/10.3390/antibiotics12071188
15. Lander AJ, Mercado LD, Li X, Taily IM, Findlay BL, Jin Y, Luk LYP. 2023. Roles of inter- and intramolecular tryptophan interactions in membrane-active proteins revealed by racemic protein crystallography. Commun Chem 6:154. https://doi.org/10.1038/s42004-023-00953-y
16. Nascimento JDS, Coelho MLV, Ceotto H, Potter A, Fleming LR, Salehian Z, Nes IF, Bastos M do C de F. 2012. Genes involved in immunity to and secretion of aureocin A53, an atypical class II bacteriocin produced by *Staphylococcus aureus* A53. J Bacteriol 194:875–883. https://doi.org/10.1128/JB.06203-11
17. Netz DJA, Bastos M do C de F, Sahl H-G. 2002. Mode of action of the antimicrobial peptide aureocin A53 from *Staphylococcus aureus*. Appl Environ Microbiol 68:5274–5280. https://doi.org/10.1128/AEM.68.11.5274-5280.2002
18. Fagundes PC, Farias FM, Santos OCS, de Oliveira NEM, da Paz JAS, Ceotto-Vigoder H, Alviano DS, Romanos MTV, Bastos MCF. 2016. The antimicrobial peptide aureocin A53 as an alternative agent for biopreservation of dairy products. J Appl Microbiol 121:435–444. https://doi.org/10.1111/jam.13189
19. Marques-Bastos SLS, Coelho MLV, de Sousa Santos IN, Moreno DSA, Barrias ES, de Mendonça JFM, Mendonça LC, Lange CC, de Paiva Brito MAV, do Carmo de Freire Bastos M. 2023. Effects of the natural antimicrobial peptide aureocin A53 on cells of *Staphylococcus aureus* and *Streptococcus agalactiae* involved in bovine mastitis in the excised teat model. World J Microbiol Biotechnol 39:1–14. https://doi.org/10.1007/s11274-022-03443-w
20. Budhwani Z, Buragina JT, Lang J, Acedo JZ. 2023. Characterization of the novel leaderless bacteriocin, bawcin, from *Bacillus wiedmannii*. Int J Mol Sci 24:16965. https://doi.org/10.3390/ijms242316965

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

21. Wang J, Xu H, Liu S, Song B, Liu H, Li F, Deng S, Wang G, Zeng H, Zeng X, Xu D, Zhang B, Xin B. 2021. Toyoncin, a novel leaderless bacteriocin that is produced by bacillus toyonensis XIN-YC13 and specifically targets *B. cereus* and *Listeria monocytogenes*. Appl Environ Microbiol 87:e0018521. https://doi.org/10.1128/AEM.00185-21
22. MsD A, MsJ S, MsZ B, MsJT B, MsL L, DrJ A. 2024. Discovery of leaderless bacteriocins through genome mining. Available from: https://doi.org/10.1139/CJC-2023-0209
23. van der Meij A, Worsley SF, Hutchings MI, van Wezel GP. 2017. Chemical ecology of antibiotic production by Actinomycetes. FEMS Microbiol Rev 41:392–416. https://doi.org/10.1093/femsre/fux005
24. Sugrue I, O'Connor PM, Hill C, Stanton C, Ross RP. 2020. *Actinomyces* produces defensin-like bacteriocins (Actifensins) with a highly degenerate structure and broad antimicrobial activity. J Bacteriol 202:e00529-19. https://doi.org/10.1128/JB.00529-19
25. de Lima Procópio RE, da Silva IR, Martins MK, de Azevedo JL, de Araújo JM. 2012. Antibiotics produced by *Streptomyces*. Braz J Infect Dis 16:466–471. https://doi.org/10.1016/j.bjid.2012.08.014
26. Richardson L, Allen B, Baldi G, Beracochea M, Bileschi ML, Burdett T, Burgin J, Caballero-Pérez J, Cochrane G, Colwell LJ, Curtis T, Escobar-Zepeda A, Gurbich TA, Kale V, Korobeynikov A, Raj S, Rogers AB, Sakharova E, Sanchez S, Wilkinson DJ, Finn RD. 2023. MGnify: the microbiome sequence data analysis resource in 2023. Nucleic Acids Res 51:D753–D759. https://doi.org/10.1093/nar/gkac1080
27. Wenning M, Breitenwieser F, Huptas C, Doll E, Bächler B, Schulz A, Dunkel A, Hofmann T, von Neubeck M, Busse H-J, Scherer S. 2020. *Brevilactibacter flavus* gen. nov., sp. nov., a novel bacterium of the family *Propionibacteriaceae* isolated from raw milk and dairy products and reclassification of *Propioniciclava sinopodophylli* as *Brevilactibacter sinopodophylli* comb. nov. Int J Syst Evol Microbiol 70:2186–2193. https://doi.org/10.1099/ijsem.0.003909
28. Hezbri K, Nouioui I, Rohde M, Schumann P, Gtari M, Klenk H-P, Montero-Calasanz M del C, Ghodhbane-Gtari F. 2017. *Blastococcus colisei* sp. nov, isolated from an archaeological amphitheatre. Antonie Van Leeuwenhoek 110:339–346. https://doi.org/10.1007/s10482-016-0804-4
29. Iturbe P, Martín AS, Hamamoto H, Marcet-Houben M, Galbaldón T, Solano C, Lasa I. 2024. Noncontiguous operon atlas for the *Staphylococcus aureus* genome. Microlife 5:uqae007. https://doi.org/10.1093/femsml/uqae007
30. Sáenz-Lahoya S, Bitarte N, García B, Burgui S, Vergara-Irigaray M, Valle J, Solano C, Toledo-Arana A, Lasa I. 2019. Noncontiguous operon is a genetic organization for coordinating bacterial gene expression. Proc Natl Acad Sci U S A 116:1733–1738. https://doi.org/10.1073/pnas.1812746116
31. Männle D, McKinnie SMK, Mantri SS, Steinke K, Lu Z, Moore BS, Ziemert N, Kaysser L. 2020. Comparative genomics and metabolomics in the genus *Nocardia*. mSystems 5:e00125-20. https://doi.org/10.1128/mSystems.00125-20
32. Borowiak M, Kreitlow A, Malorny B, Alssahen M, Lämmler C, Prenger-Berninghoff E, Ewers C, Siebert U, Plötz M, Abdulmawjood A. 2023. *Arcanobacterium pinnipediorum* strain DSM 28752 isolated from a harbour seal: complete genome sequence. Microbiol Resour Announc 12:e0118022. https://doi.org/10.1128/mra.01180-22
33. Sammra O, Rau J, Wickhorst JP, Alssahen M, Hassan AA, Lämmler C, Kämpfer P, Glaeser SP, Busse HJ, Kleinhagauer T, Knauf-Witzens T, Prenger-Berninghoff E, Abdulmawjood A, Klein G. 2017. *Arcanobacterium wilhelmae* sp. nov., isolated from the genital tract of a rhinoceros (*Rhinoceros unicornis*). Int J Syst Evol Microbiol 67:2093–2097. https://doi.org/10.1099/ijsem.0.001784
34. Jost BH, Billington SJ. 2005. *Arcanobacterium pyogenes*: molecular pathogenesis of an animal opportunist. Antonie Van Leeuwenhoek 88:87–102. https://doi.org/10.1007/s10482-005-2316-5
35. Hijazin M, Prenger-Berninghoff E, Sammra O, Alber J, Lämmler C, Kämpfer P, Glaeser SP, Busse HJ, Hassan AA, Abdulmawjood A, Zschöck M. 2012. *Arcanobacterium canis* sp. nov., isolated from otitis externa of a dog, and emended description of the genus *Arcanobacterium* Collins et al. 1983 emend. Yassin et al. 2011. Int J Syst Evol Microbiol 62:2201–2205. https://doi.org/10.1099/ijs.0.037150-0
36. Collins MD, Jones D, Schofield GM. 1982. Reclassification of'corynebacterium haemolyticum. J Gen Microbiol 128:1279–1281. https://doi.org/10.1099/00221287-128-6-1279
37. Dubey S, Diep DB, Evensen Ø, Munang'andu HM. 2022. Garvicin KS, a broad-spectrum bacteriocin protects zebrafish larvae against *Lactococcus garvieae* Infection. Int J Mol Sci 23:2833. https://doi.org/10.3390/ijms23052833
38. Kranjec C, Ovchinnikov KV, Grønseth T, Ebineshan K, Srikantam A, Diep DB. 2020. A bacteriocin-based antimicrobial formulation to effectively disrupt the cell viability of methicillin-resistant *Staphylococcus aureus* (MRSA) biofilms. NPJ Biofilms Microbiomes 6:58. https://doi.org/10.1038/s41522-020-00166-4
39. Gaussier H, Morency H, Lavoie MC, Subirade M. 2002. Replacement of trifluoroacetic acid with HCl in the hydrophobic purification steps of pediocin PA-1: a structural effect. Appl Environ Microbiol 68:4803–4808. https://doi.org/10.1128/AEM.68.10.4803-4808.2002
40. Cornish J, Callon KE, Lin CQX, Xiao CL, Mulvey TB, Cooper GJS, Reid IR. 1999. Trifluoroacetate, a contaminant in purified proteins, inhibits proliferation of osteoblasts and chondrocytes. Am J Physiol Endocrinol Metab 277:E779–E783. https://doi.org/10.1152/ajpendo.1999.277.5.E779
41. Rollema HS, Kuipers OP, Both P, de Vos WM, Siezen RJ. 1995. Improvement of solubility and stability of the antimicrobial peptide nisin by protein engineering. Appl Environ Microbiol 61:2873–2878. https://doi.org/10.1128/aem.61.8.2873-2878.1995
42. Schmid-Wendtner MH, Korting HC. 2006. The pH of the skin surface and its impact on the barrier function. Skin Pharmacol Physiol 19:296–302. https://doi.org/10.1159/000094670
43. Tax G, Urbán E, Palotás Z, Puskás R, Kónya Z, Bíró T, Kemény L, Szabó K. 2016. Propionic acid produced by *Propionibacterium acnes* strains contributes to their pathogenicity. Acta Derm Venereol 96:43–49. https://doi.org/10.2340/00015555-2154
44. Ouellette M, Bejian AA, Chen T, Jones DS, Johnston CD, Dewhirst FE. 2021. Complete genome sequence of *Arachnia rubra* strain DSM 100122T, a cultured member of the human oral microbiome. Microbiol Resour Announc 10:e0095921. https://doi.org/10.1128/MRA.00959-21
45. Conwill A, Kuan AC, Damerla R, Poret AJ, Baker JS, Tripp AD, Alm EJ, Lieberman TD. 2022. Anatomy promotes neutral coexistence of strains in the human skin microbiome. Cell Host Microbe 30:171–182. https://doi.org/10.1016/j.chom.2021.12.007
46. Pérez-Ramos A, Ladjouzi R, Benachour A, Drider D. 2021. Evidence for the involvement of pleckstrin homology domain-containing proteins in the transport of enterocin DD14 (EntDD14); a leaderless two-peptide bacteriocin. Int J Mol Sci 22:12877. https://doi.org/10.3390/-ijms222312877
47. Xu Q, Bateman A, Finn RD, Abdubek P, Astakhova T, Axelrod HL, Bakolitsa C, Carlton D, Chen C, Chiu H-J, et al. 2010. Bacterial pleckstrin homology domains: a prokaryotic origin for the PH domain. J Mol Biol 396:31–46. https://doi.org/10.1016/j.jmb.2009.11.006
48. Criado R, Diep DB, Aakra A, Gutiérrez J, Nes IF, Hernández PE, Cintas LM. 2006. Complete sequence of the enterocin Q-encoding plasmid pCIZ2 from the multiple bacteriocin producer *Enterococcus faecium* L50 and genetic characterization of enterocin Q production and immunity. Appl Environ Microbiol 72:6653–6666. https://doi.org/10.1128/AEM.00859-06
49. Shaik S, Pandey H, Thirumalasetti SK, Nakamura N. 2019. Characteristics and functions of the Yip1 domain family (YIPF), multi-span transmembrane proteins mainly localized to the golgi apparatus. Front Cell Dev Biol 7:130. https://doi.org/10.3389/fcell.2019.00130
50. Major D, Flanzbaum L, Lussier L, Davies C, Caldo KMP, Acedo JZ. 2021. Transporter protein-guided genome mining for head-to-tail cyclized bacteriocins. Molecules 26:7218. https://doi.org/10.3390/molecules26237218
51. Lobb B, Tremblay BJM, Moreno-Hagelsieb G, Doxey AC. 2020. An assessment of genome annotation coverage across the bacterial tree of life. Microb Genom 6:e000341. https://doi.org/10.1099/mgen.0.000341
52. Shen W, Le S, Li Y, Hu F. 2016. Seqkit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. PLoS ONE 11:e0163962. https://doi.org/10.1371/journal.pone.0163962
53. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:1–19. https://doi.org/10.1186/1471-2105-5-113
54. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–37. https://doi.org/10.1093/nar/gkr367

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.

55. Tietz JI, Schwalen CJ, Patel PS, Maxson T, Blair PM, Tai HC, Zakai UI, Mitchell DA. 2017. A new genome-mining tool redefines the lasso peptide biosynthetic landscape. Nat Chem Biol 13:470–478. https://doi.org/10.1038/nchembio.2319
56. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2020. GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. Bioinformatics 36:1925–1927. https://doi.org/10.1093/bioinformatics/btz848
57. Gerlt JA, Bouvier JT, Davidson DB, Imker HJ, Sadkhin B, Slater DR, Whalen KL. 2015. Enzyme function initiative-enzyme similarity tool (EFI-EST): a web tool for generating protein sequence similarity networks. Biochim Biophys Acta 1854:1019–1037. https://doi.org/10.1016/j.bbapap.2015.04.015
58. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35:4453–4455. https://doi.org/10.1093/bioinformatics/btz305
59. Xu S, Li L, Luo X, Chen M, Tang W, Zhan L, Dai Z, Lam TT, Guan Y, Yu G. 2022. *Ggtree*: a serialized data object for visualization of a phylogenetic tree and annotation data. iMeta 1:e56. https://doi.org/10.1002/imt2.56
60. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. Nature New Biol 596:583–589. https://doi.org/10.1038/s41586-021-03819-2
61. Schwengers O, Jelonek L, Dieckmann MA, Beyvers S, Blom J, Goesmann A. 2021. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. Microb Genom 7:000685. https://doi.org/10.1099/mgen.0.000685

Downloaded from https://journals.asm.org/journal/jb on 01 September 2025 by 143.239.45.11.