

Phylogenomics of novel clones of *Aeromonas veronii* recovered from a freshwater lake reveals unique biosynthetic gene clusters

Opeyemi U. Lawal,¹ Noah Bryan,^{1,2} Valeria R. Parreira,¹ Rebecca Anderson,¹ Yanhong Chen,¹ Melinda Precious,¹ Lawrence Goodridge¹

AUTHOR AFFILIATIONS See affiliation list on p. 16.

ABSTRACT Aquatic ecosystems serve as crucial reservoirs for pathogens and antimicrobial resistance genes, thus presenting a significant global health risk. Here, we investigated the phylogenomics of *Aeromonas veronii* from Lake Wilcox in Ontario. Among the 11 bacterial isolates, nine were identified as *A. veronii*. Notably, 67% of *A. veronii* isolates were potential human pathogens. Considerable genetic diversity was noted among the *A. veronii* isolates, suggesting the lake as a reservoir for multiple human pathogenic strains. Comparison of the *A. veronii* sequenced with global *A. veronii* genomes highlighted significant genetic diversity and suggests widespread dissemination of strains. All the isolates carried chromosomal genes encoding resistance to β -lactams. Although virulence gene content differed between human and non-human pathogenic strains, type III secretion systems was associated with human pathogenic isolates. The assessment of AMR genes in global isolates showed that β -lactam and tetracycline resistance genes were predominant. Although the machine learning-based pangenome-wide association approach performed did not yield any source-based genes, some genes were enriched in a few isolates from different sources. The *mrkABCD* operon that mediates biofilm formation and genes encoding resistance to colistin, chloramphenicol, trimethoprim, and tetracycline were enriched in animal products, whereas macrolide resistance genes and Inc plasmid-types were linked to the aquatic environment. Novel biosynthetic gene clusters were identified, suggesting that *A. veronii* with varying pathogenic potential could produce unique secondary metabolites. There is a need for continuous tracking of pathogens in aquatic ecosystems to contribute to our understanding of their evolutionary dynamics and the ecological roles of their genetic elements.

IMPORTANCE Lakes and other aquatic ecosystems can harbor harmful bacteria that can make people sick and resist antibiotics, posing a significant global health risk. In this study, we investigated *Aeromonas veronii*, a Gram-negative bacteria found in Lake Wilcox in Ontario. We used various techniques, including whole-genome sequencing (WGS), to analyze the bacteria and found that many of the isolates had the potential to cause human disease. We also discovered significant genetic diversity among the isolates, indicating that the lake may be a reservoir for multiple human pathogenic strains. All isolates carried genes that confer resistance to antibiotics, and some virulence genes were associated with human pathogenic isolates. This study highlights the importance of monitoring aquatic ecosystems for harmful bacteria to better understand their evolution, potential for human pathogenicity, and the ecological roles of their genetic elements. This knowledge can inform strategies for preventing the spread of antibiotic-resistant bacteria and protecting public health.

KEYWORDS *Aeromonas*, genomics, antimicrobial resistance, global population structure, biosynthetic gene cluster, water quality, freshwater lake, public health

Editor Luke R. Iwanowicz, USDA-ARS National Center for Cool and Cold Water Aquaculture, Kearneysville, West Virginia, USA

Address correspondence to Opeyemi U. Lawal, lawal@uoguelph.ca, or Lawrence Goodridge, goodridl@uoguelph.ca.

The authors declare no conflict of interest.

Received 11 May 2024

Accepted 23 September 2024

Published 8 November 2024

Copyright © 2024 Lawal et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

The role of aquatic ecosystems as reservoirs for clinically relevant pathogens and antimicrobial resistance genes (ARG) has recently gained attention as the importance of assessing the quality of these ecosystem is paramount in public health (1, 2). Freshwater bodies like lakes and rivers used for recreational purposes can significantly impact the health of communities (3, 4). Poor water quality in these settings poses a substantial risk for the transmission of various waterborne diseases, including pathogenic viruses, protozoa, and bacteria including *Aeromonas* species that thrive in such contaminated water (2–4).

Aeromonas species are Gram-negative, facultative anaerobic rods, found in various aquatic environments (5–7), and known for their ability to survive in diverse environments, ranging from freshwater to the intestinal tracts of animals (5, 7). Although some *Aeromonas* species including *Aeromonas salmonicida*, *Aeromonas hydrophila*, and *Aeromonas veronii* are known fish pathogens, *A. veronii* is one of the four species that are considered as potential human pathogens (8–10). *A. veronii* is an emerging human pathogen causing a wide range of diseases in human and animals including gastroenteritis, respiratory and skin infections, and septicemia (9–11). In addition, *A. veronii* is increasingly being recognized as a significant concern to food safety due to its frequent presence in different types of food, particularly in minimally processed ready-to-eat seafood (12, 13). Of note, the frequent and global occurrence of highly virulent strains of *A. veronii* has been detected in food samples such as meat, milk, catfish, and fish in countries including Brazil (13), Egypt (14), India (15), Israel (11), and the USA (16, 17), among others. The adaptability of *A. veronii* to various conditions poses a challenge for water quality management, especially in environments with high anthropogenic activities, where the bacterium can be a potential source of infection (5, 18).

The mechanisms of pathogenicity of *A. veronii* involve the production of various toxins and virulence factors that contribute to its ability to infect host cells and cause disease (5, 18). A significant concern with *A. veronii* is its capacity for antimicrobial resistance (AMR) (7, 19, 20). The presence of antimicrobial-resistant strains in aquatic environments is a public health concern, as it not only affects the treatment of *Aeromonas*-related infections but also represents a potential reservoir for the spread of resistance genes to other pathogenic bacteria (7, 19, 20). Studies on the population structure of *A. veronii* have described genetic diversity driven by its adaptability to various environmental conditions. These factors could drive variability in strains regarding pathogenicity and resistance to environmental stresses in this bacterium, with practical implications for public health and water management (5, 8, 18).

In recent years, advancements in sequencing technologies have greatly enhanced the genomic surveillance of known and emerging pathogens, such as *A. veronii*, across different environmental matrices (5). Despite these technological advancements, little importance has been given to *A. veronii*, especially in terms of its presence in freshwater, its impact on water quality, and its role in the dissemination of AMR in both the environment and the food chain. Understanding the genomic surveillance and population structure of this bacterium is crucial for developing effective infection treatment strategies and ensuring public health safety.

We have previously reported the detection of clinically relevant pathogens in Lake Wilcox, including novel strains of *Bacillus anthracis* (21) and *Vibrio cholerae* (22) isolated at different time points. In this study, we employed a combination of culture-based detection and whole-genome sequencing (WGS) to assess the presence of *A. veronii*, its extensive genomic fingerprint, its population structure, and the genomic characterization of stress response genes in Lake Wilcox. The genetic relatedness of *A. veronii* isolates was assessed by comparing them with previously sequenced strains in public databases using a comparative genomic approach.

MATERIALS AND METHODS

Description of sampling site

Lake Wilcox is a small kettle lake located in Richmond Hill in Ontario (43°56′56.69″ N, 79°26′9.45″ W). Historically, the lake is used for recreational purposes by the surrounding community and tourists. Despite being impacted by feces of surrounding wildlife, recreational activities have continued, and users have reported skin rashes and gastrointestinal symptoms after recreational activities (<https://projectboard.world/ysc/project/the-phage-takes-centre-stage-for-water-quality-testing>).

Sample collection and processing

Freshwater samples were obtained from Lake Wilcox in the summer of 2022 and fall of 2023. Water samples were kept at 4°C and analyzed within 48 h of collection. Samples were processed as described by Bryan et al. (23). Briefly, 1 mL of samples was serially diluted in 9 mL of lambda buffer (modified saline-magnesium buffer without gelatin) and plated onto tryptic soy agar (TSA). Following incubation for 24 h at 37°C, plates were analyzed for bacterial colonies. Distinct colonies of differing morphologies were sub-cultured onto TSA to obtain pure culture. The isolated colonies were Gram stained, and taxonomic identification was performed using VITEK (bioMérieux, Inc, Canada).

Genomic DNA extraction and whole-genome sequencing

Genomic DNA from isolated colonies was extracted using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNA libraries were prepared using the Illumina DNA prep tagmentation kit (#20018704) and IDT for Illumina DNA/RNA UD indexes (#20027213) following the manufacturer's instructions. Paired-end (2 × 150 bp) sequencing was performed using the high-output flow cell on the Illumina MiniSeq instrument as described previously (23, 24).

Genome assembly and annotation

Raw paired-end reads were quality filtered using FastQC v0.11.9 (<https://github.com/s-andrews/FastQC>) and trimmed using Trimmomatic v0.39 (25). High-quality reads with a Phred quality score above 20 were assembled *de novo* using the Skesa v2.4.0 pipeline (26). Assembly quality and genome completeness were assessed using QUAST v5.2 (27) and BUSCO (28), respectively. Taxonomic classification was performed using pubMLST and rMLST (29), as well as k-mer-based species taxonomic classification with the Kraken2 database (30). The average nucleotide identity (ANI) analysis was performed using fastANI (31). Genome annotation was performed using Prokka v1.14.6 (32).

Gene content analysis

The pathogenicity of the isolates was determined using the PathogenFinder tool (33), a machine learning-based model that compares the whole-proteome sequences to a database composed of protein families associated with either pathogenic or non-pathogenic organisms in humans and returns with a pathogenicity score. The antimicrobial resistance gene profile of all the isolates was determined using AMRFinder Plus v3.10.45 (34) and CARD (35) databases, whereas the virulence genes were identified using the updated virulence factor database (VFDB) (36, 37), which contains all known genes reported to be associated with bacterial virulence. Genes with a threshold of >70% coverage and >90% nucleotide identity were considered to be present. To define the mobile genetic elements (MGE) of the collection, the draft genomes were screened for plasmids using MOB-suite v3.1.6 (38) with default settings. Prophage regions were detected using PHASTEST (39) and PhaBox (40). The completeness [CheckV (41)] and classification [PhaGCN (42)], as well as lifestyle [BACPHLIP (43)] and the host [CHERRY (44)] of the detected prophage sequences were determined. Intact phages were screened for tailspike proteins (TSPs) using TSPDB that contains 8,105 TSPs (45). Lawal

and Goodridge (2024 Preprint). Biosynthetic gene clusters were assessed using the antiSMASH v6 pipeline (46).

Pangenome and phylogenetic analysis

To construct the phylogeny, the pangenome of all publicly available *A. veronii* genomes was generated from the annotated genomes using Roary v3.13.0 (47), and single nucleotide polymorphisms (SNPs) within the core-genome alignment were extracted using SNP-sites v2.5.1 (48) indicating GCA_008693705.1. The concatenated core-genome-based SNPs were used to construct a phylogenetic tree using FastTree (49). The general time reversible model was performed with 1,000 bootstrap resampling for node support. Except as otherwise stated, all bioinformatics tools were executed using the default settings. Genomic features that are exclusive and/or enriched in different source groups with ≥ 20 genomes were determined using the pan-genome-wide association studies (pan-GWAS) approach with Scoary v1.6.16 (50) as previously described (51, 52). Genes with a Benjamini–Hochberg P value of <0.05 and an odds ratio of >1 , with no duplicated function in the pangenome, were considered. Moreover, machine learning approach with Fisher's exact test was used as a second approach to identify genes significantly associated with different source categories. Specifically, the differences in antimicrobial resistance, virulence genes, and mobile genetic elements across various sources were assessed with an analysis of variance (ANOVA). The data set was normalized to account for variations in sample sizes, and the statistical significance of observed differences was determined with a P value <0.05 .

RESULTS

Bacterial identification and sequencing

Bacterial species were isolated from the freshwater samples recovered from a freshwater lake over a period of 1 year using the spread agar plate method. Overall, 11 distinct colonies with different morphologies were selected and further characterized using morphological, biochemical-based, and whole-genome sequencing. Taxonomic identification using the VITEK Compact system identified the colonies as *Aeromonas sobria* ($n = 8$), *A. hydrophila/punctata* ($n = 2$), and one isolate with an inconclusive result (Table 1). Sequencing of the 11 isolates yielded 1,024,248–2,725,402 paired-ended reads per isolate (Table 1). Isolates were identified as *A. veronii* ($n = 9$), *Aeromonas caviae* ($n = 1$), and *Aeromonas allosaccharophila* ($n = 1$) (Table 1). The ANI analysis with fastANI (31) using the species demarcation threshold of 95% (53) showed that the nine *A. veronii* strains had $>96\%$ ANI when *A. veronii* GCF_000820225.1 strain was used as a reference, *A. caviae* strain NB-180 had 97.9% ANI with *A. caviae* GCF_000819785.1, whereas *A. allosaccharophila* had 96.21% ANI with the reference strain *A. allosaccharophila* GCF_000819685.1. The draft genomes of *Aeromonas* species yielded between 28 and 113 contigs, with a G+C content of 58%–59%, except for *A. caviae* that had a higher G+C content of 61.26%, a value that was comparable to the reference strain *A. caviae* GCF_000819785.1. The genome size was comparable between the three *Aeromonas* species identified and ranged between 4,390,436 and 4,690,056 bp, with $>50\times$ genome coverage (Table 1).

Prediction of human pathogenicity of *A. veronii* sequenced

Considering that *Aeromonas* species are commonly associated with diseases in fish, we evaluated the potential of these isolates to be pathogenic to humans using Pathogen-Finder tool (33). Six out of the nine *A. veronii* isolates sequenced in this study had a pathogenicity score greater than 0.5, suggesting that they may be pathogenic to humans. Other isolates, including *A. caviae* and *A. allosaccharophila*, were predicted to be non-human pathogens (Table 1).

TABLE 1 Summary of sequence metrics of *Aeromonas* isolates recovered from a freshwater lake

Isolate ID	Isolation period	Coverage	# Contigs	Genome size	% GC	N50	CDS	misc_RNA	rRNA	tRNA	tmRNA	Repeat region	Vitek ID	WGS-based ID	ST	Assembly accession	SRA accession
NB-178	Fall 2023	142.68	44	45,73,701	58.62	2,22,570	4,073	46	9	84	1	0	A. sobria	A. veronii	2530	GCA_039652535.1	SRR28980392
NB-180	Fall 2023	134.57	50	45,00,871	61.26	1,73,045	4,069	48	3	84	1	0	A. hydrophila/punctata	A. caviae	2531	GCA_039652415.1	SRR28980391
NB-181	Fall 2023	53.34	90	46,01,063	58.61	1,61,830	4,101	47	9	67	1	0	Inconclusive	A. veronii	2530	GCA_039652515.1	SRR28980390
NB-185	Fall 2023	142.75	48	46,38,531	58.75	1,76,021	4,225	50	12	76	1	0	A. sobria	A. veronii	2532	GCA_039652495.1	SRR28980389
NB-187	Fall 2023	138.89	30	43,90,436	59.02	3,00,356	3,957	51	7	2	1	2	A. sobria	A. veronii	2533	GCA_039652475.1	SRR28980388
NB-188	Fall 2023	135.78	39	46,90,056	58.54	2,38,966	4,250	51	6	78	1	0	A. hydrophila/punctata	A. veronii	2534	GCA_039652455.1	SRR28980387
NB-2	Summer 2022	73.11	46	44,65,557	59.01	2,33,461	4,074	47	7	72	1	1	A. sobria	A. veronii	2535	GCA_026571335.1	SRR22013455
NB-3	Summer 2022	88.8	28	45,93,492	58.78	4,39,357	4,135	51	6	80	1	0	A. sobria	A. veronii	2536	GCA_026571285.1	SRR22013454
NB-4	Summer 2022	97.22	44	44,66,772	59.01	1,79,674	4,071	47	7	76	1	0	A. sobria	A. veronii	2535	GCA_026571345.1	SRR22013453
NB-6	Summer 2022	94.85	53	44,83,134	58.83	2,11,391	4,052	44	12	79	1	0	A. sobria	A. veronii	2537	GCA_026571275.1	SRR22013452
NB-7	Summer 2022	75.89	113	46,08,243	58.93	97,905	4,179	43	3	83	1	1	A. sobria	A. allosaccharo-phila	2538	GCA_035798095.1	SRR22013451

Population structure of *A. veronii* isolated from freshwater

To assess the genetic relatedness among isolates sequenced in this study, a combination of conventional multilocus sequence typing (MLST) and whole-genome-based phylogeny was employed. The *Aeromonas* MLST schema was used to determine the sequence types (STs) of all isolates. Of note, 41 novel alleles were identified among the 11 *Aeromonas* isolates and yielded nine unique allele profiles that were submitted together with the allele sequences and assigned to nine new STs (ST2530–ST2538) (Table 1; Table S1). Two STs (ST2530 and ST2535) contained two isolates each, whereas others were singletons suggesting the uniqueness of the isolates under study and high genetic diversity in the population. The core-genome SNP-based phylogeny of the nine *A. veronii* sequenced was constructed using the complete genome of *A. veronii* GCA_008693705.1 as a reference. *A. caviae* and *A. allosaccharophila* were used as outgroups to re-root the tree. Isolates were grouped into two main clusters irrespective of the period of isolation (Fig. 1). Isolates were distantly related by SNPs with ≥ 100 SNPs difference (Table S2) except for a pair of isolates from different timepoints (NB-2/NB-4, summer 2022, and NB-178/NB-181, fall 2023) that were highly related differing only by 9 and 11 SNPs, respectively (Fig. 1). Of note, the SNP-based clustering observed was similar to the MLST-based population structure suggesting a good concordance between these methods for typing *A. veronii*. Overall, the high genetic diversity observed in this study suggests that the freshwater lake could serve as a reservoir for multiple strains of *A. veronii* that are pathogenic to humans.

Global population structure of *A. veronii*

To assess the genetic relatedness of the sequenced isolates with global *A. veronii*, genomes and the associated metadata of 406 *A. veronii* deposited in the RefSeq database (accessed on 13 July 2024) were downloaded and re-annotated (see Methods). The 406 genomes were recovered from 30 different countries located in six continents between 1988 to 2023 from eight different sources including human, animal, aquatic environment, freshwater fish, among others (Table S3). The pangenome size of the 406 *A. veronii* genomes together with the sequenced isolates ($n = 9$) yielded 54,993 genes. A total of 2,126 core genes, defined as genes present in $\geq 95\%$ of the genomes in the collection, were identified, whereas the shell and cloud genes totaled 2,630 and 50,237, respectively. The core-genome SNP-based phylogeny was constructed using *A.*

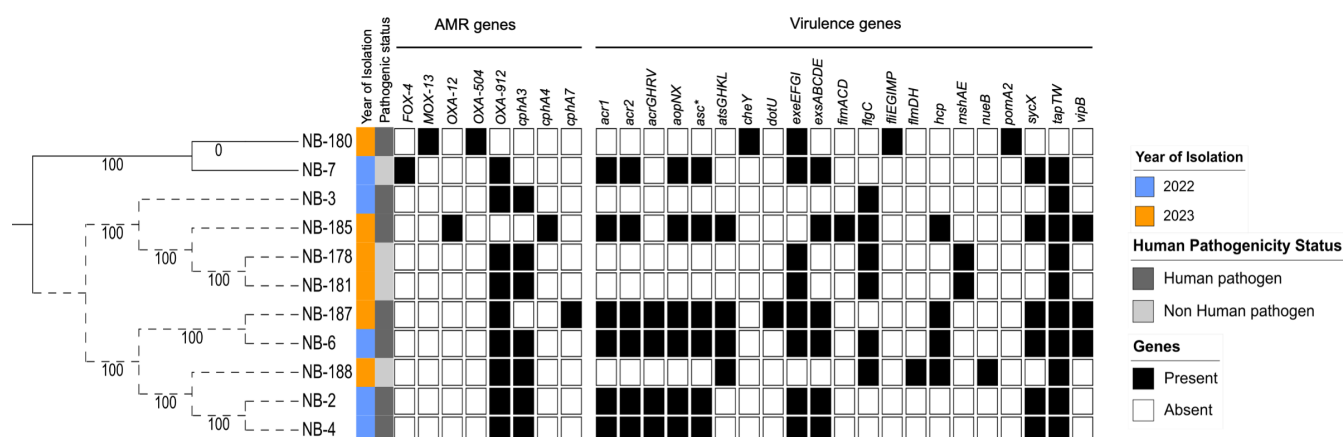


FIG 1 Maximum likelihood tree of *A. veronii* recovered from freshwater lake in Ontario. Each node represents a strain. To construct the phylogeny, pangenome was generated from the annotated genomes using Roary v3.13.0. SNPs within the core-genome alignment were extracted with SNP-sites v2.5.1 using GCA_008693705.1 as a reference. The concatenated core-genome-based SNPs were used to construct a phylogenetic tree using FastTree. The general time reversible model was performed with 1,000 bootstrap resampling for node support. The dotted lines depict *A. veronii*. NB-7 (*A. allosaccharophila*) and NB-180 (*A. caviae*) were outgroups used to re-root the tree. The draft genomes were screened for genes encoding antimicrobial resistance and virulence using CARD and VFDB databases, respectively. The tree was visualized using iTOL (<https://itol.embl.de>).

allosaccharophila as outgroup to root the phylogenetic tree. The sequenced *A. veronii* isolates compared with global *A. veronii* species showed high genetic diversity, which facilitated the clustering of the isolates into distinct clades (Fig. 2). Isolates sequenced in this study were clustered into distinct subclades, suggesting that they are distantly related to other global isolates. However, isolates from Turkey and Greece recovered in fish from different time points (2009, 2015, and 2016) were clustered together, a phenomenon that could suggest dissemination of *A. veronii* strains. Strain NB-188 belonged to the same subcluster as an isolate recovered from a leech in the USA in 2002. Although global *A. veronii* species were distantly related, we could still observe a mixture of isolates from different countries and sources within some clusters. Overall, *A. veronii* from different environments may have genetic signatures unique to pathogenic strains of this bacterium. This could also be important to determine or predict the source of isolates found in any matrix. However, the pan-GWAS approach did not yield any gene with strong association with source of isolation (Fig. 3).

Stress response genes among the sequenced isolates and in global *A. veronii*

Genes encoding resistance to β -lactams were detected in all the isolates sequenced in this study. Different alleles of *cphA* (*cphA3*, *cphA4*, and *cphA7*) gene, which belonged to the subclass B2 metallo- β -lactamase that encodes resistance to carbapenem antibiotics were detected in all *A. veronii* isolates and in *A. allosaccharophila*. OXA-912 that encodes resistance to penams, cephalosporins and carbapenems, and *cphA3* genes were predominant in the collection (Fig. 1). The rescreening of global *A. veronii* genomes for AMR genes revealed that these β -lactam resistance genes were ubiquitous in this bacterium

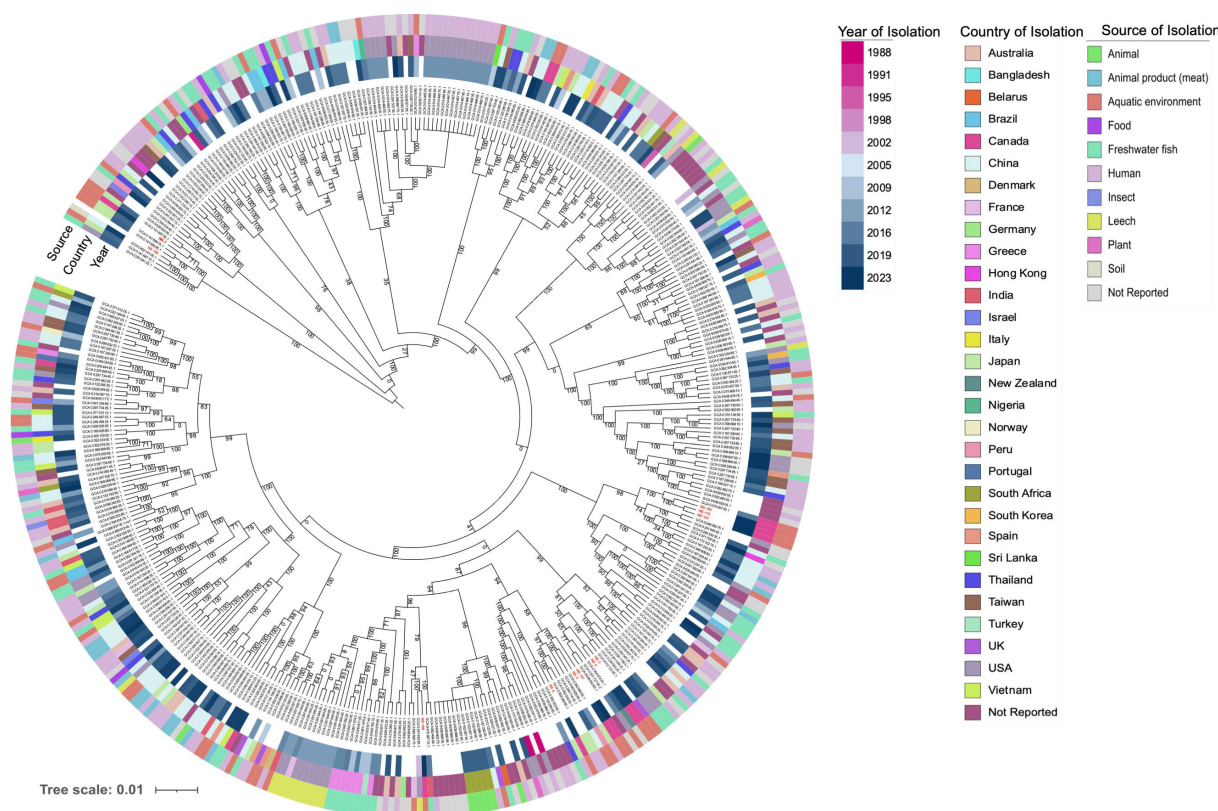


FIG 2 Core-genome-based maximum-likelihood tree of global *A. veronii* from different sources. Each node represents a strain. To construct the phylogeny, pangenome was generated from the annotated genomes using Roary v3.13.0. SNPs within the core-genome alignment were extracted using SNP-sites v2.5.1 using GCA_008693705.1 as a reference. The concatenated core-genome-based SNPs were used to construct a phylogenetic tree using FastTree. The general time reversible model was performed with 1,000 bootstrap resampling for node support. *A. allosaccharophila* (NB-7) and *A. caviae* (NB-180) were outgroups used for re-rooting the tree. The figure was generated using iTOL (<https://itol.embl.de>). *Aeromonas* isolates recovered in this study were labeled in red.

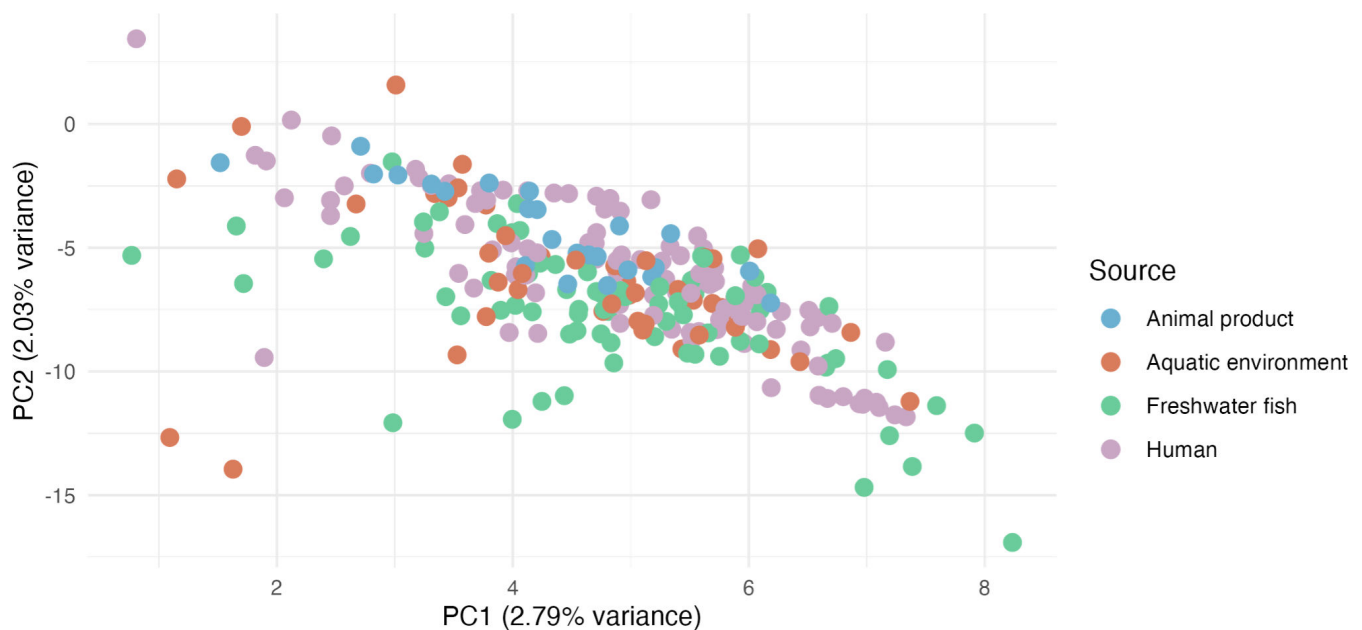


FIG 3 Assessment of the global *A. veronii* pangenome reveals limited clustering based on source in gene content variation. The principal component analysis (PCA) plot was generated from the pangenome (gene presence/absence) data, produced by Roary, and the associated source information. Each point on the plot represents an individual genome, positioned according to the principal components that account for the most variance in gene presence/absence data. Points are colored based on the source of isolation, with the spread of nodes indicating differences in gene content across genomes from various sources. The lack of distinct source-based clusters suggests that the variation in gene content among the genomes is not strongly correlated with their source of isolation.

(Fig. S1; Table S4). Genes encoding resistance to nine other classes of antibiotics were detected in the global collection, with tetracycline resistance genes being the second most prevalent after β -lactam resistance genes, found in 35% of the collection. The prevalence of resistance genes for the other classes (including aminoglycosides, chloramphenicol, colistin, macrolides, quaternary ammonium compounds, quinolones, sulfonamides, and trimethoprim) ranged from 7% to 18% (Fig. S1; Table S4). Of note, some AMR genes were enriched in a few isolates from different sources. Genes encoding resistance to colistin, chloramphenicol, trimethoprim, and tetracycline were enriched in isolates from animal products. In contrast, macrolide resistance genes show a higher prevalence in isolates from the aquatic environment. Freshwater fish and human sources exhibit a more balanced distribution of these resistance genes, with no single class being significantly dominant (Fig. 4A). Overall, these results suggest that specific environmental and/or anthropogenic factors may influence the prevalence of AMR genes in *A. veronii* from different sources.

The detection of virulence genes using a gene homology approach (see Methods) and a curated virulence gene database (VFDB [36, 37]) detected 8 to 48 virulence genes in each *A. veronii* isolate sequenced in this study. Relative to isolates predicted as human pathogens that contained 41–48 virulence genes (except NB-3), all the *A. veronii* predicted as non-human pathogens carried less virulence genes (≤ 9 virulence genes). Human and non-human pathogenic strains of *A. veronii* differed in terms of their virulence gene content. Although the flagellar and type IV pili-associated genes involved in biofilm formation (18, 37) were detected in all isolates, type III secretion system (T3SS)-associated genes were detected only in the isolates predicted as human pathogens (Fig. 1). The assessment of virulence genes in the global collection of *A. veronii* revealed a slight differential distribution of T3SS-associated genes and other virulence factors across isolates from different sources with freshwater fish isolates carrying more virulence genes compared to those from other sources (Fig. 4B). Isolates from animal products carried the least number of virulence genes except for the *nueB* gene. Also, the *mrkABCD* operon, which encodes type III fimbriae and involved in

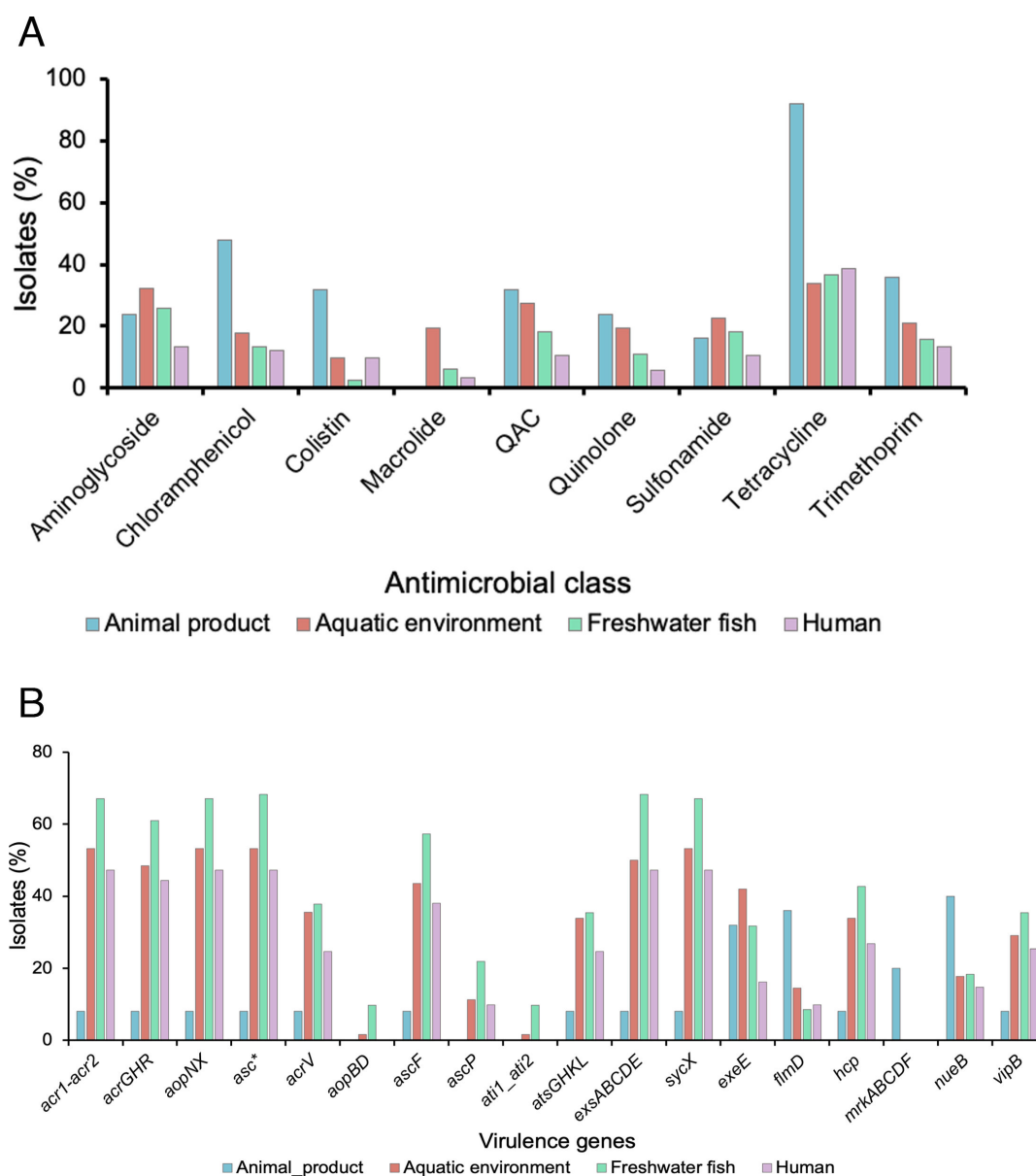


FIG 4 Distribution of genes encoding resistance to different classes of antimicrobials (A) and virulence (B) across various sources in a global collection of *A. veronii* genomes

biofilm formation on biotic and abiotic surfaces in *Klebsiella* and *Citrobacter* (36, 37), was exclusively found in isolates recovered from animal products (Fig. 4B). This overall pattern suggests a somewhat source-dependent variation in virulence gene content, highlighting the potential for distinct pathogenic profiles in a few *A. veronii* isolates from different environments.

Characterization of mobile genetic elements among the sequenced isolates and in global *A. veronii*

Plasmids were not detected in any of the *A. veronii* isolates sequenced in this study. However, rescreening of global *A. veronii* genomes identified 25 different known plasmid types in 32% ($n = 130/406$) of the genomes, with the IncU plasmid being the most predominant type in the collection (Table S5). Although isolates recovered from the aquatic environment and freshwater fish had the highest diversity of plasmid types, the

Inc plasmid types (IncC, IncP, IncQ1, IncQ2, IncU) were particularly enriched in isolates from the aquatic environment (Fig. 5A). Of note, 16% ($n = 67/406$) of the genomes carried plasmids with genes encoding resistance to at least one antimicrobial class, the great majority ($n = 39/67$) of which were non-mobilizable, suggesting that they could be cryptic plasmids. The remaining few that carried mobilizable/conjugative plasmids were predominantly recovered from the aquatic environment (Fig. 5B). Of particular interest was a previously described IncC plasmid (54) that carried 15 AMR genes that encode resistance to nine different classes of antimicrobials including aminoglycoside, chloramphenicol, macrolides, sulfonamide, among others (Fig. 5B). The analysis of the other genes in this genetic element showed that it contains phage genes. A further screening of the plasmid through the phage detection pipeline revealed an intact phage region within this plasmid, suggesting that this genetic element is indeed a phage-like plasmid (55, 56).

The detection and characterization of phage regions in the isolates sequenced in this study yielded 13 unique intact phages, among which four were predicted to be virulent phages (43). The completeness of all intact phage sequences was determined to be between 50% and 100% by CheckV (41). The phages were classified by PhaGCN (42) as Peduoviridae ($n = 10$), Chaseviridae ($n = 1$), and two others unidentified according to the International Committee on Taxonomy of Viruses (ICTV) classification (57). In addition to *A. veronii* being predicted as host of the phages, other species of *Aeromonas* (*A. australiensis*, *A. diversa*, *Aeromonas* sp.) and *Serratia marcescens* could also serve as their hosts as determined by CHERRY (40, 44), suggesting that these phages could infect multiple hosts (Table 2). Of note, the two pairs of isolates (NB-2/NB-4 and NB-178/NB-181) that were highly genetically related by SNP had the same phage content. No antibiotic resistance, toxin, or related genes were detected in the intact phages. All the

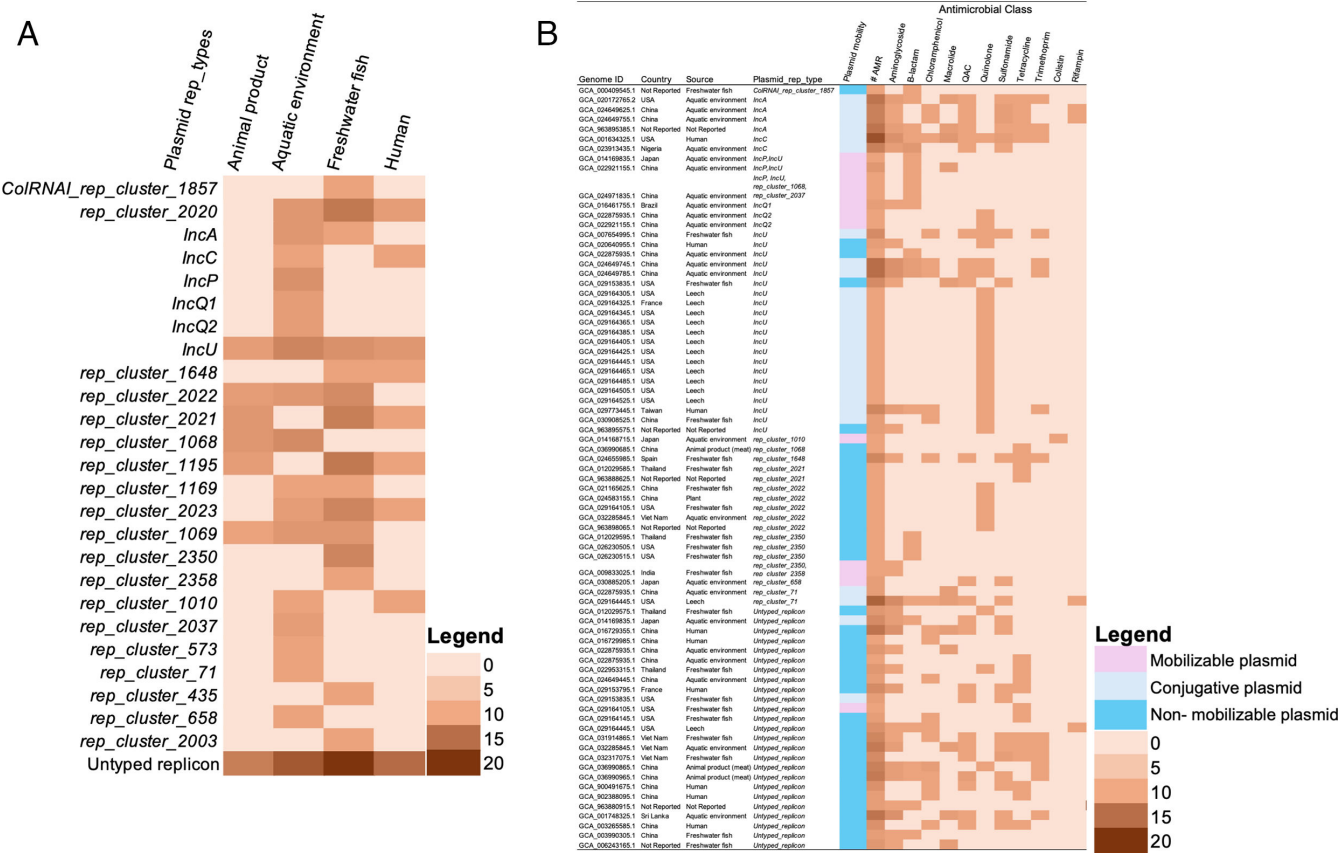


FIG 5 Source-dependent distribution of plasmid types (A) and plasmids encoding resistance to multiple classes of antimicrobials (B) in a global collection of *A. veronii*. The color shades reflect the relative abundance of each plasmid type (A) or the number of antimicrobial resistance genes (B).

intact phages detected were screened for TSPs using TSPDB that contains 8,105 TSPs (49), but none was found.

Biosynthetic gene cluster profile in *A. veronii*

The 53 biosynthetic gene clusters (BGCs) identified in all sequenced *Aeromonas* species were categorized into eight BGC families using sequence similarity network analysis with BiGSCAPE (58). The ribosomally synthesized and post-translationally modified peptides (RiPPs) were the most predominant BGC class, consisting of three gene families. In contrast, the non-ribosomal polyketide synthase (NRPS) included only one gene family. The remaining gene families were classified as “others” and included homoserine lactone ($n = 2$) and aryl polyene ($n = 2$). The three RiPPs detected were unique and conserved within the collection but exhibited low similarity scores to previously described BGCs. For example, RiPP-1 (Fig. 6A), comprising 11 open reading frames (ORFs), had a similarity score of 0.17 to angustmycin A/B/C (BGC0002621) described in *Streptomyces angustmyceticus* (accession MZ151497.1) (59). Meanwhile, RiPP-2 and RiPP-3 (Fig. 6B and C), consisting of nine and seven ORFs, respectively, had similarity scores of ≤ 0.08 to pseudopyronine A/B (BGC0001285) described in *Pseudomonas putida* (accession KT373879.1) (60). Notably, RiPP-3 was also detected in *A. allosaccharophila* (NB-7), indicating that this BGC is not exclusive to *A. veronii* (Fig. 6C). The identified NRPS had the highest similarity score of 0.9 to enterobactin (BGC0000343) previously described in *Pseudomonas* sp. J465 (accession GQ370384.1) (61). This BGC was conserved in the *A. veronii* sequenced (Fig. 6D). Further analysis of global *A. veronii* genomes confirmed that this BGC was conserved not only in this collection but also in all publicly available *A. veronii* genomes. A BGC encoding homoserine lactone, predominant in *A. veronii* ($n = 6/9$), was also detected in *A. allosaccharophila*. This BGC had a low similarity score (0.14) to thioguanine (BGC0001992) in *Erwinia amylovora* CFBP1430 (accession number: NC_013971.1) (62). Notably, a pair of *A. veronii* strains—one pathogenic (NB-6/NB-187) and one non-pathogenic to humans (NB-178/NB-181)—carried unique BGCs encoding aryl polyene (Fig. 7A and B). The pathogenic pair consisted of 17 ORFs with a similarity score of 0.44 to aryl polyene (BGC0002008) described in *Xenorhabdus doucetiae* (accession NZ_F0704550.1) (63), whereas the non-pathogenic pair contained 37 ORFs with a similarity score of 0.26 to bacilysin (BGC0000888) described in *Bacillus* sp. CS93 (accession number: GQ889493.1) (64). Overall, *A. veronii* harbored putative unique BGCs that exhibited low similarity scores to previously described compounds.

DISCUSSION

Aquatic ecosystems are continually impacted by anthropogenic activities, making the microbial quality and safety of these water bodies, especially those used for recreational activities, paramount for public health (1). In this study, we assessed the presence of *A. veronii* in a recreational lake and determined the extensive genomic features of the isolates regarding their population structure and the genomic characterization of stress response genes, mobile genetic elements, and other gene content such as biosynthetic gene clusters that confer uniqueness to different *A. veronii* strains. We inferred the global population structure of *A. veronii* by assessing the genetic relatedness of the isolates sequenced with previously sequenced strains in public databases.

In the past decade, WGS has become the gold standard method for species identification, complementing existing biochemical-based methods (65). In this study, the WGS-based approach identified isolates as *A. veronii*, *A. caviae*, and *A. allosaccharophila*, whereas the biochemical identification system misidentified all isolates as either *A. sobria* or *A. hydrophila*. Misidentification of species of environmental bacteria by biochemical approaches is not uncommon (65). Studies comparing biochemical-based bacterial species identification systems to WGS have shown that species misidentification can vary by species and is common in specific bacteria, including *Pseudomonas fluorescens*, *P. putida* (65), and *Enterococcus faecalis* (66).

TABLE 2 Features of intact phages detected in *A. veronii* sequenced in the study

Isolate ID	Intact phage region	Length (bp)	Phage lifestyle	Phage family	Host	Gene count	Viral genes	Host genes	% Phage completeness
NB-178	NB-178_00010_193626_97002-124946	27,945	Temperate	Peduvoviridae	<i>S. marcescens</i>	34	20	4	59.34
NB-178	NB-178_00010_193626_138510-165106	26,597	Virulent	Peduvoviridae	<i>A. diversa</i>	29	21	2	64.38
NB-180	NB-180p_1:101319-141612	40,294	Virulent	Unknown	<i>A. veronii</i>	39	24	2	67.14
NB-180	NB-180p_1:207148-238244	31,097	Temperate	Peduvoviridae	<i>A. australiensis</i>	41	17	1	84.82
NB-185	NB-185p_1:196668-255973	59,306	Virulent	Chaseviridae	<i>Aeromonas</i> sp. <i>DNP9</i>	82	34	2	97.53
NB-185	NB-185p_14:30853-63241	32,389	Temperate	Peduvoviridae	<i>A. veronii</i>	45	26	0	95.3
NB-185	NB-185p_3:133090-168585	35,496	Virulent	Unknown	<i>A. veronii</i>	56	29	0	100
NB-188	NB-188_00002_602127_302137-337616	35,480	Virulent	Peduvoviridae	<i>A. australiensis</i>	45	18	0	96.81
NB-188	NB-188_00010_161842_8413-46184	37,772	Temperate	Peduvoviridae	<i>A. veronii</i>	56	30	1	100
NB-2	NB-2_00001_820342_705754-750131	44,378	Temperate	Peduvoviridae	<i>A. veronii</i>	70	35	5	92.49
NB-2	NB-2_00009_103898-141465	37,568	Temperate	Peduvoviridae	<i>Aeromonas</i> sp. <i>L_1B5_3</i>	56	33	1	100
NB-3	NB-3_00002_655010_341482-372468	30,987	Temperate	Peduvoviridae	<i>A. australiensis</i>	41	18	0	84.52
NB-7	NB-7_00020_76813_28309-66369	38,061	Temperate	Peduvoviridae	<i>A. australiensis</i>	49	24	1	100

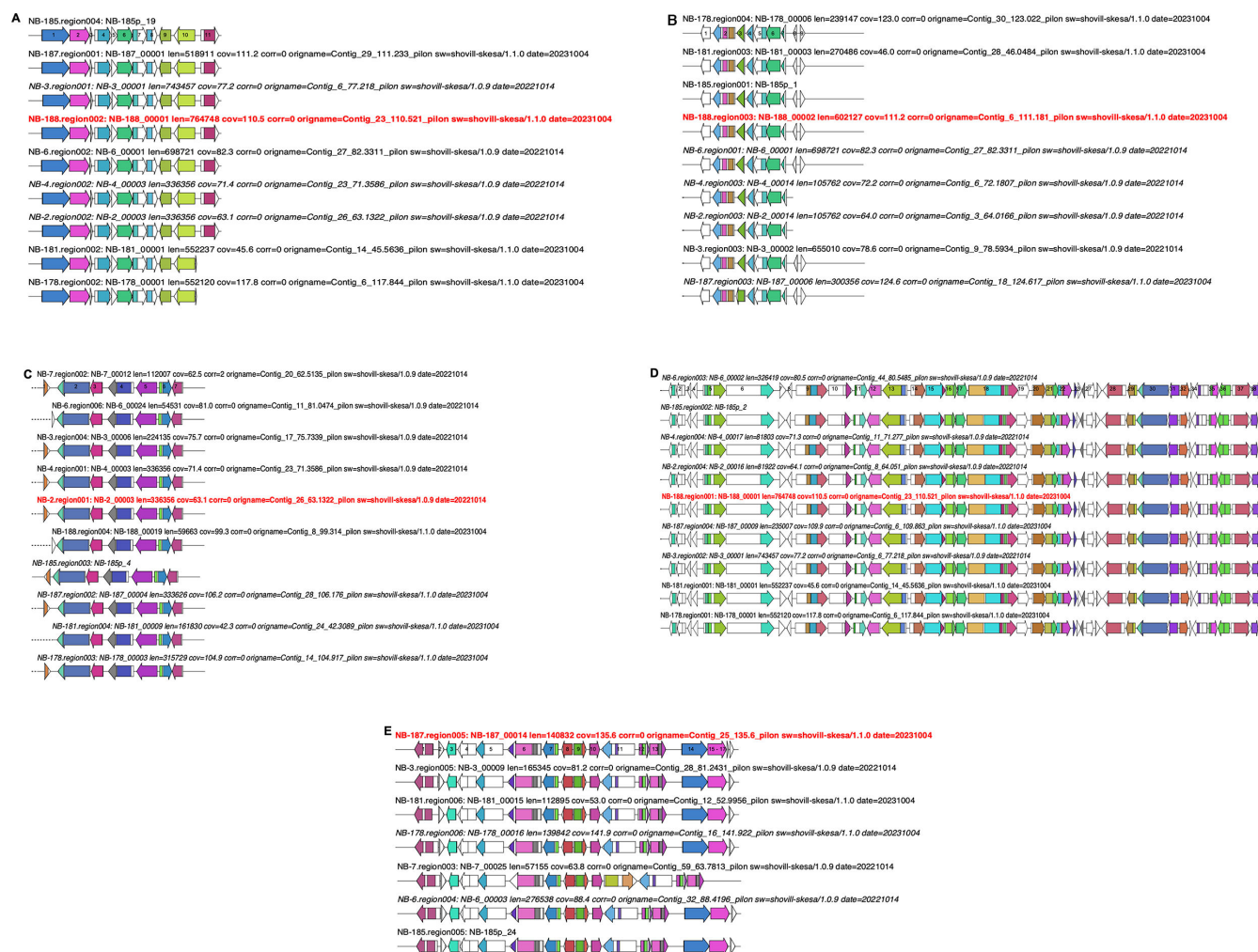


FIG 6 Novel biosynthetic gene clusters identified in *A. veronii* sequenced. (A) RiPP BGC with a low similarity score of 0.17 to angustmycin A/B/C in *S. angustmyceticus* (accession [MZ151497.1](#)). (B) RiPP gene cluster with a low similarity score of 0.04 pseudopyronine A/B found in *P. putida* (accession [KT373879.1](#)). (C) RiPP gene cluster detected in *A. veronii* and *A. allosaccharophila* with a low similarity score of 0.08 to pseudopyronine A/B found in *P. putida* (accession [KT373879.1](#)). (D) NRPS BGC with a 0.9 similarity to enterobactin in *Pseudomonas* sp. J465 (accession [GQ370384.1](#)). (E) Homoserine lactone detected in *A. veronii* and in *A. allosaccharophila* with a low similarity score of 0.14 to thioguanine found in *E. amylovora* CFBP1430 (accession number: [NC_013971.1](#)).

A. veronii strains sequenced exhibited different pathogenic potentials, with the majority (67%, $n = 6/9$) predicted to be pathogenic to humans and possessing a similar virulence determinant profile, including T3SS. T3SSs are crucial virulence mechanisms that allow bacteria to inject effector proteins directly into the host cell cytoplasm. The activity of T3SSs closely correlates with infection progression and outcome in various infection models, and its presence is considered a general indicator of virulence in *A. veronii* (67–69). The detection of human pathogenic *A. veronii* in this study, along with other clinically relevant pathogens such as *B. anthracis* (21) and *V. cholerae* (22) in this recreational lake from previous studies, emphasizes the crucial role aquatic ecosystems play in disseminating pathogens. The recreational use of this water could pose a continuous risk to public health, serving as a reservoir and facilitating the transmission of waterborne diseases. This also underscores the significance of monitoring aquatic environments as reservoirs for pathogenic bacteria.

There was a high genetic diversity among the nine *A. veronii* isolates sequenced, including the identification of novel sequence types and alleles. Although some strains were indistinguishable by SNPs, others were genetically distant. This finding could imply that different *A. veronii* strains may have been introduced into the lake multiple times

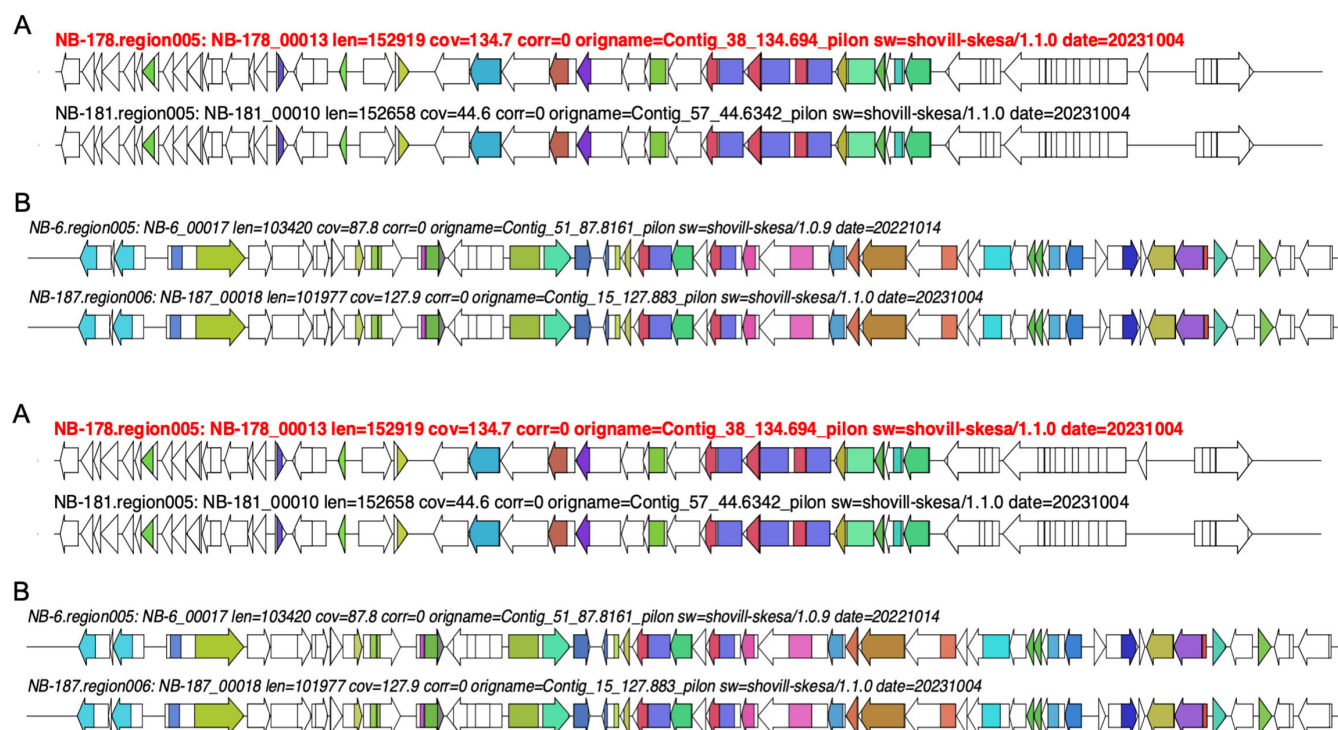


FIG 7 Biosynthetic gene clusters identified in *A. veronii* predicted to be human or non-human pathogens. (A) Aryl polyene gene cluster detected in *A. veronii* isolates predicted to be a non-human pathogen with a similarity score of 0.26 to bacilysin found in *Bacillus* sp. C593 (accession number: [GQ889493.1](#)). (B) Aryl polyene gene cluster detected in *A. veronii* isolates predicted to be a human pathogen with a similarity score of 0.44 to aryl polyene found in *X. doucetiae* (accession number: [NZ_F0704550.1](#)).

from various sources such as resident freshwater fish, domestic animals, and environmental samples (5). The integration of genomic data from the *A. veronii* isolates sequenced with global strains revealed that isolates from single sites formed smaller groups within the phylogeny. Interestingly, one isolate from this study (NB-188) was nested with an isolate recovered from leech in a crop field in the USA. A previous study assessing the core-genome-based phylogenetic analysis of *A. veronii* genomes deposited in NCBI from 18 countries revealed a high genetic diversity (5). The admixture of *A. veronii* strains from different sources was observed, suggesting a lack of source- and timepoint-based clustering in the *A. veronii* population. However, strains from a single site tend to form small groups within the phylogenetic clusters. These observations concur with our findings. The genetic diversity observed in *A. veronii* reinforces the importance of continuous genomic surveillance to monitor the emergence and spread of virulent and/or resistant strains.

The AMR determinant profile observed in the isolates sequenced in this study was comparable and included only chromosome-borne genes encoding resistance to β -lactams. The widespread presence of β -lactam resistance genes, including those conferring resistance to carbapenems, is a known phenomenon in the *A. veronii* population (10). Although these genes were chromosomal with no close proximity to mobile genetic elements, their spread to other strains or bacterial species is not entirely unlikely as bacterial cell lysis could release DNA into the environment where it could be taken up by other strains or bacterial species through the process of transformation. Indeed, natural transformation has been described as a common mechanism of horizontal gene transfer (HGT) among *Aeromonas* species, including *A. veronii*. *Aeromonas* species are capable of competence and transformation (70). In addition, *A. veronii* is known to easily acquire and exchange AMR genes (7, 20, 71). Although there was a low occurrence of AMR in *A. veronii* sequenced in this study, Lake Wilcox is a potential reservoir for AMR genes encoding resistance to multiple antibiotics as evidenced by

results from previous studies on the lake where other bacterial species isolated from the lake carried multiple AMR genes (21, 22). The assessment of AMR gene profiles in global *A. veronii* revealed variability in isolates across different sources, highlighting the influence of environmental and host-specific factors on the emergence and spread of AMR in this bacterial population (72–74). More so, the detection of virulence genes in the isolates sequenced in this study, particularly those associated with the T3SS in human pathogenic strains, further emphasizes the potential health risks posed by these bacteria. The T3SS appeared to be predominant in global isolates recovered from the aquatic environment, freshwater fish, and human but less prevalent in animal products. Although T3SS is linked to severity of disease by Gram-negative bacterial pathogen such as *Aeromonas* (75), they have been documented to play a key role in the bacterial adaptation to changing environmental conditions (68, 75–77). They have also been noted to be dynamic and constantly exchanging components and facilitating interactions between other microorganisms including fungi, depending on the ecological and evolutionary needs (68, 75–77). The predominance of the *mrkABCD* operon, which encodes type III fimbriae involved in biofilm formation on diverse surfaces (36, 37), in isolates recovered from animal products is notable. This operon may contribute to the persistence of *A. veronii* in animal processing facilities, potentially leading to microbial contamination of animal products (13, 78).

The mobilome is known to facilitate gene gain and loss, a phenomenon that plays a crucial role in bacterial evolution and ecological adaptation, and a probable change in bacterial fitness (79, 80). This change can contribute to the emergence of divergent bacterial populations with unique features, including higher pathogenic potential (79, 81, 82). Although we did not detect plasmids in our isolates, global data revealed their presence and association with multidrug resistance genes. Of note was a phage-like plasmid that was found to contain genes encoding resistance to nine classes of antimicrobials including heavy metal (mercury) (54). Phage-like plasmids are plasmids that share structural and functional features of phage and enable them to replicate and facilitate the HGT of genes (including AMR) between bacteria (55, 83, 84). Other MGEs including prophages and insertion sequences were identified. The majority of the intact prophages were identified as P2-like phages (Pseudoviridae) (57), and a few of them were predicted to have multiple host bacterial species. This observation is interesting and could suggest a broad host range of these phages, which could have applications in biocontrol (85–87), but further studies on the host range of these phages would be needed to ascertain this. Another factor that contributes to the rapid evolution and ecological adaptation and that could influence the pathogenicity of bacterial species is BGCs that encode the production of various secondary metabolites (60, 88, 89). This phenomenon is seldom studied in *A. veronii*.

In this study, we found a high abundance of novel BGCs and identified unique NRPS and RiPP that were conserved in *A. veronii*. Notably, NRPS with high similarity (0.9) to enterobactin was found in *Pseudomonas* sp. J465 (53), which mediates high affinity for iron acquisition in stringent conditions (90, 91). Angustmycin A/B/C (51) and pseudopyrroline A/B (52) homologs were found to be conserved in *A. veronii*. These RiPP products encode antimicrobial properties and contribute to the survival of their producers in their ecological niche (59, 92). These conserved clusters could be promising genomic markers for typing *A. veronii*. Of note, a bacilysin homolog gene (64) was detected in a pair of non-human pathogenic strains. Bacilysin is an antimicrobial dipeptide produced by *Bacillus* spp. that exhibits antagonistic activity against both Gram-negative and Gram-positive bacteria (64, 93, 94). Further studies would be required to decipher the antimicrobial activity of the bacilysin homolog identified in this study against human pathogenic strains of *A. veronii* and other pathogens, as well as their mechanism of actions.

Conclusion

The study presents a genomic analysis of *A. veronii* strains isolated from a freshwater lake, defines the population structure, and characterizes the genetic factors associated with stress and ecological adaptation. A significant finding is the pathogenic potential of *A. veronii* to humans that underscores the public health implications, especially considering the recreational use of the lake. The MGEs identified that could contribute to the genetic diversity, adaptability, and pathogenicity to human, as well as the role of *A. veronii* as a reservoir for AMR genes, while the BGCs identified presents opportunities for the discovery of novel bioactive compounds. Overall, this study not only contributes to our understanding of the genetic diversity and ecological dynamics of *A. veronii* but also highlights the potential public health risks and AMR reservoir role of this bacterium. It underscores the need for continuous surveillance for pathogens in aquatic ecosystems.

ACKNOWLEDGMENTS

We acknowledge the support from the Canada Research Chairs Program and Canada First Research Excellence fund in support of this project.

O.L., N.B., V.P., R.A., M.S., Y.C., and M.P. conducted the sampling, isolation, and whole-genome sequencing. O.L. performed the bioinformatics analysis and wrote the original draft of the manuscript. O.L. and L.G. designed the project. L.G. provided funding and resources. O.L., V.P., and L.G. supervised the study. All authors read and approved the final manuscript.

AUTHOR AFFILIATIONS

¹Canadian Research Institute for Food Safety (CRIFS), University of Guelph, Guelph, Ontario, Canada

²Bayview Secondary School, Richmond Hill, Ontario, Canada

AUTHOR ORCIDs

Opeyemi U. Lawal  <http://orcid.org/0000-0003-2352-2832>

Lawrence Goodridge  <http://orcid.org/0009-0008-9965-3509>

AUTHOR CONTRIBUTIONS

Opeyemi U. Lawal, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review and editing | Noah Bryan, Investigation | Valeria R. Parreira, Investigation, Methodology, Writing – review and editing | Yanhong Chen, Investigation | Melinda Precious, Investigation, Writing – review and editing | Lawrence Goodridge, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review and editing.

DATA AVAILABILITY

The whole-genome sequences reported in this study were deposited at DDBJ/ENA/GenBank under the BioProject accession numbers [PRJNA893208](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA893208). The raw sequence reads and genome assembly accession numbers are listed in Table 1. In addition, accession numbers and associated metadata of genomes retrieved from NCBI are listed in Table S3.

ADDITIONAL FILES

The following material is available [online](#).

Supplemental Material

Table S1 (Spectrum01171-24-s0001.csv). Summary of the novel alleles and sequence types identified in *Aeromonas* isolates.

Table S2 (Spectrum01171-24-s0002.csv). Genetic distance among *Aeromonas veronii* isolates sequenced in this study.

Table S3 (Spectrum01171-24-s0003.csv). List of accession number and associated metadata of publicly available *Aeromonas veronii* genomes.

Table S4 (Spectrum01171-24-s0004.csv). Antimicrobial resistance genes in a global collection of *Aeromonas veronii*.

Table S5 (Spectrum01171-24-s0005.csv). Distribution of plasmid types in a global collection of *Aeromonas veronii*.

Fig S1 (Spectrum01171-24-s0006.tiff). Distribution of plasmid types in a global collection of *Aeromonas veronii*.

REFERENCES

- Rodrigues C, Cunha MÂ. 2017. Assessment of the microbiological quality of recreational waters: indicators and methods. *Euro-Mediterr J Environ Integr* 2:25. <https://doi.org/10.1007/s41207-017-0035-8>
- Stec J, Kosikowska U, Mendrycka M, Stępień-Pyśniak D, Niedźwiedzka-Rystwej P, Bębnowska D, Hryniewicz Z, Ziętara-Wysocka J, Grywalska E. 2022. Opportunistic pathogens of recreational waters with emphasis on antimicrobial resistance—a possible subject of human health concern. *Int J Environ Res Public Health* 19:7308. <https://doi.org/10.3390/ijerph19127308>
- Ramírez-Castillo FY, Loera-Muro A, Jacques M, Garneau P, Avelar-González FJ, Harel J, Guerrero-Barrera AL. 2015. Waterborne pathogens: detection methods and challenges. *Pathogens* 4:307–334. <https://doi.org/10.3390/pathogens4020307>
- Steven JAC, Thorn RMS, Robinson GM, Turner D, Lee JE, Reynolds DM. 2022. The control of waterborne pathogenic bacteria in fresh water using a biologically active filter. *npj Clean Water* 5:30. <https://doi.org/10.1038/s41545-022-00169-y>
- Liu F, Yuwono C, Tay ACY, Wehrhahn MC, Riordan SM, Zhang L. 2022. Analysis of global *Aeromonas veronii* genomes provides novel information on source of infection and virulence in human gastrointestinal diseases. *BMC Genomics* 23:166. <https://doi.org/10.1186/s12864-022-08402-1>
- Ramsamy Y, Mlisana KP, Amoako DG, Abia ALK, Allam M, Ismail A, Singh R, Essack SY. 2020. Comparative pathogenomics of *Aeromonas veronii* from pigs in South Africa: dominance of the novel ST657 clone. *Microorganisms* 8:2008. <https://doi.org/10.3390/microorganisms8122008>
- Sakulworakan R, Chokmangmeepisarn P, Dinh-Hung N, Sivaramasamy E, Hirono I, Chuanchuen R, Kayansamruaj P, Rodkhum C. 2021. Insight into whole genome of *Aeromonas veronii* isolated from freshwater fish by resistome analysis reveal extensively antibiotic resistant traits. *Front Microbiol* 12:733668. <https://doi.org/10.3389/fmicb.2021.733668>
- Janda JM, Abbott SL. 2010. The genus *Aeromonas*: taxonomy, pathogenicity, and infection. *Clin Microbiol Rev* 23:35–73. <https://doi.org/10.1128/CMR.00039-09>
- Lamy B, Kodjo A, colBVH Study Group, Laurent F. 2009. Prospective nationwide study of *Aeromonas* infections in France. *J Clin Microbiol* 47:1234–1237. <https://doi.org/10.1128/JCM.00155-09>
- Drk S, Puljko A, Dželalija M, Udiković-Kolić N. 2023. Characterization of third generation cephalosporin- and carbapenem-resistant *Aeromonas* isolates from municipal and hospital wastewater. *Antibiotics (Basel)* 12:513. <https://doi.org/10.3390/antibiotics12030513>
- Senderovich Y, Ken-Dror S, Vainblat I, Blau D, Izhaki I, Halpern M. 2012. A molecular study on the prevalence and virulence potential of *Aeromonas* spp. recovered from patients suffering from diarrhea in Israel. *PLoS ONE* 7:e30070. <https://doi.org/10.1371/journal.pone.0030070>
- Hoel S, Vadstein O, Jakobsen AN. 2019. The significance of mesophilic *Aeromonas* spp. in minimally processed ready-to-eat seafood. *Microorganisms* 7:91. <https://doi.org/10.3390/microorganisms7030091>
- Martins LM, Marquez RF, Yano T. 2002. Incidence of toxic *Aeromonas* isolated from food and human infection. *FEMS Immunol Med Microbiol* 32:237–242. <https://doi.org/10.1111/j.1574-695X.2002.tb00559.x>
- Zh E, Rn Z, Rr S. 2021. Prevalence and characterization of *Aeromonas* spp. isolated from some meat products in Egypt. *Nutr Food Technol Open Access* 7. <https://doi.org/10.16966/2470-6086.170>
- Manna SK, Maurye P, Dutta C, Samanta G. 2013. Occurrence and virulence characteristics of *Aeromonas* species in meat milk and fish in India. *J Food Saf* 33:461–469. <https://doi.org/10.1111/jfs.12077>
- Nawaz M, Khan SA, Khan AA, Sung K, Tran Q, Kerdahi K, Steele R. 2010. Detection and characterization of virulence genes and integrons in *Aeromonas veronii* isolated from catfish. *Food Microbiol* 27:327–331. <https://doi.org/10.1016/j.fm.2009.11.007>
- Nawaz M, Sung K, Khan SA, Khan AA, Steele R. 2006. Biochemical and molecular characterization of tetracycline-resistant *Aeromonas veronii* isolates from catfish. *Appl Environ Microbiol* 72:6461–6466. <https://doi.org/10.1128/AEM.00271-06>
- Talagrand-Reboul E, Colston SM, Graf J, Lamy B, Jumas-Bilak E. 2020. Comparative and evolutionary genomics of isolates provide insight into the pathoadaptation of *Aeromonas*. *Genome Biol Evol* 12:535–552. <https://doi.org/10.1093/gbe/evaa055>
- Hatrongjit R, Kerdsin A, Takeuchi D, Wongsurawat T, Jenjaroenpun P, Chopjitt P, Boueroy P, Akeda Y, Hamada S. 2020. Genomic analysis of *Aeromonas veronii* C198, a novel *Mcr-3.41*-harboring isolate from a patient with septicemia in Thailand. *Pathogens* 9:1031. <https://doi.org/10.3390/pathogens9121031>
- Zhu Z, Wu S, Zhu J, Wang T, Wen Y, Yang C, Lv J, Zhang H, Chen L, Du H. 2023. Emergence of *Aeromonas veronii* strain co-harboring blaKPC–2, mcr–3.17, and tmexC3.2–tmexD3.3–toprJ1b cluster from hospital sewage in China. *Front Microbiol* 14:1115740. <https://doi.org/10.3389/fmicb.2023.1115740>
- Bryan N, Anderson R, Lawal OU, Parreira VR, Goodridge L. 2023. Draft genome sequence of *Bacillus anthracis* N1, isolated from a recreational freshwater Kettle Lake in Ontario, Canada. *Microbiol Resour Annot* 12:e0126222. <https://doi.org/10.1128/mra.01262-22>
- Lawal OU, Bryan N, Soni M, Chen Y, Precious M, Parreira VR, Goodridge L. 2024. Whole genome sequence of *Vibrio cholerae* NB-183 isolated from freshwater in Ontario, Canada harbors a unique gene repertoire. *BMC Genom Data* 25:18. <https://doi.org/10.1186/s12863-024-01204-2>
- Bryan N, Anderson R, Lawal OU, Parreira VR, Goodridge L. 2023. Draft genome sequence of *Exiguobacterium* sp. strain N5, isolated from a recreational freshwater Kettle Lake in Ontario. *Microbiol Resour Annot* 12:e0126122. <https://doi.org/10.1128/mra.01261-22>
- Karama M, Lawal OU, Parreira VR, Soni M, Chen Y, Cenci-Goga BT, Grispoldi L, Greyling J, Goodridge L. 2024. Draft genome sequences of two *Salmonella* Uzaramo isolates from poultry in South Africa. *Microbiol Resour Annot* 13:e0102623. <https://doi.org/10.1128/mra.01026-23>
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>

26. Souvorov A, Agarwala R, Lipman DJ. 2018. SKESA: strategic k-mer extension for scrupulous assemblies. *Genome Biol* 19:153. <https://doi.org/10.1186/s13059-018-1540-z>
27. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29:1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>
28. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
29. Jolley KA, Bray JE, Maiden MCJ. 2018. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 3:124. <https://doi.org/10.12688/wellcomeopenres.14826.1>
30. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 15:R46. <https://doi.org/10.1186/gb-2014-15-3-r46>
31. Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>
32. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30:2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>
33. Cosentino S, Voldby Larsen M, Möller Aarestrup F, Lund O. 2013. PathogenFinder—distinguishing friend from foe using bacterial whole genome sequence data. *PLoS ONE* 8:e77302. <https://doi.org/10.1371/journal.pone.0077302>
34. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, Tyson GH, Zhao S, Hsu C-H, McDermott PF, Tadesse DA, Morales C, Simmons M, Tillman G, Wasilenko J, Folster JP, Klimke W. 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrob Agents Chemother* 63:e00483-19. <https://doi.org/10.1128/AAC.00483-19>
35. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, Huynh W, Nguyen A-LV, Cheng AA, Liu S, et al. 2020. CARD 2020: antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 48:D517–D525. <https://doi.org/10.1093/nar/gkz935>
36. Chen L, Zheng D, Liu B, Yang J, Jin Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res* 44:D694–7. <https://doi.org/10.1093/nar/gkv1239>
37. Liu B, Zheng D, Zhou S, Chen L, Yang J. 2022. VFDB 2022: a general classification scheme for bacterial virulence factors. *Nucleic Acids Res* 50:D912–D917. <https://doi.org/10.1093/nar/gkab1107>
38. Robertson J, Nash JHE. 2018. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom* 4:1–7. <https://doi.org/10.1099/mgen.0.000206>
39. Wishart DS, Han S, Saha S, Oler E, Peters H, Grant JR, Stothard P, Gautam V. 2023. PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res* 51:W443–W450. <https://doi.org/10.1093/nar/gkad382>
40. Shang J, Peng C, Liao H, Tang X, Sun Y. 2023. PhaBOX: A web server for identifying and characterizing phage contigs in metagenomic data. *Bioinf Adv* 3. <https://doi.org/10.1093/bioadv/vbad101>
41. Nayfach S, Camargo AP, Schulz F, Eloie-Fadrosch E, Roux S, Kyrpides NC. 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. *Nat Biotechnol* 39:578–585. <https://doi.org/10.1038/s41587-020-00774-7>
42. Shang J, Jiang J, Sun Y. 2021. Bacteriophage classification for assembled contigs using graph convolutional network. *Bioinformatics* 37:i25–i33. <https://doi.org/10.1093/bioinformatics/btab293>
43. Hockenberry AJ, Wilke CO. 2021. BACPHLP: predicting bacteriophage lifestyle from conserved protein domains. *PeerJ* 9:e11396. <https://doi.org/10.7717/peerj.11396>
44. Shang J, Sun Y. 2022. CHERRY: a computational method for accurate prediction of virus–prokaryotic interactions using a graph encoder–decoder model. *Brief Bioinformatics* 23. <https://doi.org/10.1093/bib/bbac182>
45. Lawal OU, Goodridge L. n.d. TSPDB: A curated resource of tailspike proteins with potential applications in phage research. *Genomics*. <https://doi.org/10.1101/2024.05.23.595625>
46. 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:W29–W35. <https://doi.org/10.1093/nar/gkab335>
47. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31:3691–3693. <https://doi.org/10.1093/bioinformatics/btv421>
48. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T, Keane JA, Harris SR. 2016. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2:e000056. <https://doi.org/10.1099/mgen.0.000056>
49. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>
50. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. 2016. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol* 17:1–9. <https://doi.org/10.1186/s13059-016-1108-8>
51. Lawal OU, Fraqueza MJ, Bouchami O, Worning P, Bartels MD, Gonçalves ML, Paixão P, Gonçalves E, Toscano C, Empel J, Urbaś M, Domínguez MA, Westh H, de Lencastre H, Miragaia M. 2021. Foodborne origin and local and global spread of *Staphylococcus saprophyticus* causing human urinary tract infections. *Emerg Infect Dis* 27:880–893. <https://doi.org/10.3201/eid2703.200852>
52. Lawal OU, Fraqueza MJ, Worning P, Bouchami O, Bartels MD, Gonçalves L, Paixão P, Gonçalves E, Toscano C, Empel J, Urbaś M, Domínguez MA, Westh H, de Lencastre H, Miragaia M. 2021. *Staphylococcus saprophyticus* causing infections in humans is associated with high resistance to heavy metals. *Antimicrob Agents Chemother* 65:e0268520. <https://doi.org/10.1128/AAC.02685-20>
53. 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* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
54. Hughes HY, Conlan SP, Lau AF, Dekker JP, Michelin AV, Youn J-H, Henderson DK, Frank KM, Segre JA, Palmore TN. 2016. Detection and whole-genome sequencing of carbapenemase-producing *Aeromonas hydrophila* isolates from routine perirectal surveillance culture. *J Clin Microbiol* 54:1167–1170. <https://doi.org/10.1128/JCM.03229-15>
55. Cadieux B, Lawal OU, Emond-Rheault J-G, Jeukens J, Freschi L, Kukavica-Ibrulj I, Levesque RC, Austin JW, Goodridge L. 2022. Draft genome sequences of two *Clostridium botulinum* group II strains carrying phage-like plasmids. *Microbiol Resour Announc* 11:e0009122. <https://doi.org/10.1128/mra.00091-22>
56. Colavecchio A, Jeukens J, Freschi L, Edmond Rheault J-G, Kukavica-Ibrulj I, Levesque RC, LeJeune J, Goodridge L. 2017. Complete genome sequences of two phage-like plasmids carrying the CTX-M-15 extended-spectrum β -lactamase gene. *Genome Announc* 5:e00102-17. <https://doi.org/10.1128/genomeA.00102-17>
57. Turner D, Shkoporov AN, Lood C, Millard AD, Dutilh BE, Alfenas-Zerbini P, van Zyl LJ, Aziz RK, Oksanen HM, Poranen MM, et al. 2023. Abolishment of morphology-based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial viruses subcommittee. *Arch Virol* 168:74. <https://doi.org/10.1007/s00705-022-05694-2>
58. Navarro-Muñoz JC, Selem-Mojica N, Mallowney MW, Kautsar SA, Tryon JH, Parkinson EI, De Los Santos ELC, Yeong M, Cruz-Morales P, Abubucker S, Roeters A, Lokhorst W, Fernandez-Guerra A, Cappellini LTD, 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* 16:60–68. <https://doi.org/10.1038/s41589-019-0400-9>
59. Yu L, Zhou W, She Y, Ma H, Cai Y-S, Jiang M, Deng Z, Price NPJ, Chen W. 2021. Efficient biosynthesis of nucleoside cytokinin angustmycin A containing an unusual sugar system. *Nat Commun* 12:6633. <https://doi.org/10.1038/s41467-021-26928-y>
60. Bauer JS, Ghequire MGK, Nett M, Josten M, Sahl H-G, De Mot R, Gross H. 2015. Biosynthetic origin of the antibiotic pseudopyronines A and B in

- Pseudomonas putida* BW11M1. *ChemBiochem* 16:2491–2497. <https://doi.org/10.1002/cbic.201500413>
61. Curson ARJ, Sullivan MJ, Todd JD, Johnston AWB. 2010. Identification of genes for dimethyl sulfide production in bacteria in the gut of Atlantic Herring (*Clupea harengus*). *ISME J* 4:144–146. <https://doi.org/10.1038/ismej.2009.93>
 62. Litomska A, Ishida K, Dunbar KL, Boettger M, Coyne S, Hertweck C. 2018. Enzymatic thioamide formation in a bacterial antimetabolite pathway. *Angew Chem Int Ed* 57:11574–11578. <https://doi.org/10.1002/anie.201804158>
 63. Grammbitter GLC, Schmalhofer M, Karimi K, Shi Y-M, Schöner TA, Tobias NJ, Morgner N, Groll M, Bode HB. 2019. An uncommon type II PKS catalyzes biosynthesis of Aryl polyene pigments. *J Am Chem Soc* 141:16615–16623. <https://doi.org/10.1021/jacs.8b10776>
 64. Moran S, Robertson K, Paradisi F, Rai DK, Murphy CD. 2010. Production of lipopeptides in *Bacillus* sp. CS93 isolated from Pozol. *FEMS Microbiol Lett* 304:69–73. <https://doi.org/10.1111/j.1574-6968.2009.01882.x>
 65. Tohya M, Teramoto K, Watanabe S, Hishinuma T, Shimojima M, Ogawa M, Tada T, Tabe Y, Kirikae T. 2022. Whole-genome sequencing-based re-identification of *Pseudomonas putida/fluorescens* clinical isolates identified by biochemical bacterial identification systems. *Microbiol Spectr* 10:e0249121. <https://doi.org/10.1128/spectrum.02491-21>
 66. Garza-González E, Bocanegra-Ibarias P, Dinh A, Morfin-Otero R, Camacho-Ortiz A, Rojas-Larios F, Rodríguez-Zulueta P, Arias CA. 2020. Species identification of *Enterococcus* spp: whole genome sequencing compared to three biochemical test-based systems and two matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) systems. *J Clin Lab Anal* 34:e23348. <https://doi.org/10.1002/jcla.23348>
 67. Chacón MR, Soler L, Groisman EA, Guarro J, Figueras MJ. 2004. Type III secretion system genes in clinical *Aeromonas* isolates. *J Clin Microbiol* 42:1285–1287. <https://doi.org/10.1128/JCM.42.3.1285-1287.2004>
 68. Stuber K, Frey J, Burnens AP, Kuhnert P. 2003. Detection of type III secretion genes as a general indicator of bacterial virulence. *Mol Cell Probes* 17:25–32. [https://doi.org/10.1016/s0890-8508\(02\)00108-1](https://doi.org/10.1016/s0890-8508(02)00108-1)
 69. Coburn B, Sekirov I, Finlay BB. 2007. Type III secretion systems and disease. *Clin Microbiol Rev* 20:535–549. <https://doi.org/10.1128/CMR.00013-07>
 70. Huddleston JR, Brokaw JM, Zak JC, Jeter RM. 2013. Natural transformation as a mechanism of horizontal gene transfer among environmental *Aeromonas* species. *Syst Appl Microbiol* 36:224–234. <https://doi.org/10.1016/j.syapm.2013.01.004>
 71. Algammal AM, Ibrahim RA, Alfifi KJ, Ghabban H, Alghamdi S, Kabrah A, Khafagy AR, Abou-Elala GM, Abu-Elala NM, Donadu MG, El-Tarabili RM. 2022. A first report of molecular typing, virulence traits, and phenotypic and genotypic resistance patterns of newly emerging XDR and MDR *Aeromonas veronii* in *Mugil sehelii*. *Pathogens* 11:1262. <https://doi.org/10.3390/pathogens11111262>
 72. Goryluk-Salmonowicz A, Popowska M. 2022. Factors promoting and limiting antimicrobial resistance in the environment - Existing knowledge gaps. *Front Microbiol* 13:992268. <https://doi.org/10.3389/fmicb.2022.992268>
 73. Bengtsson-Palme J, Kristiansson E, Larsson DGJ. 2018. Environmental factors influencing the development and spread of antibiotic resistance. *FEMS Microbiol Rev* 42:fux053. <https://doi.org/10.1093/femsre/fux053>
 74. Fletcher S. 2015. Understanding the contribution of environmental factors in the spread of antimicrobial resistance. *Environ Health Prev Med* 20:243–252. <https://doi.org/10.1007/s12199-015-0468-0>
 75. Rahmatelahi H, El-Matbouli M, Menanteau-Ledouble S. 2021. Delivering the pain: an overview of the type III secretion system with special consideration for aquatic pathogens. *Vet Res* 52:146. <https://doi.org/10.1186/s13567-021-01015-8>
 76. Nazir R, Mazurier S, Yang P, Lemanceau P, van Elsas JD. 2017. The ecological role of type three secretion systems in the interaction of bacteria with fungi in soil and related habitats is diverse and context-dependent. *Front Microbiol* 8:38. <https://doi.org/10.3389/fmicb.2017.00038>
 77. Milne-Davies B, Wimmi S, Diepold A. 2021. Adaptivity and dynamics in type III secretion systems. *Mol Microbiol* 115:395–411. <https://doi.org/10.1111/mmi.14658>
 78. Erickson VI, Alfifi A, Hounmanou YGM, Sana MJ, Christensen JP, Dalsgaard A. 2023. Genomic traits of *Aeromonas veronii* isolated from slaughtered Danish broilers. *Vet Microbiol* 283:109772. <https://doi.org/10.1016/j.vetmic.2023.109772>
 79. Hazen TH, Pan L, Gu J-D, Sobczyk PA. 2010. The contribution of mobile genetic elements to the evolution and ecology of *Vibrios*. *FEMS Microbiol Ecol* 74:485–499. <https://doi.org/10.1111/j.1574-6941.2010.00937.x>
 80. Vale FF, Lehours P, Yamaoka Y. 2022. Editorial: the role of mobile genetic elements in bacterial evolution and their adaptability. *Front Microbiol* 13:849667. <https://doi.org/10.3389/fmicb.2022.849667>
 81. Huang S, Li H, Ma L, Liu R, Li Y, Wang H, Lu X, Huang X, Wu X, Liu X. 2023. Insertion sequence contributes to the evolution and environmental adaptation of *Acidithiobacillus*. *BMC Genomics* 24:282. <https://doi.org/10.1186/s12864-023-09372-8>
 82. Kozitskaya S, Cho SH, Dietrich K, Marre R, Naber K, Ziebuhr W. 2004. The bacterial insertion sequence element IS256 occurs preferentially in nosocomial *Staphylococcus epidermidis* isolates: association with biofilm formation and resistance to aminoglycosides. *Infect Immun* 72:1210–1215. <https://doi.org/10.1128/IAI.72.2.1210-1215.2004>
 83. Pfeifer E, Rocha EPC. 2024. Phage-plasmids promote recombination and emergence of phages and plasmids. *Nat Commun* 15:1545. <https://doi.org/10.1038/s41467-024-45757-3>
 84. Pfeifer Eugen, Moura de Sousa JA, Touchon M, Rocha EPC. 2021. Bacteria have numerous distinctive groups of phage-plasmids with conserved phage and variable plasmid gene repertoires. *Nucleic Acids Res* 49:2655–2673. <https://doi.org/10.1093/nar/gkab064>
 85. Ross A, Ward S, Hyman P. 2016. More is better: selecting for broad host range bacteriophages. *Front Microbiol* 7:1352. <https://doi.org/10.3389/fmicb.2016.01352>
 86. Myers J, Davis li J, Lollo M, Hudec G, Hyman P. 2023. More's the same-multiple hosts do not select for broader host range phages. *Viruses* 15:518. <https://doi.org/10.3390/v15020518>
 87. Fong K, Wong CWY, Wang S, Delaquis P. 2021. How broad is enough: the host range of bacteriophages and its impact on the agri-food sector. *PHAGE (N Rochelle)* 2:83–91. <https://doi.org/10.1089/phage.2020.0036>
 88. Gavrilidou A, Kautsar SA, Zaburannyi N, Krug D, Müller R, Medema MH, Ziemert N. 2022. Compendium of specialized metabolite biosynthetic diversity encoded in bacterial genomes. *Nat Microbiol* 7:726–735. <https://doi.org/10.1038/s41564-022-01110-2>
 89. Xia L, Miao Y, Cao A, Liu Y, Liu Z, Sun X, Xue Y, Xu Z, Xun W, Shen Q, Zhang N, Zhang R. 2022. Biosynthetic gene cluster profiling predicts the positive association between antagonism and phylogeny in *Bacillus*. *Nat Commun* 13:1023. <https://doi.org/10.1038/s41467-022-28668-z>
 90. Moynié L, Milenkovic S, Mislin GLA, Gasser V, Mallocci G, Baco E, McCaughan RP, Page MGP, Schalk IJ, Ceccarelli M, Naismith JH. 2019. The complex of ferric-enterobactin with its transporter from *Pseudomonas aeruginosa* suggests a two-site model. *Nat Commun* 10:3673. <https://doi.org/10.1038/s41467-019-11508-y>
 91. Saha P, Xiao X, Yeoh BS, Chen Q, Katkere B, Kirmanjeswara GS, Vijay-Kumar M. 2019. The bacterial siderophore enterobactin confers survival advantage to *Salmonella* in macrophages. *Gut Microbes* 10:412–423. <https://doi.org/10.1080/19490976.2018.1546519>
 92. Bouthillette LM, Darcey CA, Handy TE, Seaton SC, Wolfe AL. 2017. Isolation of the antibiotic pseudopyronine B and SAR evaluation of C3/C6 alkyl analogs. *Bioorg Med Chem Lett* 27:2762–2765. <https://doi.org/10.1016/j.bmcl.2017.04.067>
 93. Islam T, Rabbee MF, Choi J, Baek K-H. 2022. Biosynthesis, molecular regulation, and application of bacilysin produced by *Bacillus* species. *Metabolites* 12:397. <https://doi.org/10.3390/metabo12050397>
 94. Nannan C, Vu HQ, Gillis A, Caulier S, Nguyen TTT, Mahillon J. 2021. Bacilysin within the *Bacillus subtilis* group: gene prevalence versus antagonistic activity against Gram-negative foodborne pathogens. *J Biotechnol* 327:28–35. <https://doi.org/10.1016/j.biotech.2020.12.017>