

Conserved *Vibrio* Repeats (VXRs) mediate horizontal transfer of superintegron-associated antibiotic resistance genes across *Vibrio* species

Sergio Morgado

sergio.morgado@ioc.fiocruz.br

Instituto Oswaldo Cruz, FIOCRUZ

Erica Fonseca

Instituto Oswaldo Cruz, FIOCRUZ

Ana Carolina Vicente

Instituto Oswaldo Cruz, FIOCRUZ

Research Article

Keywords: attC, recombination, cassette, evolution, sedentary, integrase

Posted Date: August 24th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9861061/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Functional & Integrative Genomics on October 6th, 2026. See the published version at <https://doi.org/10.1007/s10142-026-02068-8>.

Abstract

Superintegrons (SIs) are large chromosomal integrons that harbor extensive arrays of gene cassettes and constitute major reservoirs of bacterial genetic diversity. In *Vibrio*, SI cassettes are flanked by species-associated attC repeats, here referred to as *Vibrio* species repeats (VXRs). Although SIs have been proposed as ancestral sources of antibiotic resistance genes (ARGs), the extent to which SI-associated ARGs are exchanged among *Vibrio* species remains poorly understood. To investigate the dynamics of SI-associated ARGs, we analyzed 37,812 genomes representing 181 *Vibrio* species. Using VXR mapping, comparative phylogenetics, and genome screening, we identified 8,717 putative ARG cassettes distributed across 8,523 genomes from 11 species. Despite the extensive diversity of VXRs identified in SIs, ARG cassettes were predominantly associated with a small subset of conserved VXRs shared among phylogenetically distant species. Moreover, phylogenetic incongruence between VXR and host species trees supported horizontal transfer of ARG-associated cassettes across the genus. Comparative analyses further suggested that members of the Cholera clade may act as potential sources of SI-associated ARG dissemination. Five major ARG families were identified within SIs, including *catB*, *qnrVC*, *bla*_{CARB}, *dfrA*, and *vatF*, some of which displayed phylogenetic patterns consistent with interspecies transfer. Functional characterization of a SI-associated *qnrVC4* cassette from *Vibrio mimicus* confirmed its ability to confer reduced susceptibility to ciprofloxacin in a heterologous host. Together, our findings indicate that a conserved subset of VXR backbones, mainly from the Cholera clade, mediates ARG dissemination across *Vibrio* species, highlighting superintegrons as active evolutionary reservoirs contributing to the emergence and spread of antimicrobial resistance.

1. Introduction

Integrons are genetic platforms specialized in the capture, excision, and expression of gene cassettes, flanked by species-specific direct repeats, through the action of a site-specific recombinase, the integrase IntI (Fluit and Schmitz 2004). While mobile integrons are well known for their crucial role in disseminating antimicrobial resistance genes (ARGs) among clinical pathogens, the vast majority of cassette diversity resides in sedentary chromosomal integrons, also called superintegrons (SIs), large and stable arrays of gene cassettes (up to 200 or more), which act as biobanks (Darracq et al. 2025). First characterized in *Vibrio cholerae*, SIs represent one of the richest reservoirs of genetic innovation in bacteria and have been proposed as the ancestral source from which mobile integrons evolved (Rowe-Magnus et al. 2001). In *Vibrio*, the species-specific direct repeats that flank the gene cassettes are called *Vibrio* species repeats (VXR).

A key feature of integron biology is the mobility of gene cassettes, which may undergo excision and reinsertion within an SI or transfer between species. Experimental work has provided direct evidence that SIs in *V. cholerae* are dynamic structures capable of cassette recruitment. Rowe-Magnus et al. (2002) showed real-time recruitment of a *V. cholerae* SI-derived *cat* gene into a mobile class 1 resistance integron during antibiotic selection, after which it could be disseminated to clinically relevant pathogens via conjugation. SI diversification can also occur under natural transformation conditions through competence-mediated exchange, including environmental settings, resulting in extensive cassette reshuffling and acquisition of novel genes (Debatisse et al. 2025; Righi et al. 2026). Comparative genomic studies also indicate that SIs serve as sources of gene cassettes for mobile integrons. In a *V. cholerae* O1 strain, Fonseca et al. (2008) identified a *qnrVC* cassette, in a class 1 integron, flanked by attC sites highly similar to repeats found in the SIs of *V. parahaemolyticus* and *V. cholerae*, suggesting movement of cassettes among species and between chromosomal SIs and mobile integrons. More broadly, reviews of SI diversity in *Vibrio* highlight the frequent presence of cassettes associated with metabolism, stress resistance, and antimicrobial tolerance, many of which also occur in mobile integrons circulating among human pathogens (Rapa and Labbate 2013). Together, these observations support the view of SIs as evolutionary hubs that generate, diversify, and disseminate gene cassettes, including ARGs, across bacterial populations.

Despite these advances, major questions remain unresolved. In particular, it is unclear to what extent ARG-containing cassettes within SIs are endogenous to specific species or acquired horizontally from distantly related lineages. In addition, most knowledge of SIs comes from *V. cholerae*. The genus *Vibrio* is exceptionally diverse, with numerous species harboring extensive SIs of varying size, structure, and repeat composition (Gillings 2014). However, no large-scale study has evaluated the diversity and origin of SI-associated ARGs across the genus or determined whether particular clades function as donors or recipients of cassette flow.

In this study, by combining VXR identification, genome mapping of VXR-flanked ARGs, and phylogenetics of repeats, cassettes, and host genomes, we aim to identify the SI-associated ARGs and their distribution among *Vibrio* species.

2. Methods

2.1. Genome sampling, superintegron detection and VXR extraction

From NCBI (Sept 2025), *Vibrio* genomes exhibiting completeness > 90%, contamination < 6%, and ≤ 3 contigs were screened to identify representative genomes with complete SI, resulting in the selection of 903 genomes. From this dataset, one random genome per species (n = 85 species) was retained for integron analysis.

The SI of these 85 representative genomes was searched using Integron Finder v2.0.6 with `-local_max` parameter (Néron et al. 2022). Additionally, the putative SIs had their integrase analyzed to confirm them as Int4, characteristic of SIs. Integrase sequences were recovered and submitted to phylogenetic clustering, identifying 47 Int4-like integrases. Thus, we considered 47 of 85 sequences as SIs. All *Vibrio* attC-associated repeats (VXRs) from these 47 SIs were extracted, yielding 531 VXR sequences.

Since VXR prediction by Integron Finder could generate false-positive attC sites (such as degenerate VXRs or palindromic motifs), sequences shared across different species were considered to represent true integron repeats. Thus, cd-hit v4.8.1 (Li et al. 2006) was used for VXR clustering.

2.3. Large-scale genome screening for ARG cassettes

A total of 37,812 *Vibrio* genomes, encompassing 181 species, were obtained from NCBI (Sept 2025). Among these, *V. cholerae* (14,351 genomes) and *V. parahaemolyticus* (14,034 genomes) were the most represented.

First, ARGs were identified using abricate tool (<https://github.com/tseemann/abricate>) with Comprehensive Antibiotic Resistance Database (CARD) database, considering a minimum threshold of $\geq 70\%$ sequence identity and $\geq 70\%$ coverage. Next, for each ARG, its nucleotide sequence plus 500 bp of each flanking region were extracted.

The ARGs and their flanking regions were subjected to BLASTN using the 531 VXR, predicted previously, as queries. Hits were retained if: coverage $> 70\%$, identity $> 70\%$, alignment length > 100 bp, and bordering both sides of the ARG. In these conditions, ARGs were classified as putative SI-associated ARGs.

2.4. Phylogenetic analyses

For phylogenomic reconstruction of representative genomes, the core genes were identified across the dataset using Roary v3.13.0 (Page et al. 2015). The core gene alignment was processed with snp-sites v2.5.1 (Page et al. 2016) to extract single nucleotide polymorphisms (SNPs). Maximum-likelihood phylogenetic trees (including core genome, gene, and VXR trees) were inferred with IQ-TREE v2.4.0 (Minh et al. 2020) considering 1,000 ultrafast bootstrap replicates and MFP + ASC parameters.

2.5. Strain, genomic sequencing, and functional analyses in a heterologous system

To assess the functionality of ARGs of SIs, we examined a *V. mimicus* strain (VM343) isolated from the environment (water) in 1990 in Japan (Morgado et al. 2024). Genomic sequencing of the strain was performed using ONT technology with a PromethION P2 Solo and a PromethION Flow Cell (FLO-PRO114M), alongside the Native Barcoding Kit 96 V14 (SQK-NBD114.96). The MinKNOW program was used for sequencing management, including real-time monitoring, and basecalling was performed using the maximum accuracy model v5.0.0–400 bps. The genome was assembled using Flye v2.9.6 with the nano-hq mode, and deposited in the Genbank (JAZHPO000000000.2).

Specific primers were used to amplify the complete *qnrVC* gene (5'-GAACGTCAATTTGGCACCT-3' and 5'-CTGTCAGGATTGCGACTCAT-3') from strain VM343. The PCR product was cloned into the pGEM T-Easy Cloning Vector System (Promega) and used to transform *Escherichia coli* DH5 α . The competent cells were plated onto LB containing ampicillin 100 $\mu\text{g/mL}$. After confirming the insert orientation relative to the vector promoter, the ciprofloxacin MICs for the transformants were determined using the Etest method (bioMérieux) on Mueller-Hinton agar. *E. coli* DH5 α was used as a control, exhibiting a ciprofloxacin MIC of 0.05 $\mu\text{g/mL}$.

3. Results

3.1. *Vibrio* species carrying ARGs in the SI

Screening of 37,812 *Vibrio* genomes (181 species) for antibiotic resistance gene cassettes associated with SIs (ARG/SI) revealed putative 8,717 cassette sequences from 8,523 genomes ($\sim 22\%$). These sequences were distributed in the 11 species, in addition to two sp. genomes (Supplementary Material 1, Table S1): *Vibrio fujianensis* (1/1), *Vibrio misgurnus* (1/2), *Vibrio navarrensis* (1/46), *Vibrio vulnificus* (4/2,806), *Vibrio metoecus* (6/41), *Vibrio metschnikovii* (8/191), *Vibrio paracholerae* (19/63), *Vibrio mimicus* (20/116), *Vibrio parahaemolyticus* (22/14,351), *Vibrio anguillarum* (58/301), and *Vibrio cholerae* (8381/14,034). Although *V. cholerae* and *V. parahaemolyticus* genomes are overrepresented, a distinction is visible among species concerning ARG/SIs. Despite both species having thousands of sequenced genomes, *V. parahaemolyticus* harbored only a minor fraction of genomes containing ARG/SIs. Of these 8,523 genomes presenting ARG/SIs, 4,355 ($\sim 51\%$) presented metadata associated with their source, where 3,670, 523, and 162 were from humans, environments, and animals, respectively (Supplementary Material 1, Table S1). ARGs have been presented within *Vibrio* SIs for several decades, including in a *V. paracholerae* genome from 1916 (pre-antibiotic era) (Supplementary Material 1, Table S1). Most genomes carried only one ARG cassette in SI, but 186 genomes carried two or three, almost exclusively from *V. cholerae* (173 genomes). A core genome tree of these species shows that they comprise distinct *Vibrio* clades, where most of them would be closely related to *V. cholerae* (Fig. 1).

3.2. Dominance of a small set of highly conserved VXRs

Among the initial 531 VXR sequences used for screening ARG/SIs (see methods), only 63 were identified associated with ARGs, where a small set of eight VXRs was highly prevalent (8,312–8,554 out of 8,717 ARG cassettes), while most of the VXR sequences were present in less than 100 ARG cassettes (Table 1). These highly prevalent VXRs are 129–131 bp length (compatible with the size of already characterized VXRs, such as VCR), share $\geq 90\%$ nucleotide identity with each other, and appear across multiple species, suggesting a conserved backbone that is widespread among *Vibrio* species (since it was only observed the association of these VXRs with ARGs, other kind of genes might be associated with other VXRs). In addition, comparison of the VXR cladogram (Fig. 2) with the species phylogeny (Fig. 1) revealed a topological incongruence, suggesting horizontal transfer between *Vibrio* species.

Table 1
Origin and count of the most prevalent VXR sequences asso

Genome origin	# ARG cassettes	VXR sequence
GCA_009665215.1_V.navarrensis_seq3	575/8,717	ACATAACAAACGCCTCAAGTGGGACTGCCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGGT
GCA_002216685.1_V.tarraie_seq4	1,067/8,717	GCCTAACAAACGCCTCAAGATGGACTGTCAACGCGTGGCGTTTCCAGTCCCAATGAGCCGCGGT
GCA_029635775.1_V.misgurnus_seq8	1,126/8,717	GCTATAACAAATGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAACCGCGGT
GCA_000969235.1_V.cholerae_seq8	8,312/8,717	ACCTAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCTATTGAGCCGCGGT
GCA_001558475.2_V.mimicus_seq11	8,341/8,717	GGCTAACAAACGCCTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCATTGAGCCTCGGT
GCA_002211505.1_V.anguillarum_seq6	8,348/8,717	ATTTAACAAACGCCTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGT

Genome origin	# ARG cassettes	VXR sequence
GCA_051352415.1_V.cidicii_seq4	8,505/8,717	AGCTAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_002257545.1_V.qinghaiensis_seq14	8,529/8,717	ACCTAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_009665395.1_V.cincinnatiensis_seq11	8,529/8,717	CTATAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_900538065.1_V.paracholerae_seq2	8,539/8,717	GCATAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_009665255.1_V.metoecus_seq7	8,554/8,717	GCATAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1

Genome origin	# ARG cassettes	VXR sequence

Only exhibiting VXR sequences flanking more than 100 ARG cassettes.

We then evaluated the SI of specific genomes from which each prevalent VXR sequence originated (Table 1) to determine whether that repeat type was also highly represented within its native SI. For instance, is the repeat sequence GCA_000969235.1_V.cholerae_seq8 dominant within the SI of the GCA_000969235.1 genome? This analysis enabled us to infer whether a prevalent VXR (Table 1) was likely endogenous to its source SI or instead represented a repeat associated with cassettes horizontally acquired. This analysis showed that genomes belonging to the cholera clade (*V. cholerae*, *V. metoecus*, *V. paracholerae*, and *V. tarriae*) had, as the most abundant repeats in their SIs, the same VXR type that were identified as prevalent in the dataset (Table 1). The only partial exception was *V. mimicus* (GCA_001558475.2), in which the VXR GCA_001558475.2_V.mimicus_seq11 was not the dominant repeat but the second most abundant within its SI. Indeed, *V. mimicus* is the most basal lineage of the Cholera clade (Fig. 1). In contrast, the prevalent VXRs from *V. anguillarum*, *V. cidicii*, *V. cincinnatiensis*, *V. misgurnus*, *V. navarrensis*, and *V. qinghaiensis* (Table 1) were not dominant in their respective native SIs. Taken together, these results indicate that, for these small set of genomes analyzed, ARG cassettes linked to the prevalent VXRs are likely endogenous to members of the Cholera clade, whereas they appear to be exogenous acquisitions in other *Vibrio* species.

When restricting the analysis to ARG cassettes from non-*V. cholerae* genomes only (155 cassettes from 142 genomes), we found that all previously defined VXR types (set 1; Table 1) remained represented, indicating that their prevalence is not driven by the numerical overrepresentation of *V. cholerae* genomes. In addition, a distinct group of VXR sequences (set 2; Table 2) was detected sharing among them $\geq 89\%$ nucleotide identity, and exhibiting approximately 74% nucleotide identity to set 1, consistent with a more pronounced evolutionary divergence between these repeat families (Fig. 2). In fact, almost all sequences from set 2 were identified in *V. anguillarum* (Table 2), which suggests a specificity of these VXRs. The analysis of the prevalence of these repeat types in their native SIs showed that only VXRs of *V. anguillarum* (seq5) and *V. qinghaiensis* (seq6) were the most prevalent in their SIs. Indeed, both species are closely-related in the species phylogeny (Fig. 1).

Table 2
Origin and count of the most prevalent VXR sequences associated with ARC

Genome origin	# ARG cassettes	VXR sequence
GCA_002216685.1_V.tarraie_seq4	43/155	GCCTAACAAACGCCTCAAGATGGACTGTCAACGCGTGGCGTTTCCAGTCCCAATGAGCCGCGGT
GCA_000969235.1_V.cholerae_seq8	47/155	ACCTAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCTATTGAGCCGCGGT
GCA_029635775.1_V.misgurnus_seq8	47/155	GCTATAACAAATGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAACCGCGG
GCA_002211505.1_V.anguillarum_seq6	56/155	ATTTAACAAACGCCTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCAATGAGCCGCGGT
GCA_001558475.2_V.mimicus_seq11	57/155	GGCTAACAAACGCCTCAAGAGGGACTGTCAACGCGCGGCGTTTCCAGTCCCATTGAGCCTCGGT
GCA_000969235.1_V.cholerae_seq2	59/155	TCCTAACAAACGGCTCAAGAGGGATTGCCAACGCGTGGCGGCTTTAGTACAAAATTTAAATTCAG
GCA_001558475.2_V.mimicus_seq5	59/155	ACATAACAAACGGCTCAAGAGGGATTGCCAACGCGTGGCGACTTTAGTGCAAAATTTAGTTTCAG
GCA_002211505.1_V.anguillarum_seq5	59/155	AGCTAACAAACGGCTCAAGAGGGATTGCCAACGCGTGGCGACTTTAGTGCAAAATTGAAGTTTCAG

Genome origin	# ARG cassettes	VXR sequence
GCA_002216685.1_V.tarriae_seq7	59/155	GCCTAACAAACGGCTCAAGAGGGATTGCCAACGCGTGGCGACTTTAGTGCAAAATTGAAGTTCAG
GCA_024347295.1_V.artabrorum_seq10	59/155	AACTAACAAACGGCTCAAGAGGGATTGCCAACGCTTGGCGACTTTAGTACAAAATTGAAGCCCAG
GCA_051352415.1_V.cidicii_seq9	59/155	GCATAACAAACGGCTCAAGAGGGATTGCCAACGCGTGGCGACTTTAGCGCAAAATTGAAGTTCA
GCA_002257545.1_V.qinghaiensis_seq6	61/155	GCATAACAAACGGCTCAAGAGGGATTGCCAACGCTTGGCGGTTTTAGTGCAAAATTGCAGTTCAG
GCA_002257545.1_V.qinghaiensis_seq14	61/155	ACCTAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_009665395.1_V.cincinnatiensis_seq11	61/155	CTATAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_051352415.1_V.cidicii_seq4	61/155	AGCTAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1
GCA_900538065.1_V.paracholerae_seq2	61/155	GCATAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1

Genome origin	# ARG cassettes	VXR sequence
GCA_009665255.1_V.metoecus_seq7	62/155	GCATAACAAACGCCTCAAGAGGGACTGTCAACGCGTGGCGTTTCCAGTCCCATTGAGCCGCGG1

Sequences in bold correspond to the set 1, also identified in *V. cholerae*.

3.3. Diversity of ARGs carried by SIs

The 8,717 ARG/SIs comprised five genes *dfr*, *cat*, *qnr*, *bla*_{CARB} and *vat*, though allelic diversity was observed (Table 3). These genes were distributed unevenly in the species, with *vatF* only in two; *dfrA*, *cat* and *bla*_{CARB} in three; and *qnrVC* in nine species. In particular, *catB9* had an overwhelming dominance (7,703/8,717; 88%), however, most were from *V. cholerae* (n = 7,700). For some of these genes, their cladograms (Supplementary Material 2, Figures S3A-E) showed polyphyly relative to the species tree (Fig. 1), providing another evidence of horizontal movement of SI-associated ARGs among *Vibrio* species. Indeed, some of these species (Table 3) belongs to distinct ecological niches.

Table 3
ARG diversity in *Vibrio* SIs

Gene (count)	Alleles-like (count)	Species (count)	Most prevalent VXR associated (count)
<i>dfrA</i> (56)	-6 (32); -31 (24)	<i>V. cholerae</i> (38), <i>V. metschnikovii</i> (8), <i>V. parahaemolyticus</i> (10)	GCA_000969235.1_V.cholerae_seq1 (27)
<i>bla</i> _{CARB} (330)	-7 (315); -9 (15)	<i>V. cholerae</i> (307), <i>V. metoecus</i> (4), <i>V. paracholerae</i> (19)	GCA_051352415.1_V.cidicii_seq4 (328)
<i>catB9</i> (7,703)	-9 (7,703)	<i>V. cholerae</i> (7,700), <i>V. navarrensis</i> (1), <i>V. parahaemolyticus</i> (1), <i>Vibrio</i> sp. N03-D2	GCA_900538065.1_V.paracholerae_seq2 (7,691)
<i>qnrVC</i> (622)	-4 (455); -5 (29); -6 (74); -7 (64)	<i>V. anguillarum</i> (60), <i>V. cholerae</i> (517), <i>V. mimicus</i> (20), <i>V. fujianensis</i> (1), <i>V. metoecus</i> (3), <i>V. misgurnos</i> (1), <i>V. paracholerae</i> (2), <i>V. parahaemolyticus</i> (12), <i>V. vulnificus</i> (5), <i>Vibrio</i> sp. 1151_11	GCA_009665255.1_V.metoecus_seq7 (539)
<i>vatF</i> (6)		<i>V. metschnikovii</i> (1), <i>V. paracholerae</i> (5)	GCA_900538065.1_V.paracholerae_seq2 (6)

3.4. Functional analyses of a SI-associated ARG

To assess the functionality of an ARG cassette within SI, we examined a *V. mimicus* strain VM343. Using long-read genome sequencing, we confirmed that the *qnrVC4* gene is located within the VM343 SI, positioned 135 cassette genes downstream of *intI4* (approximately 103 kb apart). The VXR sequence flanking this gene exhibited high similarity to the 11 VXRs inferred in this study (Table 1). This *qnrVC* allele was cloned and expressed in a heterologous system to evaluate its functionality. The resulting transformants displayed a ciprofloxacin MIC of 0.38 µg/mL, representing a nearly 8-fold increase compared to the control strain (0.05 µg/mL), a resistance increase typical of *qnrVC* alleles (Miranda et al. 2022), thereby confirming the expression and activity of the cassette.

Discussion

Superintegrons have long been recognized as major reservoirs of gene cassette diversity, including ARGs and antiphage defense systems (Rowe-Magnus et al. 2001; Darracq et al. 2025; Righi et al. 2026). However, their contribution to the evolution and dissemination of ARGs across *Vibrio* species remains poorly understood. By combining VXR profiling, large-scale BLAST-based mapping of cassette boundaries, and phylogenetic analyses across thousands of *Vibrio* genomes, our study provides an overview of SI-associated ARGs within the genus. Despite the extensive diversity of VXRs identified, ARG cassettes were predominantly associated with a small subset of highly conserved VXR types, most of which were enriched in members of the Cholera clade, suggesting preferential mobilization through specific VXR backbones.

SI attC sites are known to be highly homogeneous (Mazel 2006), and such conservation would be expected to restrict cassette recruitment between divergent SIs, ultimately resulting in species-specific VXR repertoires (Rowe-Magnus et al. 2001; Mazel 2006; Debatisse et al. 2025). However, our broad comparative analysis revealed marked incongruence between the VXR cladogram and species phylogeny, with clusters of highly similar VXR sequences shared among phylogenetically distant *Vibrio* species. These patterns are consistent with interspecies cassette exchange. The sharing of highly similar VXR sequences across divergent species suggests that these sites may retain structural compatibility with multiple *IntI4* integrases, thereby facilitating

interspecies cassette dissemination among SIs. Indeed, Vit et al. (2021) demonstrated that integrase recognition of attC sites depends primarily on secondary structure rather than primary nucleotide sequence.

Vibrio species carrying ARGs within their SIs were identified across multiple One Health compartments, even in contexts apparently lacking antibiotic selective pressure, such as the presence of ARGs in isolates from the pre-antibiotic era. These observations reinforce the role of SIs as long-term reservoirs of adaptive genes (Gillings 2014). Moreover, given the promiscuous attC recognition activity of Int1, the integrase associated with mobile class 1 integrons, SIs are thought to represent ancestral sources of ARG cassettes subsequently recruited into mobile integrons (Mazel 2006). Indeed, some ARG cassettes identified here have also been reported in class 1 integrons, often as distinct allelic variants. Examples include *bla*_{CARB} in *Pseudomonas aeruginosa* (AY913771) and *Salmonella enterica* (AY123251); *catB* in *Acinetobacter bereziniae* (EF125010) and *Klebsiella pneumoniae* (FJ594766); *dfrA* in *Bacteroides fragilis* (KF188200) and *K. pneumoniae* (CP003225); *qnrVC* in *Aeromonas punctata* (GQ891757) and *V. cholerae* (EU436855).

Interestingly, by comparing the prevalence of ARG-associated VXR types with their abundance within native SI arrays, we were able to infer potential directional patterns of ARG dissemination. Specifically, the predominant ARG-associated VXR types were also highly abundant within the SIs of members of the Cholera clade, suggesting that these repeats, and likely their associated ARG cassettes, are endogenous to this lineage. In contrast, the same VXR types appeared only sporadically in the SIs of other *Vibrio* species, consistent with horizontal acquisition. Together, these findings support the hypothesis that members of the Cholera clade may act as major sources of SI-associated ARG dissemination across the genus. In fact, Righi et al. (2026) demonstrated cross-species transfer of SI cassettes from *Vibrio furnissii* and *V. parahaemolyticus* to *V. cholerae*, providing direct evidence that this phenomenon can occur naturally. Our findings extend these observations to a genus-wide evolutionary scale and support the hypothesis that horizontal transfer contributes to the long-term dissemination of SI-associated adaptive genes across *Vibrio* species. Moreover, our functional analyses demonstrated that an ARG cassette carried within the SI of a *V. mimicus* strain remains fully functional.

Future studies investigating non-ARG-associated SI cassettes, such as the antiphage defense systems (Darracq et al. 2025), will help determine whether the cross-species VXR patterns observed here are specific to ARG-associated cassettes or instead reflect broader structural and evolutionary constraints intrinsic to integron biology.

Declarations

CRedit authorship contribution statement

Sergio Morgado: Methodology, Formal analysis, Writing - Original Draft, Writing - Review & Editing. Érica L. Fonseca: Investigation, Writing - Original Draft, Writing - Review & Editing. Ana Carolina Vicente: Conceptualization, Writing - Original Draft, Writing - Review & Editing. All authors have read and approved the manuscript.

Data availability

The authors confirm that the data supporting the findings of this study are available within the article or its supplementary materials.

Competing Interests

None to declare.

Funding

This study was financed by FAPERJ - Fundação Carlos Chagas Filho de Amparo à

Pesquisa do Estado do Rio de Janeiro, Processo SEI-260003/019688/2022.

References

1. Darracq B, Littner E, Brunie M, et al. (2025) Sedentary chromosomal integrons as biobanks of bacterial antiphage defense systems. *Science* 388(6747):eads0768. doi:10.1126/science.ads0768
2. Debatisse K, Brunie M, Darracq B, et al. (2025) Bacterial natural transformation drives cassette shuffling and simplifies recombination in chromosomal integrons. *Nucleic Acids Res* 53(21):gkaf1162. doi:10.1093/nar/gkaf1162
3. Fluit AC, Schmitz FJ (2004) Resistance integrons and super-integrons. *Clin Microbiol Infect* 10(4):272-288. doi:10.1111/j.1198-743X.2004.00858.x
4. Fonseca EL, Dos Santos Freitas F, Vieira VV, Vicente AC (2008) New qnr gene cassettes associated with superintegron repeats in *Vibrio cholerae* O1. *Emerg Infect Dis* 14(7):1129-1131. doi:10.3201/eid1407.080132
5. Gillings MR (2014) Integrons: past, present, and future. *Microbiol Mol Biol Rev* 78(2):257-277. doi:10.1128/MMBR.00056-13
6. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22(13):1658-1659. doi:10.1093/bioinformatics/btl158
7. Mazel D (2006) Integrons: agents of bacterial evolution. *Nat Rev Microbiol* 4(8):608-620. doi:10.1038/nrmicro1462

8. Minh BQ, Schmidt HA, Chernomor O, et al. (2020) IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol Biol Evol* 37(5):1530-1534. doi:10.1093/molbev/msaa015
9. Miranda CD, Concha C, Godoy FA, Lee MR (2022) Aquatic Environments as Hotspots of Transferable Low-Level Quinolone Resistance and Their Potential Contribution to High-Level Quinolone Resistance. *Antibiotics (Basel)* 11(11):1487. doi:10.3390/antibiotics11111487
10. Morgado SM, Dos Santos Freitas F, Lourenço da Fonseca E, Vicente ACP (2024) *Vibrio mimicus* Lineage Carrying Cholera Toxin and *Vibrio* Pathogenicity Island, United States and China. *Emerg Infect Dis* 30(8):1729-1732. doi:10.3201/eid3008.240252
11. Néron B, Littner E, Haudiquet M, Perrin A, Cury J, Rocha EPC (2022) IntegronFinder 2.0: Identification and Analysis of Integrons across Bacteria, with a Focus on Antibiotic Resistance in *Klebsiella*. *Microorganisms* 10(4):700. doi:10.3390/microorganisms10040700
12. Page AJ, Cummins CA, Hunt M, et al. (2015) Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31(22):3691-3693. doi:10.1093/bioinformatics/btv421
13. Page AJ, Taylor B, Delaney AJ, et al. (2016) SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genom* 2(4):e000056. doi:10.1099/mgen.0.000056
14. Rapa RA, Labbate M (2013) The function of integron-associated gene cassettes in *Vibrio* species: the tip of the iceberg. *Front Microbiol* 4:385. doi:10.3389/fmicb.2013.00385
15. Righi L, Stutzmann S, Bader L, Lemopoulos A, Blokesch M (2026) Competence-mediated DNA uptake diversifies *Vibrio cholerae* sedentary chromosomal integrons. *Science* 392(6794):194-201. doi:10.1126/science.aed0645
16. Rowe-Magnus DA, Guerout AM, Mazel D (2002) Bacterial resistance evolution by recruitment of super-integron gene cassettes. *Mol Microbiol* 43(6):1657-1669. doi:10.1046/j.1365-2958.2002.02861.x
17. Rowe-Magnus DA, Mazel D (2001) Integrons: natural tools for bacterial genome evolution. *Curr Opin Microbiol* 4(5):565-569. doi:10.1016/s1369-5274(00)00252-6
18. Vit C, Richard E, Fournes F, et al. (2021) Cassette recruitment in the chromosomal Integron of *Vibrio cholerae*. *Nucleic Acids Res* 49(10):5654-5670. doi:10.1093/nar/gkab412

Figures

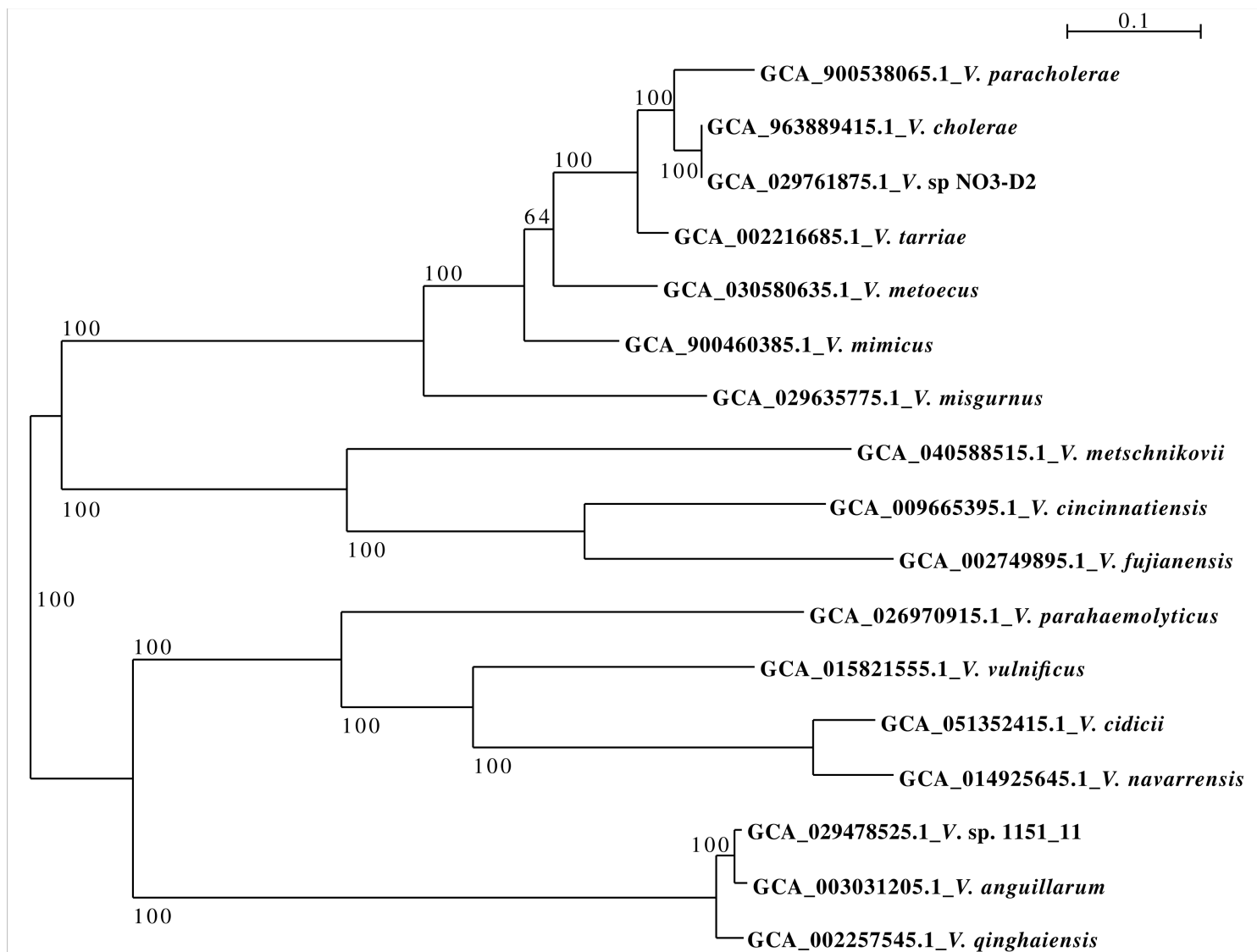


Figure 1

Maximum-likelihood phylogenomic tree of 17 *Vibrio* genomes representing species that harbor ARGs within their SIs.

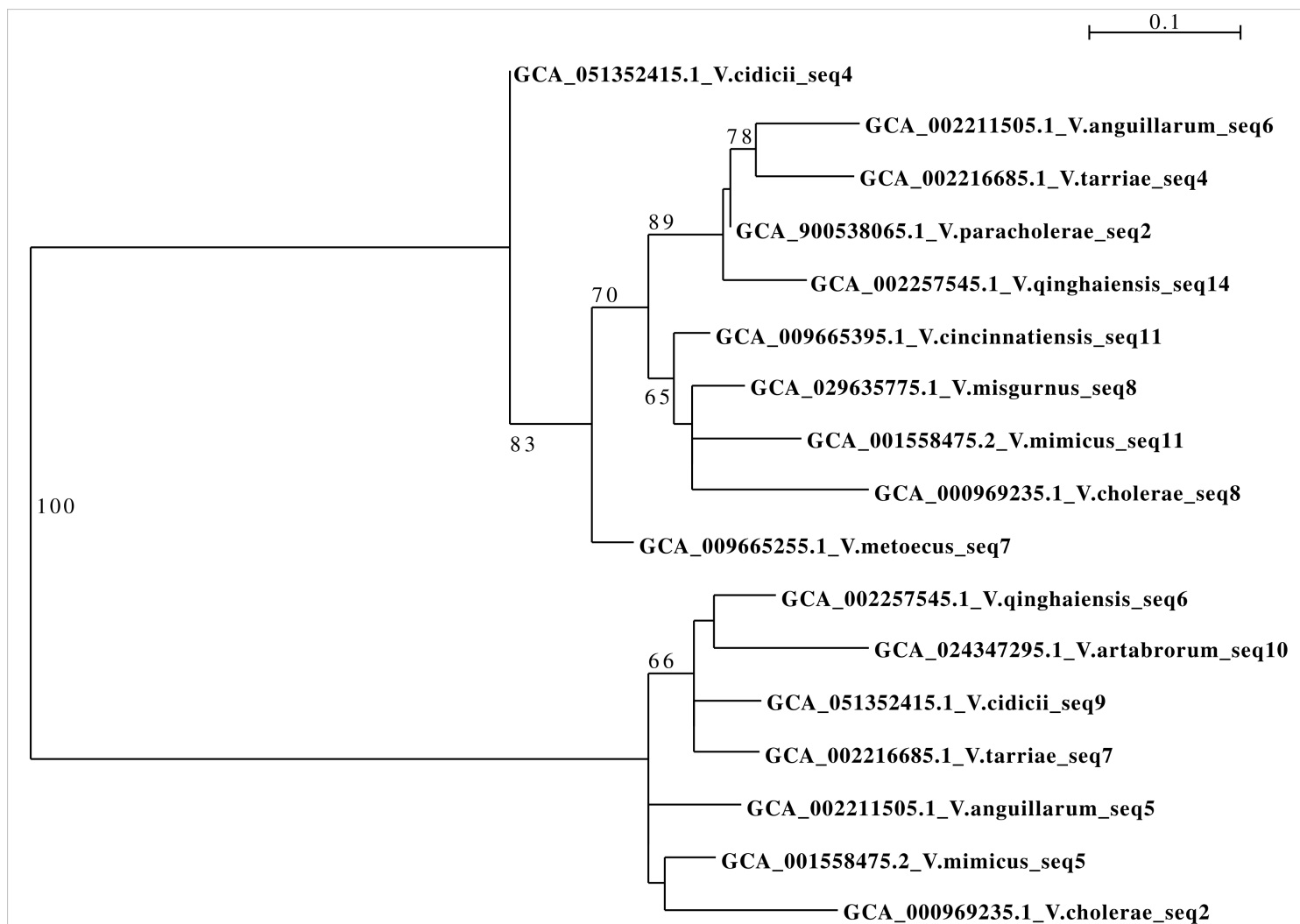


Figure 2

Maximum-likelihood cladogram of the most prevalent VXR sequences associated with ARG cassettes of SIs, including VXRs from set 1 (upper clade) and set 2 (lower clade).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial1.xls](#)
- [Supplementarymaterial2.pdf](#)