



Novel configurations of type I-E CRISPR-Cas system in *Corynebacterium striatum* clinical isolates

Juliana Nunes Ramos¹ · Paulo Victor Pereira Baio² · João Flávio Carneiro Veras³ · Érica Miranda Damásio Vieira³ · Ana Luiza Mattos-Guaraldi¹ · Verônica Viana Vieira³

Received: 12 April 2022 / Accepted: 21 November 2022 / Published online: 7 December 2022
© The Author(s) 2022

Abstract

Clustered regularly interspaced short palindromic repeats (CRISPR) are a prokaryotic adaptive immune system that, through Cas proteins, promote the degradation of foreign nucleic acids such as phages and plasmids. We analyzed 10 genomes of *Corynebacterium striatum* clinical isolates from a public hospital in Rio de Janeiro, Brazil, the most emergent multidrug-resistant *Corynebacterium* species. All isolates were submitted to antimicrobial susceptibility testing. The occurrence and diversity of the CRISPR system were investigated by bioinformatics tools. Our analysis revealed that the isolates exhibited type I-E gene arrangements, and 3 more multidrug-resistant isolates, alternative type I-E gene arrangements, showing a divergent gene arrangement within the *cas* operon. Phylogenetic analysis of the *cas1* gene of this type I-E CRISPR-Cas system alternative arrangement, termed here type I-E', showed a cluster in a distinct clade of the type I-E CRISPR-Cas system. The systems' guanine-cytosine (GC) content is lower than the genomic DNA's GC content, and mobile genetic elements were found in some isolates near the CRISPR-Cas system. Most CRISPR spacers are unknown indicating that there is a reservoir of unexplored corynebacteriophages and plasmids. Some spacers showed perfect homologies with phage and plasmid sequences. Intact phage regions were found in 3 of our isolates, ranging from 9.1 to 43.8 kb, with regions showing similarity to *Rhodococcus* and *Corynebacterium* phages. Our results may contribute to research about the CRISPR-Cas system diversity in *C. striatum*, where there are no published data to date.

Keywords CRISPR-Cas · *Corynebacterium striatum* · Multidrug-resistant · Health care infections · Emergent pathogen

Juliana Nunes Ramos and Paulo Victor Pereira Baio contributed equally as the first authors of this manuscript.

Responsible Editor: Luis Nero

✉ Juliana Nunes Ramos
jnr.uerj@gmail.com

- ¹ Laboratório de Difteria E Corinebactérias de Importância Clínica (LDCIC), Faculdade de Ciências Médicas, Universidade Do Estado Do Rio de Janeiro, Av. 28 de Setembro, 87, Fundos, 3º Andar, Vila Isabel, Rio de Janeiro, RJ, Brazil
- ² Laboratório Químico-Farmacêutico Do Exército Brasileiro (LQFEx), Ministério da Defesa, Brasília, Brazil
- ³ Laboratório Interdisciplinar de Pesquisas Médicas (LIPMED), Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil

Introduction

Corynebacterium species are widely distributed in the environment and are part of the skin and mucosal microbiota [1, 2]. *Corynebacterium striatum* is a Gram-positive rod, non-sporulating recognized as a true pathogen in specific circumstances when isolated from patients with chronic diseases, indwelling medical devices, and several samples from sterile body sites [3–5]. *C. striatum* has been cited in several reports as a multidrug-resistant (MDR) health care infection pathogen, including septicemia, valvular damage, pulmonary infection, meningitis, endocarditis, osteomyelitis, and other invasive infections [1, 3, 6, 7]. Clonal MDR *C. striatum* has been affecting both immunocompromised and immunocompetent patients [1, 3, 6]. Multidrug resistance and patient-to-patient transmission have also been reported [3, 6].

The prokaryotic CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) system resembles an immunity system to host cells against invading nucleic acids and they represent a barrier to recombination. This system has been detected in bacteria and archaea. CRISPR-associated proteins (Cas) consist of a combination of Cas effector proteins that are involved in target DNA or RNA cleavage, CRISPR loci transcript processing, and novel spacer integration [8–11]. The CRISPR system uses mechanisms divided into three stages, the first being the adaptation phase, when a novel spacer is acquired from an invading nucleic acid. At this stage, there are Cas1 and Cas2 that form a complex involved in the adaptation step and are conserved in almost all CRISPR-Cas types. The second stage is the expression phase, where the CRISPR locus is transcribed into pre-crRNA, and *cas* genes are expressed and involved in pre-crRNA processing and in mature short crRNAs that form a complex with these genes. The last step is the interference phase, where invading nucleic acids will be recognized and cleaved by crRNA-Cas complexes [12, 13].

The CRISPR-Cas system prevents the spread of plasmids and bacteriophages and limits horizontal gene transfer by these mobile genetic elements [14]. In many bacterial species, antibiotic resistance is mediated by the acquisition of genes from plasmids and transposons. The presence of the CRISPR-Cas system and the acquisition of antibiotic resistance genes may have an inverse correlation. Studies have shown a significant association between antibiotic resistance in *Enterococcus faecalis* and the absence of the CRISPR-Cas system [15]. However, in a study performed on *Escherichia*

coli, CRISPR-Cas system appears to ineffectively block plasmid dissemination and antibiotic resistance [16].

CRISPR and associated *cas* genes have been detected in some *Corynebacterium* species, such as *Corynebacterium diphtheriae* [8], *Corynebacterium bovis* [17], *Corynebacterium pseudotuberculosis* [18], *Corynebacterium ulcerans* [19], and *Corynebacterium urealyticum* [20]. In other *Corynebacterium* species, CRISPR-Cas system studies are scarce, except for *Corynebacterium glutamicum*, which is an important metabolite producer in the biotechnology industry [21]. The type I-E CRISPR-Cas system was found in most corynebacteria. In the pathogen that causes diphtheria diseases, *Corynebacterium diphtheriae*, type II-C system and type I-E variant system were found [8]. To date, there are no published data on diversity about the *C. striatum* CRISPR-Cas system, an opportunistic hospital-associated pathogen. In this study, we explored 10 *C. striatum* genomes isolated from a nosocomial outbreak in the city of Rio de Janeiro, Brazil, and 21 *C. striatum* genomes available at NCBI were analyzed for the presence and characterization of the CRISPR-Cas system and their spacers.

Methods

Bacterial isolates

We analyzed 10 *C. striatum* genomes, deposited at NCBI (Table 1), isolated from nosocomial outbreaks that occurred

Table 1 Details of CRISPR-Cas loci in Brazilian *C. striatum* isolates

Isolates ^a	Isolation date ^a	Isolation sites ^a	PFGE profiles ^a	CRISPR-Cas system (number of spacers)	GenBank access number
1954	October 2009	Surgical wound secretion	IV	–	PGGF01000000.1
1961	December 2009	Urine	III	I-E (15)	LAYR01000000.1
2023	August 2009	Blood	I	I-E (18) I-E' (9)	LBCN01000000.1
2038	September 2009	Blood	II	I-E (51) I-E' (59)	PGGG00000000.1
2130	August 2010	Blood	V	I-E (20)	NRIL00000000.1
2230	February 2011	Blood	VI	I-E (44) I-E (47)	LTBF00000000.1
2237	March 2011	Blood	VI	I-E (47)	NRIM01000000.1
2296	July 2011	Central venous catheter	VII	I-E (100)	NRIN00000000.1
2308	August 2011	Blood	I	I-E (117) I-E' (58)	NRIO00000000.1
2425	August 2012	Blood	IX	I-E (55)	NRIP00000000.1

^aAccording to Ramos et al. [4]

at a public university hospital in the city of Rio de Janeiro, Brazil, for 42 months (January 2009 to February 2013). In addition to the 10 genomes, we have included 21 *C. striatum* genomes, available at GenBank/NCBI, for comparative purposes (Supplementary Table 1).

Antimicrobial susceptibility testing

Antimicrobial susceptibility profiles were determined by minimum inhibitory concentration (MIC) using Etest strips on cation-adjusted Mueller–Hinton agar supplemented with 5% sheep blood using inoculum equivalent to a 0.5 McFarland standard. Interpretation of penicillin, ciprofloxacin, moxifloxacin, vancomycin, clindamycin, tetracycline, linezolid, and rifampin values was performed according to the current BrCAST/EuCAST guideline [22] (Table 2). MDR was defined as acquired non-susceptibility to at least one agent in three or more antimicrobial categories [23].

Characterization of CRISPR-Cas diversity

CRISPRCasFinder¹ (version CRISPR-Cas + + 1.1.2) was used to identify the CRISPR-Cas system of genomes. CRISPR arrays with low evidence, equal to 0 or 1, were not included in the analyses [24]. CRISPR-Cas cassette type was determined following the nomenclature and classification described by Makarova et al. [11]. CRISPR array spacers were extracted from CRISPRFinder outputs. Spacer sequences were analyzed for their identity: at the CRISPRTarget² database, which contained A Classification of Mobile Genetic Elements (ACLAME), Genbank-Phage, RefSeq-Plasmid, RefSeq-Viral, Island-Viewer, PHAST and Community Cyber infrastructure for Advanced Microbial Ecology Research & Analysis (CAMERA) sequences, and the cut-off score was the default parameter value [25]; and at the ViroBLAST³ server, version 2.6 against viral databases using default parameters [26] and against spacers databases in the CRISPR-Cas + + database¹ with E-value = 0.01 [24]. Spacer hits were selected from the ViroBlast, CRISPRTarget, and CRISPR-Cas + + databases with a cut-off identity cover according to Sangal et al. [8]. Direct repeat conservation was represented by WebLogo,⁴ version 2.82

[27]. Prophage sequences were identified from assembled contigs using the PHASTER⁵ webserver (PHAge Search Tool Enhanced Release) [28].

Core genome and phylogenetic analysis

Prokka,⁶ version 1.14.6, was used for whole-genome annotation to produce standards-compliant GFF3 output files required for pangenome calculation [29]. The annotated gene repertoires of the studied genomes were grouped using Roary,⁷ version 3.13.3 using the parameters as follows: –mafft; –i 60 to calculate pan-genome and core-genome [30]. A neighbor-joining (NJ) phylogenetic tree was constructed from *C. striatum* core genome sequence alignment using Mega X⁸ [31]. NJ trees were also generated from nucleotide sequence alignments of *cas* genes and direct repeats consensus of CRISPR-Cas systems of all isolates by p-distance with 500 iterations for bootstrap. Geneious,⁹ version 2021.2.2, was used to assess Cas protein conservation by multiple sequence global alignment (Needleman-Wunsch) with standard parameters.

Results

Antimicrobial resistance profiles

Table 2 shows antimicrobial susceptibility profiles of 10 *C. striatum* isolates. Eight *C. striatum* isolates showed non-susceptibility to at least one agent in three or more antimicrobial categories and were identified as MDR pathogens. Isolates no. 1954 and 1961 were susceptible to all antimicrobials tested and showed intermediate susceptibility to ciprofloxacin, according to the BrCAST/EuCAST guideline [22]. Isolates no. 2023, 2038, and 2308 have an additional CRISPR-Cas system, termed here type I-E' CRISPR-Cas, and were susceptible only to vancomycin, tetracycline, and linezolid. Our isolates showed resistance to penicillin (70%), ciprofloxacin (100%), moxifloxacin (80%), clindamycin (80%), and rifampin (30%). All isolates were susceptible to vancomycin, linezolid, and tetracycline.

¹ <https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index>

² http://crispr.otago.ac.nz/CRISPRTarget/crispr_analysis.html

³ <https://indra.mullins.microbiol.washington.edu/viroblast/viroblast.php>

⁴ <https://weblogo.berkeley.edu/logo.cgi>

⁵ <http://phaster.ca/>

⁶ <https://github.com/tseemann/prokka/issues>

⁷ <https://sanger-pathogens.github.io/Roary/>

⁸ <https://www.megasoftware.net/>

⁹ <https://www.geneious.com>

Table 2 Minimum inhibitory concentration and antimicrobial susceptibility profiles of 10 *Corynebacterium striatum* isolates from blood ($n=7$), central venous catheter ($n=1$), urine ($n=1$), and surgical wound secretion ($n=1$) infections

Antimicrobial agents	Isolates									
	1954	1961	2130	2230	2237	2296	2425	2023	2038	2308
Penicillin	0.12	0.03	0.5	2-4	2-4	0.5	0.12	>256	>256	>256
Vancomycin	0.25	0.5	0.5	0.5	0.5	0.5	0.5	0.25	0.25	0.5
Ciprofloxacin	0.094	0.125	2	> 32	> 32	> 32	>32	>32	>32	>32
Moxifloxacin	0.125	0.125	0.75	8	4-6	> 32	>32	8	6	6
Tetracycline	2	2	2	2	1	0.5	0.5	1	1	1
Clindamycin	0.12	0.12	>256	> 256	> 256	> 256	> 256	>256	>256	>256
Rifampin	<0.002	<0.002	0.003	0.002	0.002	<0.002	<0.002	>32	>32	>32
Linezolid	0.38	0.094	0.38-0.5	0.19	0.38	0.25	0.064	0.25	0.75	0.25
	Negative CRISPR-Cas (n=1)	Type I-E CRISPR-Cas (n=6)						Type I-E CRISPR-Cas plus type I-E' CRISPR-Cas		

According to the BrCAST guideline (2021), orange is resistant, blue is intermediate, and green is susceptible to antimicrobial agents

Isolation sites: surgical wound secretion (1954); urine (1961); blood (2023, 2038, 2130, 2230, 2237, 2308, and 2425); and central venous catheter (2296)

CRISPR-Cas system diversity and arrangement in *C. striatum*

A total of 13 CRISPR-Cas systems among 9 Brazilian *C. striatum* genomes and their features were listed in Table 1. Only multidrug-susceptible (MDS) isolate 1954 did not contain CRISPR-Cas system. All systems scored highly accurate (level of evidence = 4; as high as possible) based on parameters from the CRISPRFinder database, which assigns levels of evidence from 1 to 4 for spacer repetition and similarity [24]. The CRISPRFinder server identified type I-E system in all isolates. Four isolates presented two CRISPR-Cas systems, and the CRISPR-Cas system arrangement in three *C. striatum* isolates showed differences (Fig. 1a and b). All three MDR *C. striatum* isolates susceptible only to vancomycin, tetracycline, and linezolid present a novel type I-E system configuration, exhibiting a divergent gene arrangement within the *cas* operon, termed here as type I-E' (Fig. 1b).

C. striatum draft or complete genomes from other countries available in the GenBank/NCBI database, including

MDR 215 and 216 complete genomes [3], were used to compare with our isolates. Three draft genomes were excluded from the analysis, for presenting evidence level zero or 1 in the CRISPRFinder database. Furthermore, LK37 and 1329-caur draft genomes were not used because the CRISPR-Cas system was divided into distinct contigs. Thus, it was possible to observe that 6 *C. striatum* genomes from the GenBank/NCBI database also presented the type I-E' system (Supplementary Table 1).

To further distinguish the CRISPR-Cas system in *C. striatum*, the *casI* gene's NJ tree was constructed. Results showed that all *casI* genes of the type I-E CRISPR system formed a separated branch from those of the type I-E' CRISPR system (Fig. 2). Moreover, the amino acid identities of all *cas* proteins of Brazilian *C. striatum* genomes for types I-E ($n=10$) and I-E' ($n=3$) CRISPR systems were detected. Within a single type I-E or I-E', most CasI proteins share more than 99% amino acid identity and some up to 100% amino acid identity. When amino acid sequences of type I-E CasI proteins were compared with those of type I-E', they shared 30% of amino acid identity (Table 3).

Fig. 1 Type I-E CRISPR-Cas system (a) classic configuration and alternative gene arrangement named type I-E' (b) found in Brazilian *C. striatum* isolates. The similarity scale between the *C. striatum* isolate CRISPR-Cas systems is represented below, ranging from 0 to 100% (white to black). The numbers in the red box indicate the quantity of spacer sequences. Isolates bearing the contig number have two CRISPR-Cas systems

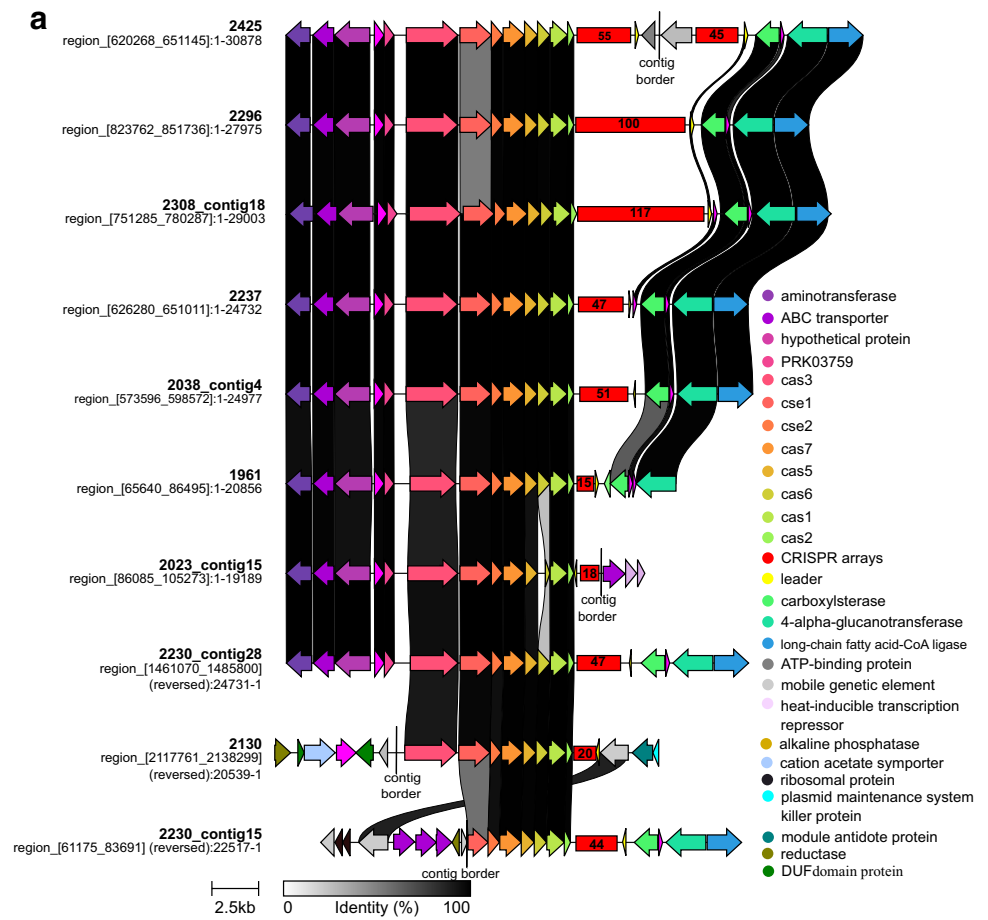


Fig. 1 (continued)

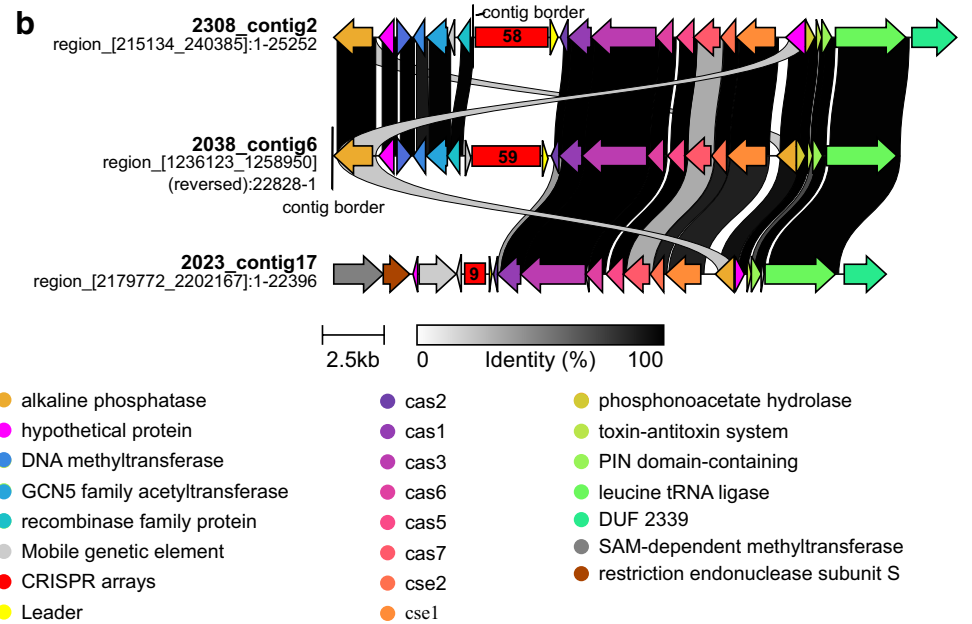


Fig. 2 NJ tree for *cas1* gene in *C. striatum*. *Cas1* gene alignment was performed using the ClustalW algorithm in BioEdit Sequence Alignment, and the tree was generated in MegaX software. The inferred distance was calculated using p-distance. Bootstrap values (> 60%) based on 500 replicates are shown. 0.05% scale bar estimated sequence divergence. Isolates LK37 and 1329-caur were excluded from this analysis because the CRISPR-Cas system is divided into different contigs. Brazilian isolates contain an asterisk in the name. Type I-E' and type I-E CRISPR-Cas systems were colored blue and black, respectively. FDAARGOS 1054^T corresponds to *C. striatum* type strain

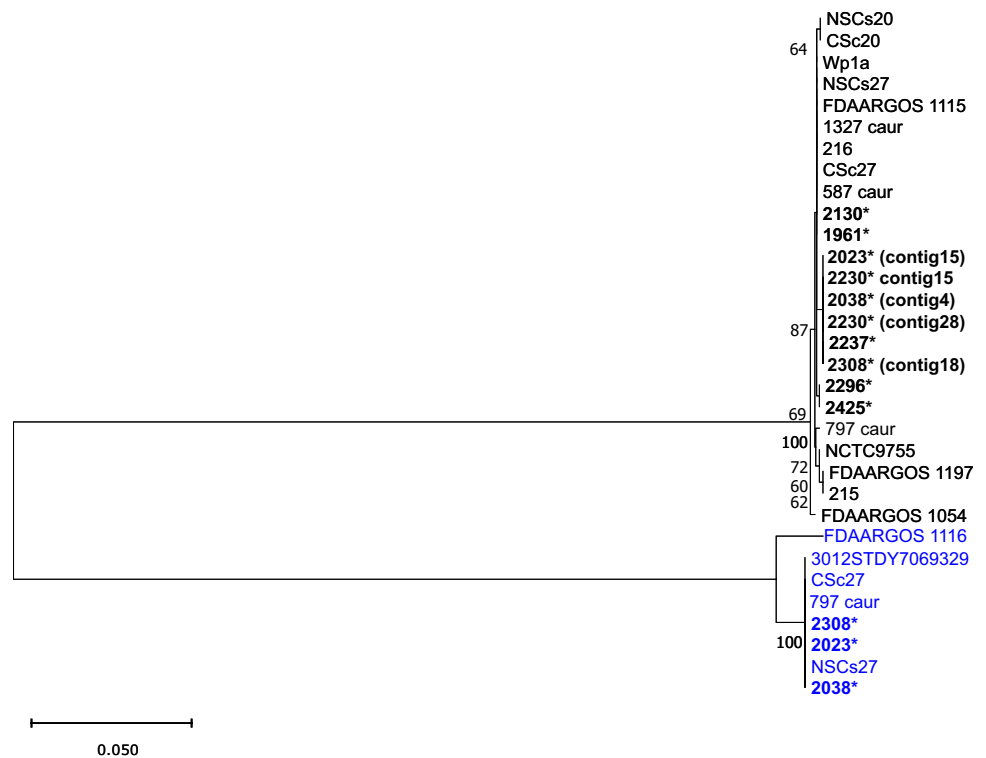


Table 3 Cas protein conservation of Brazilian *Corynebacterium striatum* isolates

CRISPR type	Isolates	% amino acid identity								
		Cas3	Cse1	Cse2	Cas7	Cas5	Cas6	Cas1	Cas2	
Type I-E	1961									
	2023									
	2038									
	2130									
	2230									
	2230									
	2237									
	2296									
	2308									
	2425									
Consensus		99.7	98.6	98.7	99.8	99.6	99.6	99.9	99.8	
Type I-E'	2023									
	2038									
	2308									
Consensus		99.9	100	100	100	100	100	100	100	
% amino acid identity			23.5	19.8	17.9	24.9	19.1	22.5	30.7	28.3

Geneious global alignment (Needleman-Wunsch) was used for multiple sequence alignment. For each protein, global alignment was performed within each type (I-E or I-E'), and consensus similarity between them was given in percentage

The amino acid identity percentage indicates the alignment between type I-E protein consensus sequence and type I-E' protein consensus sequence

As *cas3* was cleaved in isolate 1961, it was not considered to calculate amino acid identity

As *cse1* was very short in isolate 2230 (contig 15) and showed a frameshift in isolate 2296, it was not considered to calculate amino acid identity

As *cas7* showed a frameshift in isolate 2038 (type I-E'), it was not considered to calculate amino acid identity

As *cas2* was broken in isolated 2023 (type I-E'), it was not considered to calculate amino acid identity

Source of spacer sequences

A total of 640 spacer sequences were found in *C. striatum* CRISPR arrays from Brazilian isolates (Supplementary Table 2). The ViroBlast server found no hits for spacers of isolates no. 1961 and 2023. The CRISPRTarget database returned the highest number of hits ($n = 226$) among the databases used. All hits ($n = 16$) found in CRISPRCasdb returned only for CRISPR spacers found in *C. striatum* genomes with the parameters previously described in methods. About 379 spacers (59.2%) are unknown according to the database used: 73 and 306 spacers in type I-E' and I-E systems, respectively.

The CRISPR loci found in *C. striatum* vary in length and spacer content. The longest CRISPR locus contains 117 spacers and was found in the type I-E system of MDR isolate 2308. The smallest CRISPR locus was found in MDR isolate 2023, with 9 spacers. Of the 117 spacers found in isolate 2308, five were duplicated. None of these spacer sequences was found in the additional type I-E' system of this isolate. Forty-nine of the 117 spacer sequences were found in isolates 2230 and 2237, some with a 1 bp difference in length. A significant amount of 117 spacer sequences were found only in the type I-E system of MDR isolates 2023 and 2038, totaling 20 and 45 spacer sequences, respectively.

The 9 Brazilian *C. striatum* isolates shared spacers with 100% similarity (Supplementary Table 3), but between type I-E' and type I-E systems, there was no sharing of spacer sequences. Isolate 2130 was the only one that did not share spacer sequence with other isolates. The type I-E CRISPR system spacer sequences of isolate 2308 were the most shared among the isolates, with 44 of them found in the type I-E CRISPR system of isolates 2038, 2230, and 2237.

Repeat consensus sequences were also different among the two CRISPR system configurations (Fig. 3a and b), while they were conserved within the same layout (Supplementary Fig. 1). Within a single type I-E or I-E', most

repeat consensus shared more than 96.5% nucleotide similarity and some up to 100% nucleotide similarity. When repeat consensus nucleotide sequences of the type I-E system were compared with those of type I-E', they shared 50% nucleotide identity.

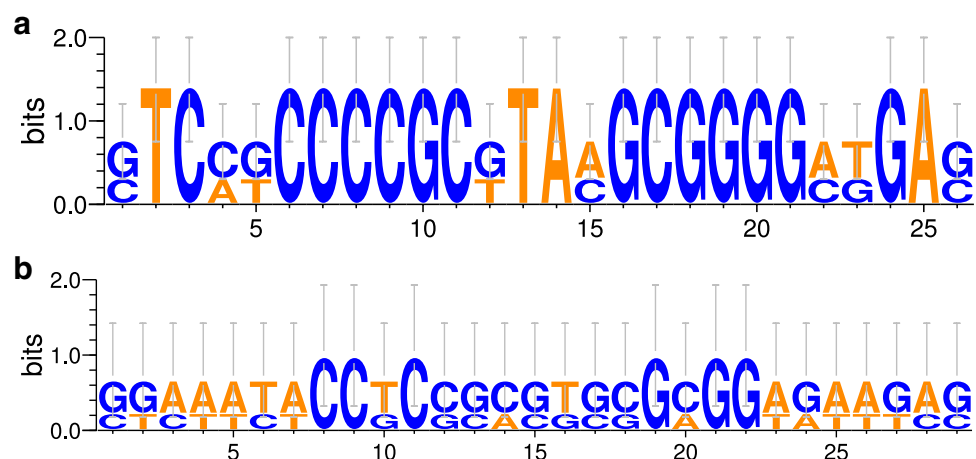
Genetic relationship between *C. striatum* isolates

The *C. striatum* core genome consisted of 1619 coding sequences (CDS), when calculated using 31 genomes. Phylogenetic analysis (Fig. 4) of the conserved core genome was contradictory to the type of CRISPR-Cas systems and phylogeny of *cas1* genes (Fig. 2). Furthermore, by phylogenetic analysis of the core genome, Brazilian MDR isolates were grouped in a distinct clade from MDS isolates 1954 and 1961 and from isolates of other countries (Fig. 4; Supplementary Table 1).

Prophage regions

Isolates were submitted to PHASTER analysis for identification and annotation of prophage sequences. The results revealed the presence of 3 intact phages in isolates 1961, 2130, and 2425. An intact prophage region of 9.1 kb was identified in MDS isolate 1961, with a score of 100% containing 12 proteins. The phage with the highest number of proteins like those identified in this region was Rhodoc_Sleepyhead (GenBank NC048782_2), with 52.22% of GC content. Isolate 2130 also has a 43.2 kb intact prophage region containing 60 proteins, scoring 140%. The phage with the highest number of proteins like those in the region is Rhodoc_Jace (GenBank NC047974_15), with 54.7% of GC content. Finally, isolate 2425 showed an intact prophage region of 43.8 kb containing 64 proteins, with a score of 110%, and the most common phage found was Coryne_Poushou (GenBank NC042139_11), with 57.2% of GC content. When

Fig. 3 Direct repeat conservation in **a** type I-E and **b** type I-E' CRISPR configurations of *C. striatum*. The sequence logo was created by WebLogo 2.8.2. The height of the letters shows the relative frequency of the corresponding nucleotide at that position



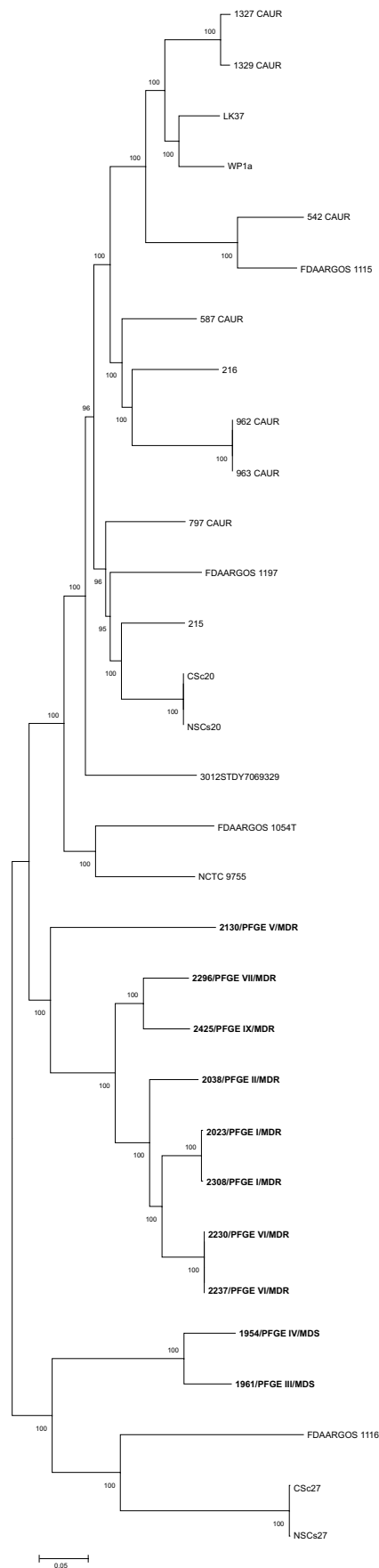


Fig. 4 NJ tree from variation in the core genome (1619 CDS) of 31 *C. striatum* isolates. The scale bar represents the number of replacements per site. Brazilian isolates are highlighted in bold. The PFGE profiles and the strength profile according to Magiorakos et al. [23] were added. FDAARGOS 1054^T corresponds to the strain of type *C. striatum*

submitted to NCBI BLAST using nucleotide collection, the intact prophage sequence of isolate 1961 aligned with a small chromosomal region of *C. striatum* 215, about 7 kb, approximately. Of the 43.2 kb of the intact region of the prophage found in isolate 2130, about 21.7 kb were aligned with the chromosomal region of *C. striatum* strain, with correspondence with proteins commonly found on phages, such as phage tail and phage portal protein. Finally, the intact prophage region of isolate 2425 aligned only 5.4 kb, approximately, with the chromosome region of *C. striatum* DSM45711 (FDAARGOS 1197). The similarity between the spacers and the phages found in isolates 1961, 2130, and 2425 was verified. There was no match between the CRISPR spacers from isolate 1961 and the phage found in the same isolate. Four unknown CRISPR spacers from isolate 2130 matched with the phage found in this isolate. Three CRISPR spacers from isolate 2425 matched with the phage found in this isolate. Spacers 6, 10, and 31 were found in *C. simulans* (100% identity), *Gordonia* phage Easley (84.37% identity), and *Corynebacterium glyciniphilum* (87.50% identity)/Turkey Adenovirus 1 (81.25% identity), respectively.

Discussion

C. striatum isolates expressing different MDR profiles have been identified as the etiologic agent of healthcare-associated infections in several countries [1, 3, 4, 7, 32] demonstrating the involvement of genetic mobile elements in the resistome of this bacterial species. *C. striatum* isolate genomes present antimicrobial resistance genes acquired, disseminated, and conserved by vertical transfer through plasmids and transposons [32]. Bacterial CRISPR-Cas systems are important elements for inhibiting horizontal gene transfer between bacterial cells [8, 33]. However, information on CRISPR-Cas systems in *C. striatum* remains unknown. In this study, we analyzed CRISPR-Cas systems, and all elements for this system (*cas* genes, repeats, leader sequences, and adjacent protospacer motifs) were characterized. In some studies, there was a reverse correlation between the CRISPR-Cas system presence and antibiotic resistance in some species such as *Enterococci* [33], but, in MDR *C. striatum*, such as *E. coli* [34], there was no significant relation.

In this study, we have identified a different configuration of the type I-E CRISPR-Cas system in 3 Brazilian isolates.

To our knowledge, this is the first comprehensive study of the CRISPR-Cas system in *C. striatum*, an important MDR pathogen associated with nosocomial outbreaks in several countries [1, 3, 4, 7].

Previous studies reported the emergence of *C. striatum* isolates as novel clones by PFGE genotyping as etiologic agents of invasive infections [4, 7]. The CRISPR-Cas system was found in 9 Brazilian genomes. Of these, 8 isolates are MDR with different levels of resistance, but resistant to at least one agent in three or more antimicrobial categories [23]. Isolates 2023, 2038, and 2308, classified as PFGE profiles I and II, are the most resistant, susceptible only to tetracycline, vancomycin, and linezolid [4] (Table 2). It is important to emphasize that these 3 more resistant isolates presented an additional system with an alternative arrangement named here type I-E'. Although the resistance profiles of *C. striatum* isolates from other countries used in this study are unknown, some showed the type I-E' CRISPR-Cas system too (Supplementary Table 1). A similar configuration was found in *C. diphtheriae*, the causative agent of diphtheria disease, named type I-Ea [8, 35]. In this case, the CRISPR loci are located between the *cas3* and *cse1* genes.

This study revealed a diversity of spacers among CRISPR arrays of Brazilian *C. striatum* genomes. Of the 640 spacers, 113 were shared among the 9 isolates, being 104 spacers between the type I-E system and 9 spacers between the type I-E' system. There was no sharing of spacers between the two types of systems. CRISPR loci found in our isolates presented a significant amount of spacer sequences ($n = 135$), with similarities ranging from 81.2 to 100% to prophage of *Rhodococcus* (phage Rhodoc REQ3—GenBank NC016654), belonging to family Siphoviridae, isolated from wastewater. This prophage is also found in the *C. striatum* type strain. Some spacers showed similarities above 81.2% with corynebacteriophages or unannotated regions of other *Corynebacterium* species, such as *Corynebacterium aurimucosum*, *Corynebacterium simulans*, and *Corynebacterium ulcerans*. Eleven spacer sequences from our Brazilian isolates, except for isolates 1961, 2296, and 2425, matched with *Corynebacterium* phage phi673 (GenBank NC042354), a lytic phage of *Corynebacterium glutamicum* ATCC 13,032.

Thirteen spacer sequences matched with *C. striatum* 215, 216, and type strain by the CRISPR-Cas + + database, with similarities ranging from 84.3 to 100%. Interestingly, 5 spacers showed similarity with non-corynebacteria plasmids, such as *Bacillus*, *Sinorhizobium*, and *Rhodobacter*, and the last had the same score as the phage Salmon SPN1S spacer. This diversity of spacers may reflect a history of previous invasions, with hits ranging from uncultivable viruses, *Rhodococcus*, *Mycobacterium*, *Gordonia*, and *Corynebacterium* phages to plant pathogenic bacteria plasmids, such as *Ralstonia solanacearum*, although for the majority (59.7%) no

hits were found in the databases, indicating that there is a reservoir of unexplored plasmids and corynebacteriophages.

Isolates 2296 and 2425 shared 3 spacers with 84% similarity to *C. simulans* chromosome region. Moreover, spacer 26 from isolate 2296 had similarity to phage *Propionibacterium* PFR1 found in isolate 2425 (spacer 22) (Supplementary Table 3). Some type I-E' system spacers from isolate 2038 ($n = 5$) and type I-E system spacers from isolates 2130 ($n = 1$), 2296 ($n = 2$), and 2308 ($n = 5$) were duplicated within the CRISPR arrays. This event has also been reported in *Corynebacterium diphtheriae*, in which 9 duplicated spacers were found within CRISPR arrays [8]. The relationship between spacer duplication and increased efficiency of CRISPR-Cas-mediated immunity against invasive DNA is controversial [8].

It is important to emphasize that the total of 640 spacers are those that were present near the CRISPR-Cas system. Except for isolates 2038 and 2296, there are other CRISPR arrays located in other regions of the genomes. Therefore, there may be more than 113 spacers shared between the 9 Brazilian isolates. The presence of mobile genetic elements, such as insertion sequences close to some CRISPR-Cas loci, may indicate that these systems may have been acquired horizontally. This observation is supported by the lower DNA GC content of *cas* operons (53% for both type I-E and type I-E' systems) when compared with the average GC content of the *C. striatum* genome (59.3%). Furthermore, in 6 isolates, the *cas* operons in both types I-E and I-E' systems are close to transposases of families 21, 30, and 481 (Fig. 1a and b). The IS21 element is carried by plasmid IncP R68 and is close to the kanamycin resistance gene therein [36]. IS30 works on a structure of compound transposons and has been identified as part of compound transposons that flank the colistin resistance gene *mcr-1* [37]. The IS481 family transposase actively mobilizes the TnRErm46 transposable element that contains the *erm* gene responsible for emerging macrolide resistance [38]. According to Sangal et al. [8], the type I-E system is commonly flanked by mobile genetic elements, indicating the potential mobility of this system between isolates.

According to Ramos et al. [4], Brazilian isolates 2230 and 2237 are clones belonging to PFGE profile VI and were isolated from the same patient with an interval of 1 month in 2011 (Table 1). However, both isolates showed differences in the number of CRISPR-Cas systems. Isolate 2230 presented two type I-E CRISPR-Cas systems, and isolate 2237, only one type I-E CRISPR-Cas system. One of the CRISPR-Cas systems of isolate 2230 has 47 spacer sequences, as well as the number of spacers found in the CRISPR-Cas system of isolate 2237. The 47 spacer sequences found in both clones of the same PFGE profile are similar to each other (Supplementary Table 3).

The leader sequence is adjacent to the first spacer of the CRISPR locus and acts as a promoter for locus transcription and as a guide for novel spacer incorporation [39, 40]. This is a region rich in AT (adenine and thymine) and in Brazilian isolates that ranged from 44 to 58%, and no difference between %AT of the leader sequence of the type I-E system and the type I-E' system was observed. The leader sequence of isolates 2296, 2308 (contig 18) and isolate 2425 shared 100% similarity (Fig. 1a) as well as the leader sequence of two CRISPR-Cas systems of isolate 2237.

The core genome is defined as the content of genes present in all representatives of a species [41]. Core genome analysis was performed to verify if there was a correlation between the CRISPR system type and the core genome. Our core genome results showed that there is no such correlation. Additionally, there was no correlation between the core genome and the *cas1* gene phylogeny (Fig. 2) and direct repeat consensus (Supplementary Fig. 1). This fact has already been observed for the *C. diphtheriae* species [8]. By the core genome phylogenetic analysis, MDS 1954 and 1961 isolates were grouped with one pre-1992 surgical incision isolate and two Chinese sputum and nasopharyngeal isolates in 2018. MDR isolates 2023 and 2308, classified as PFGE profile I in previous studies [4, 7], were grouped in the same clade. Similarly, MDR isolates 2230 and 2237, classified as PFGE profile VI [4], were grouped in the same clade. *C. striatum* Brazilian MDR isolates are grouped into a distinct clade from those isolates from other countries, but we do not know their antimicrobial susceptibility profiles, except for MDR isolates 215 and 216 from the USA mentioned above [3].

Searching for the CRISPR system of the Cas1 protein type I-E' by BLASTp/NCBI, we have found that it exists only in one other *Corynebacterium* species: *C. simulans* (data not shown), phylogenetically related to *C. striatum*. Also, *cas* operon arrangement in the *C. simulans* genome is equivalent to the type I-E' system found in our *C. striatum* isolates. When all type I-E system Cas proteins were compared with those of type I-E' system, they showed low identity (Table 3). In addition to the low similarity (30.7%) between the type I-E and type I-E' system Cas1 proteins (Table 3), phylogenetic analysis revealed that the type I-E' system Cas1 protein branched independently from the type I-E system Cas1 protein (Fig. 2).

Cas3 protein is the signature of type I systems, responsible for target DNA cleavage and degradation [43]. *Cas3* gene (Supplementary Fig. 2) phylogenetic analyses were also carried out and showed that the *cas3* gene of the isolates that have the type I-E' system separated into a different clade from the isolates that have the type I-E system.

Horizontal gene transfer favors the survival and adaptation of bacteria and archaea that acquire virulence factors, the ability to degrade toxic compounds, and antibiotic

resistance. Although CRISPR systems provide bacterial immunity against horizontally acquired elements, this system is not 100% effective. Some acquired elements are maintained when they confer selective advantage [40], and this could explain the presence of intact prophages found in 3 of our isolates. To determine whether a bacterial lineage has been previously exposed to a specific phage, a matching CRISPR spacer sequence must be found. Phages can regulate host population size and can alter bacterial physiology as well bacteria phenotype in toxin production, antibiotic resistance genes, virulence factors, photosystem components, CRISPR-Cas system, and other metabolic and genes with unknown functions [43]. The intact prophage regions of isolates 2130 and 2425 are similar in length, 43.2 kb and 43.8 kb, respectively, both do not share sequence similarity and also differ in GC content. A large part of the prophage region is found in isolate MDS 1961 (6.9 kb to 9.1 kb), aligned with one of the 4 intact prophage regions *C. striatum* 215, an MDR sputum isolate from the USA. Approximately 21 kb of the phage region of isolate 2130 was aligned with various phage regions found in *C. striatum* strain, and only approximately 5 kb of the phage region was found in isolate 2425, aligned with the phage region of *C. striatum* DSM 45,711 (FDAARGOS 1197), isolated from blood in 2011 in Italy, according to information available in the German Collection of Microorganisms DSMZ.¹⁰ In addition to the intact phages found in 3 of our isolates previously described, all isolates showed incomplete phage sequences with a total of 8 in isolate 2308 (data not shown). These finds show that the CRISPR-Cas system is not always effective against foreign DNA invasions.

As important emergent pathogens, the CRISPR-Cas system in *C. striatum* was found in 9 genomes from a nosocomial outbreak that occurred at a public hospital in Rio de Janeiro, Brazil. Of these 9 clinical isolates, 8 are MDR with different resistance levels. According to the current classification method [11], the 13 CRISPR-Cas loci detected in our isolates are classified as type I-E system. However, additional *cas* operon alternative gene arrangement in 3 isolates (Fig. 1b) was distinct from the classic type I-E system, despite the low identity between all proteins of both systems (Table 3). Phylogenetic analyses of *cas1* (Fig. 2) and *cas3* genes (Supplementary Fig. 2) and of direct repeat consensus sequences (Supplementary Fig. 1) separated both type I-E and type I-E' systems into two distinct clades. Furthermore, no spacers were shared between the two systems. As observed by Ramos et al. [4], the profile PFGE I was the most frequently observed in the hospital, with susceptibility only to tetracycline, vancomycin, and linezolid, and it also

¹⁰ <https://www.dsmz.de/collection/catalogue/details/culture/DSM-45711>

presented a second CRISPR-Cas system called here type I-E' system. As observed for other species, such as *E. coli* and *Enterococci* [15, 16], there is no correlation between the presence of CRISPR-Cas system and multidrug resistance in *C. striatum*. This is the first study about the CRISPR-Cas system occurrence, arrangement, and diversity in *C. striatum*, and these finds may contribute to further investigations, in particular about the role of the CRISPR-Cas system in *C. striatum*.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42770-022-00881-4>.

Author contribution JNR, PVPB, and VVV conceived and designed the study, generated the data, and wrote the manuscript. ALMG, EMDV, and JFCV analyzed the data and wrote the manuscript. All authors read and approved the submitted manuscript.

Funding This study was financed in part by Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ), Fundação Oswaldo Cruz (Programa Inova Fiocruz/VPPCB-007-FIO-18-2-103), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), and Sub-Reitoria de Pós-Graduação e Pesquisa da Universidade do Estado do Rio de Janeiro (SR-2/UERJ).

Data availability All data generated or analyzed during this study are included in this published article (and its supplementary information files).

Code availability Not applicable.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

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

References

- Alibi S, Ferjani A, Gaillot O, Marzouk M, Courcol R, Boukadida J (2015) Identification of clinically relevant *Corynebacterium* strains by Api Coryne, MALDI-TOF-mass spectrometry and molecular approaches. *Pathol Biol (Paris)* 63(4–5):153–157. <https://doi.org/10.1016/j.patbio.2015.07.007>
- Daisuke U, Oishi T, Yamane K, Terada K (2017) *Corynebacterium striatum* bacteremia associated with a catheter-related blood stream infection. *Case Rep Infect Dis* 2017(2):10–12. <https://doi.org/10.1155/2017/2682149>
- Nudel K, Zhao X, Basu S et al (2018) Genomics of *Corynebacterium striatum*, an emerging multi-drug resistant pathogen of immunocompromised patients. *Clin Microbiol Infect* 24(9):1016.e7. <https://doi.org/10.1016/J.CMI.2017.12.024>
- Ramos J, Souza C, Faria Y et al (2019) Bloodstream and catheter-related infections due to different clones of multidrug-resistant and biofilm producer *Corynebacterium striatum*. *BMC Infect Dis* 19(1):1–11. <https://doi.org/10.1186/s12879-019-4294-7>
- Alibi S, Ferjani A, Boukadida J et al (2017) Occurrence of *Corynebacterium striatum* as an emerging antibiotic-resistant nosocomial pathogen in a Tunisian hospital. *Sci Rep* 7(1):1–8. <https://doi.org/10.1038/s41598-017-10081-y>
- Navas J, Fernández-Martínez M, Salas C, Cano M, Martínez-Martínez L (2016) Susceptibility to aminoglycosides and distribution of aph and aac(3)-XI genes among *Corynebacterium striatum* clinical isolates Cloeckaert A ed. *PLoS One* 11(12):0167856. <https://doi.org/10.1371/journal.pone.0167856>
- Baio P, Mota H, Freitas A et al (2013) Clonal multidrug-resistant *Corynebacterium striatum* within a nosocomial environment, Rio de Janeiro. *Brazil Mem Inst Oswaldo Cruz* 108(1):23–29. <https://doi.org/10.1590/S0074-02762013000100004>
- Sangal V, Fineran P, Hoskisson P (2013) Novel configurations of type I and II CRISPR–Cas systems in *Corynebacterium diphtheriae*. *Microbiol* 159(10):2118–2126. <https://doi.org/10.1099/MIC.0.070235-0>
- Zhao N, Li L, Luo G et al (2020) Multiplex gene editing and large DNA fragment deletion by the CRISPR/Cpf1-RecE/T system in *Corynebacterium glutamicum*. *J Ind Microbiol Biotechnol* 47(8):599–608. <https://doi.org/10.1007/S10295-020-02304-5>
- Makarova K, Koonin E (2015) Annotation and classification of CRISPR-Cas systems. *Methods Mol Biol* 1311:47–75. https://doi.org/10.1007/978-1-4939-2687-9_4
- Makarova K, Wolf Y, Iranzo J et al (2020) Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat Rev Microbiol* 18(2):67–83. <https://doi.org/10.1038/S41579-019-0299-X>
- Koonin E, Makarova K, Zhang F (2017) Diversity, classification and evolution of CRISPR-Cas systems. *Curr Opin Microbiol* 37:67–78. <https://doi.org/10.1016/J.MIB.2017.05.008>
- Mosterd C, Rousseau G, Moineau S (2021) A short overview of the CRISPR-Cas adaptation stage. *Can J Microbiol* 67(1):1–12. <https://doi.org/10.1139/CJM-2020-0212>
- Duan C, Cao H, Zhang L, Xu Z (2021) Harnessing the CRISPR-Cas systems to combat antimicrobial resistance. *Front Microbiol* 12:2369. <https://doi.org/10.3389/FMICB.2021.716064/BIBTEX>
- Price V, McBride S, Hullahalli K, Chatterjee A, Duerkop B, Palmer K (2019) *Enterococcus faecalis* CRISPR-Cas is a robust barrier to conjugative antibiotic resistance dissemination in the murine intestine. *MSphere* 4(4). <https://doi.org/10.1128/MSPHERE.00464-19>
- Touchon M, Charpentier S, Pognard D et al (2012) Antibiotic resistance plasmids spread among natural isolates of *Escherichia coli* in spite of CRISPR elements. *Microbiol (Reading)* 158(Pt 12):2997–3004. <https://doi.org/10.1099/MIC.0.060814-0>
- Cheleuitte-Nieves C, Gulvik C, McQuiston J et al (2018) Genotypic differences between strains of the opportunistic pathogen *Corynebacterium bovis* isolated from humans, cows, and rodents. *PLoS One* 13(12). <https://doi.org/10.1371/JOURNAL.PONE.0209231>

18. Parise D, Parise M, Viana M, et al (2018) First genome sequencing and comparative analyses of *Corynebacterium pseudotuberculosis* strains from Mexico. *Stand Genomic Sci* 13(1). <https://doi.org/10.1186/S40793-018-0325-Z>
19. Guimarães L, Viana M, Benevides L et al (2016) Draft genome sequence of toxigenic *Corynebacterium ulcerans* strain 04–7514, isolated from a dog in France. *Genome Announc* 4(2). <https://doi.org/10.1128/GENOMEA.00172-16>
20. Guimarães L, Soares S, Trost E et al (2015) Genome informatics and vaccine targets in *Corynebacterium urealyticum* using two whole genomes, comparative genomics, and reverse vaccinology. *BMC Genomics* 16(Suppl 5):S7. <https://doi.org/10.1186/1471-2164-16-S5-S7>
21. Wang Q, Zhang J, al Makishah N, et al (2021) Advances and perspectives for genome editing tools of *Corynebacterium glutamicum*. *Front Microbiol*. <https://doi.org/10.3389/FMICB.2021.654058>
22. BrCAST. Brazilian Committee on Antimicrobial Susceptibility Testing - BrCAST Tabelas de pontos de corte para interpretação de CIMs e diâmetros de halos - versão 11. BrCAST. Published 2021. Accessed October 26, 2021. <http://brcast.org.br/documentos/>
23. Magiorakos A, Srinivasan A, Carey R et al (2012) Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect* 18(3):268–281. <https://doi.org/10.1111/J.1469-0691.2011.03570.X>
24. Couvin D, Bernheim A, Toffano-Nioche C et al (2018) CRISPR-CasFinder, an update of CRISPRfinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res* 46(W1):W246–W251. <https://doi.org/10.1093/NAR/GKY425>
25. Biswas A, Gagnon J, Brouns S, Fineran P, Brown C (2013) CRISPRtarget: bioinformatic prediction and analysis of crRNA targets. *RNA Biol* 10(5):817–827. <https://doi.org/10.4161/RNA.24046>
26. Deng W, Nickle D, Learn G, Maust B, Mullins J (2007) ViroBLAST: a stand-alone BLAST web server for flexible queries of multiple databases and user's datasets. *Bioinformatics* 23(17):2334–2336. <https://doi.org/10.1093/BIOINFORMATICS/BTM331>
27. Crooks G, Hon G, Chandonia J, Brenner S (2004) WebLogo: a sequence logo generator. *Genome Res* 14(6):1188–1190. <https://doi.org/10.1101/GR.849004>
28. Arndt D, Grant J, Marcu A et al (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44(W1):W16–W21. <https://doi.org/10.1093/NAR/GKW387>
29. Seemann T (2014) Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30(14):2068–2069. <https://doi.org/10.1093/BIOINFORMATICS/BTU153>
30. Page A, Cummins C, Hunt M et al (2015) Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 31(22):3691–3693. <https://doi.org/10.1093/BIOINFORMATICS/BTV421>
31. Kumar S, Stecher G, Li M, Knyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol* 35(6):1547–1549. <https://doi.org/10.1093/MOLBEV/MSY096>
32. Leyton B, Ramos J, Baio P, et al (2021) Treat me well or will resist: uptake of mobile genetic elements determine the resistome of *Corynebacterium striatum*. *Int J Mol Sci* 22(14). <https://doi.org/10.3390/IJMS22147499>
33. Gholizadeh P, Köse Ş, Dao S et al (2020) How CRISPR-Cas system could be used to combat antimicrobial resistance. *Infect Drug Resist* 13:1111–1121. <https://doi.org/10.2147/IDR.S247271>
34. Bozic B, Repac J, Djordjevic M (2019) Endogenous gene regulation as a predicted main function of type I-E CRISPR/Cas system in *E. coli*. *Molecules* 24(4). <https://doi.org/10.3390/MOLECULES24040784>
35. Nudel K, Zhao X, Basu S et al (2018) Genomics of *Corynebacterium striatum*, an emerging multidrug-resistant pathogen of immunocompromised patients. *Clin Microbiol Infect* 24(9):1016.e7–1016.e13. <https://doi.org/10.1016/J.CMI.2017.12.024>
36. Tagini F, Pillonel T, Croxatto A, et al (2018) Distinct genomic features characterize two clades of *Corynebacterium diphtheriae*: proposal of *Corynebacterium diphtheriae* subsp. *diphtheriae* subsp. nov. and *Corynebacterium diphtheriae* subsp. *lausannense* subsp. nov. *Front Microbiol*. 9. <https://doi.org/10.3389/FMICB.2018.01743>
37. Reimann C, Moore R, Little S, Savioz A, Willetts N, Haas D (1989) Genetic structure, function and regulation of the transposable element IS21. *Mol Gen Genet* 215(3):416–424. <https://doi.org/10.1007/BF00427038>
38. Vandecraen J, Chandler M, Aertsen A, van Houdt R (2017) The impact of insertion sequences on bacterial genome plasticity and adaptability. *Crit Rev Microbiol* 43(6):709–730. <https://doi.org/10.1080/1040841X.2017.1303661>
39. Álvarez-Narváez S, Giguère S, Anastasi E, Hearn J, Scotti M, Vázquez-Boland J (2019) Clonal confinement of a highly mobile resistance element driven by combination therapy in *Rhodococcus equi*. *mBio* 10(5). <https://doi.org/10.1128/MBIO.02260-19>
40. Shen J, Lv L, Wang X, Xiu Z, Chen G (2017) Comparative analysis of CRISPR-Cas systems in *Klebsiella* genomes. *J Basic Microbiol* 57(4):325–336. <https://doi.org/10.1002/JOBM.201600589>
41. Richter C, Chang J, Fineran P (2012) Function and regulation of clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR associated (Cas) systems. *Viruses* 4(10):2291–2311. <https://doi.org/10.3390/V4102291>
42. Trost E, Blom J, Soares S et al (2012) Pangenomic study of *Corynebacterium diphtheriae* that provides insights into the genomic diversity of pathogenic isolates from cases of classical diphtheria, endocarditis, and pneumonia. *J Bacteriol* 194(12):3199–3215. <https://doi.org/10.1128/JB.00183-12>
43. Luo M, Mullis A, Leenay R, Beisel C (2015) Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. *Nucleic Acids Res* 43(1):674. <https://doi.org/10.1093/NAR/GKU971>
44. Michael J, Tisza, Christopher B. Buck (2021) A catalog of tens of thousands of viruses from human metagenomes reveals hidden associations with chronic diseases. *Proc Natl Acad Sci USA*. 118(23). <https://doi.org/10.1073/PNAS.2023202118>

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