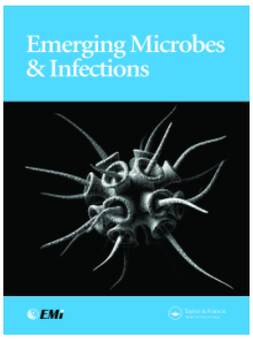

Emerging Microbes & Infections

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/temi20

Genomic evolution and rearrangement of CTX-Φ prophage elements in *Vibrio cholerae* during the 2018–2024 cholera outbreaks in eastern Democratic Republic of the Congo

Leonid M. Irenge, Jérôme Ambroise, Bertrand Bearzatto, Jean-François Durant, Maxime Bonjean, Louisette K. Wimba & Jean-Luc Gala

To cite this article: Leonid M. Irenge, Jérôme Ambroise, Bertrand Bearzatto, Jean-François Durant, Maxime Bonjean, Louisette K. Wimba & Jean-Luc Gala (2024) Genomic evolution and rearrangement of CTX-Φ prophage elements in *Vibrio cholerae* during the 2018–2024 cholera outbreaks in eastern Democratic Republic of the Congo, Emerging Microbes & Infections, 13:1, 2399950, DOI: 10.1080/22221751.2024.2399950

To link to this article: https://doi.org/10.1080/22221751.2024.2399950

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd

View supplementary material

Published online: 12 Sep 2024.

Submit your article to this journal

Article views: 843

View related articles

View Crossmark data

Citing articles: 3 View citing articles

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=temi20

Emerging Microbes & Infections
2024, VOL. 13, 2399950 (10 pages)
https://doi.org/10.1080/22221751.2024.2399950

OPEN ACCESS Check for updates

Genomic evolution and rearrangement of CTX-Φ prophage elements in *Vibrio cholerae* during the 2018–2024 cholera outbreaks in eastern Democratic Republic of the Congo

Leonid M. Irenge[a], Jérôme Ambroise[a], Bertrand Bearzatto[a], Jean-François Durant[a], Maxime Bonjean[a], Louisette K. Wimba[b] and Jean-Luc Gala[a]

[a]Centre for Applied Molecular Technologies, Institute of Clinical and Experimental Research, Université catholique de Louvain (UCLouvain), Woluwe-Saint-Lambert, Belgium; [b]Institut Supérieur des Techniques Médicales/Bukavu, Bukavu, The Democratic Republic of the Congo

ABSTRACT

Between 2018 and 2024, we conducted systematic whole-genome sequencing and phylogenomic analysis on 263 *V. cholerae* O1 isolates from cholera patients across four provinces in the Democratic Republic of Congo (North-Kivu, South-Kivu, Tanganyika, and Kasai Oriental). These isolates were classified into the AFR10d and AFR10e sublineages of AFR10 lineage, originating from the third wave of the seventh El Tor cholera pandemic (7PET). Compared to the strains analysed between 2014 and 2017, both sublineages had few genetic changes in the core genome but recent isolates (2022-2024) had significant CTX prophage rearrangement. AFR10e spread across all four provinces, while AFR10d appeared to be extinct by the end of 2020. Since 2022, most *V. cholerae* O1 isolates exhibited significant CTX prophage rearrangements, including a tandem repeat of an environmental satellite phage RS1 downstream the *ctxB* toxin gene of the CTX-Φ-3 prophage on the large chromosome, as well as two or more arrayed copies of an environmental pre-CTX-Φ prophage precursor on the small chromosome. We used Illumina data for mapping and coverage estimation to identify isolates with unique CTX-Φ genomic features. Gene localization was then determined on MinION-derived assemblies, revealing an organization similar to that of non-O1 *V. cholerae* isolates found in Asia (O139 VC1374, and environmental O4 VCE232), but never described in *V. cholerae* O1 El Tor from the third wave. In conclusion, while the core genome of AFR10d and AFR10e showed minimal changes, significant alterations in the CTX-Φ and pre-CTX-Φ prophage content and organization were identified in AFR10e from 2022 onwards.

ARTICLE HISTORY Received 20 June 2023; Revised 22 November 2023; Accepted 11 December 2023

KEYWORDS *Vibrio cholerae*; DRC; genomic evolution; whole genome; prophages; CTX-Φ rearrangement

Introduction

Cholera is a severe watery diarrheal disease caused by *Vibrio cholerae* serogroups O1 and/or O139. Symptoms result from a dimeric toxin encoded by CTX-φ, a filamentous bacteriophage [1–3]. Cholera remains a significant healthcare concern in developing countries [1–4]. It is estimated to cause up to 3 million cases of cholera and approximately 95,000 deaths annually [5,6]. Over the past decade, Sub-Saharan Africa has become the most afflicted by cholera [7,8], now referred to as the "new cholera homeland" [9], with countries in this region reporting 83% of global cases between 2000 and 2015 [10]. The DRC is among the leading countries in terms of cholera cases, contributing 5-14% of the global case count annually [5,11,12]. In 2020, DRC reported 19,789 cases, ranking second only to Yemen in terms of numbers [13]. Recent data from 2023 indicate a deepening cholera crisis in the country, with over 41,000 cases and 314 deaths reported [14].

Cholera has traditionally been most prevalent in the Great Lakes Region (GLR) [15], but recent reports show it has spread to most DRC provinces, including areas well beyond the GLR [16]. For example, between 2018-2022, Kasai Oriental province in the central region of DRC reported 2,176 cholera cases [17]. The initial introduction of *V. cholerae* O1 strains into the DRC occurred during the first wave of P7ET, known as the T5 introduction (1970–1972) [18,19]. A subsequent introduction, in the early nineties, as part of the P7ET's third wave and designated as the T10 introduction event, brought a strain from southern Asia [20], which has persisted in the GRL through successive outbreaks [21,22]. Previous studies, including our own, have documented the division of the T10 (AFR10) lineage into two sublineages, AFR10d, and

CONTACT Jean-Luc Gala jean-luc.gala@uclouvain.be Centre for Applied Molecular Technologies, Institute of Clinical and Experimental Research, Université catholique de Louvain (UCLouvain), Avenue Hippocrate 54/B1.54.01, 1200 Woluwe-Saint-Lambert, Belgium

Supplemental data for this article can be accessed online at https://doi.org/10.1080/22221751.2024.2399950.

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Shanghai Shangyixun Cultural Communication Co., Ltd
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

AFR10e, and identified a regional focus of endemic disease derived from the T10 introduction [21].

In this study conducted between 2018 and 2024, we performed systematic whole-genome sequencing and phylogenetic analyses on *V. cholerae* isolates collected from cholera cases in the North-Kivu, South-Kivu, and Tanganyika provinces, which are part of the GLR. Additionally, we included strains collected in 2020 in the province of Kasai Oriental, located in the centre of the DRC. By comparing these genomes with data from *V. cholerae* isolates collected between 2014 and 2017, we aimed to assess the genetic evolution of *V. cholerae* in the DRC and explore potential genetic links between isolates from the GLR and those from the central Kasai Oriental province.

Material and methods

Ethical clearance

The study protocols received approval from the Ethical Review Board (ERB) of the ISTM/Bukavu, DRC (ISTM-BUKAVU/CRPS/CIES/ML/0016/2023). The ERB granted a waiver of informed consent, determining that the study complied with international guidelines applicable to research during severe outbreaks. To protect confidentiality, samples were analysed anonymously.

Phenotypic analysis

The isolation of *V. cholerae* bacteria and subsequent phenotypical tests and antimicrobial susceptibility assays followed protocols previously described [21]. Rectal swabs were collected from patients exhibiting clinical symptoms of cholera, notably severe watery diarrhoea. These samples were incubated in alkaline peptone water for 6–8 h, then streaked onto thiosulfate-citrate-bile salts-sucrose (TCBS) agar Petri dishes and incubated at 37°C overnight. Colonies appearing large yellow, flattened with opaque centres and translucent peripheries were further cultured on nutritive alkaline agar.

For phenotypical confirmation of *V. cholerae* standard microbiological tests were carried out [23], alongside immunological tests to determine the serotypes of the isolates. Identification of *V. cholerae* was further confirmed by a duplex real-time PCR targeting the *V. cholerae*-specific outer membrane protein W (*omp*W) gene (forward: 5'-CAAACCATTTGCGGCCTAGCC 3'; reverse: 5'-CCCGCGCGCACAATAAAGTC-3'; probe: Yakima-Yellow-5'-CTTGCAGCCCTACTAGCCGCTCCTGTAT-3'-BHQ-1), and the cholera toxin subunit (*ctx*A) gene (forward 5'-TTTGTTAGGCACGATGATGGAT-3'; reverse: 5'-ACCAGACAATATAGTTTGACCCACTAAG-3'; probe: 6-FAM-5'-TGTTTCCACCTCAATTAGTTTGAGAAGTGCCC-3'-BHQ-1). The duplex real-time quantitative PCR was carried out in a 25 µL reaction mixture containing 2.5 µL of extracted DNA as template, 300 nM of each primer, 100 nM of each probe, and 12.5 µL of Takyon No ROX Probe 2x Master Mix Blue dTTP (Eurogentec, Ougrée, Belgium). Amplification was performed on a Biorad CFX96 Touch™ Real-Time PCR Detection System (Biorad, Hercules, California, USA). The reaction was initiated at 95°C for 5 min followed by 45 cycles of denaturation at 95°C for 10 s, annealing/extension at 60°C for 30 s. Data were recorded as crossing points (Cq) using the analytical Biorad CFX manager 3.1 software.

Antimicrobial susceptibility was assessed using the diffusion method on Mueller-Hinton agar plates, using antibiotic disks. Bacterial isolates were prepared at a density equivalent to 0.5 to 0.6 McFarland standard, and the zone of inhibition around antibiotics discs was measured after 16–20 h of incubation.

Minimum inhibitory concentration (MIC) for ciprofloxacin was determined with MIC test strips (Liofilchem, Roseto degli Abruzzi, Italy) following the manufacturer's guidelines. The interpretation of MIC values inhibition zone diameters was based on the EUCAST clinical breakpoint values (version 13.0, https://www.eucast.org/clinical_breakpoints).

Whole-genome sequencing

Genomic DNA (gDNA) was isolated from all *V. cholerae* isolates in our collection using an EZ1 Advanced XL Biorobot and the Tissue DNA Kit (QIAGEN, Hilden, Germany), employing the Bacterial Card according as per the manufacturer's instructions. Quantification of the isolated DNA was performed using a Qubit® fluorometer (Thermo Fisher Scientific, Eugene, OR, USA).

For whole-genome sequencing (WGS) of *V. cholerae,* we initially used Illumina short-reads sequencing. Short-read libraries were prepared from 70 ng DNA using an Illumina DNA Prep kit (Illumina, San Diego, CA, USA), and then sequenced using paired-end (2 × 300 bp) reads on a MiSeq platform (Illumina, San Diego, CA).

To investigate the genomic organization within the prophage *V. cholerae* CTX-Φ, WGS was also carried out using the long-reads sequencing on the GridION platform (Oxford Nanopore Technologies, UK) for a subset of 24 samples, selected across different collection years. Long-reads libraries were prepared from 400 ng of genomic DNA using the rapid barcoding kit 96 V14 (SQK-RBK114) and sequenced on an R10.4.1 flow cell run in a GridION device over a 72-h period.

Bioinformatics

(a) Phylogenetic analysis

Raw sequencing data from all *V. cholerae* isolates were submitted to the European Nucleotide Archive (ENA)

under accession number PRJEB66194 (http://www.ebi.ac.uk/ena). The accession number of each isolate is reported in Supplementary Table S1.

For Illumina short-reads, our bioinformatics pipeline started with quality checking via FastQC v.0.11.9 [24] followed by trimmomatic v.0.39 [25]. De novo assembly of paired-end reads for each *V. cholerae* isolate was performed using Spades v.3.13.0 [26], generating a draft genome sequence. Assembly quality was assessed using QUAST 5.0.2 [27] and completeness was verified with the BUSCO v.4.1.1 algorithm, using the bacteria_odb10 dataset [28]. Assembled genomes were submitted to NCBI and the accession number of each isolate is reported in Supplementary Table S1.

Snippy v4.6.0 was used to detect Single Nucleotide Polymorphisms (SNPs), insertions and deletions (indels) between each assembled genome and the O1 *V. cholerae* isolate N16961 (accession NZ_LT906614), which served as a reference. Individual reports were then used to identify SNPs and indels found in all isolates from a particular sublineage (e.g. ST69 or ST515). In parallel, Snippy was used to identify core SNPs, and recombined regions were identified and excluded from the pseudo-genome alignment using Gubbins v.1.4.10 [29]. Finally, the phylogenetic tree was built using FastTree v.2.1.0 [30] and rooted at the reference genome (O1 *V. cholerae* isolate N16961) with Dendroscope 3.8.5 [31]. The final phylogenomic tree was represented using the R ggtree v.3.12.0 Bioconductor package [32]. *In-silico* Multi-Locus Sequence Typing (MLST) was performed on all genomes by using the screen-Blast-mlst function of the custom Pathogenomics R package (https://github.com/UCL-CTMA/Pathogenomics), requiring 100% sequence identity and coverage for allele type assignment. All genomes were screened for *V. cholerae*-associated virulence factor genes as per the VFDB core database (http://www.mgc.ac.cn/VFs) [33], with an identity cut-off set at 80%. Genotypic AMR profiles were assessed using AMRFinderPlus v3.10.24.

(b) Gene annotation

To identify exclusive marker genes in either ST515 or ST69, draft genomes based on Illumina sequences were annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP). The amino acid sequences of the predicted genes were used as input for the BPGA (bacterial pan-genome analysis tool) comparative genomics pipeline v1.3 [34].

(c) Monitoring of the *V. cholerae* O1 CTX-Φ prophage

(i) Monitoring and characterizing the genomic organization of the CTX-Φ prophage involved a two-step process that began with mapping and normalizing Illumina data against representative sequences of CTX-Φ prophage genes, including *rst*R, *rst*A, *rst*B, *rst*C, *psh*, *cep*, *pIIIctx*, *ace*, *zot*, *ctx*A, *ctx*B. The resulting coverages were normalized using the single-copy reference gene *pnt*A. This approach facilitated the estimation of copy numbers for each CTX-Φ prophage gene.

(ii) Following this mapping, isolates that exhibited unusual coverage profiles in CTX-Φ prophage genes were identified and selected for further analysis. The selected isolates underwent de novo assembly genome analysis using GridION long reads with Canu v2.2.1 [35]. The genomic organization of the prophage and the localization of prophage genes on the large or small chromosome were determined using these assembled genomes. CTX-Φ prophage sequences were localized within the contigs using Blastn v2.12.0 [36]. The graphical representation of the prophage organization was generated using the Gviz v1.21.1 Bioconductor package.

Results

V. cholerae serotyping and antimicrobial susceptibility

In this study, 269 clinical isolates presumably carrying *V. cholerae* were analysed phenotypically and by real-time PCR. Of these, *V. cholerae* was identified in 263 isolates (97.8%). A map of DR Congo displays the sample collection sites by time and province (Figure 1). The phenotypic characterization data and antimicrobial susceptibility profiles of the 263 *V. cholerae* isolates are presented in Table 1 and Supplementary Table S2. All isolates tested positive for agglutination with the O1 antiserum. Among these, 236 isolates agglutinated with the Ogawa and 27 with the Inaba antisera, respectively. All *V. cholerae* isolates O1 were found to be susceptible to ampicillin, chloramphenicol, and tetracycline. However, they showed resistance to cotrimoxazole and polymyxin B, the latter being indicative of the El Tor biotype of *V. cholerae* [37]. While all Inaba isolates (n = 27) were susceptible to ciprofloxacin, the Ogawa isolates (n = 236) demonstrated reduced susceptibility to this antibiotic.

Phylogenetics of *V. cholerae* O1 isolates in the DRC provinces under study

Phylogenetic analyses revealed that all *V. cholerae* O1 isolates belonged to the AFR10 lineage (comprising AFR10d and AFR10e sublineages), aligning with previous reports from the same region [21,22]. In line with an previous study [21], the two DRC AFR10

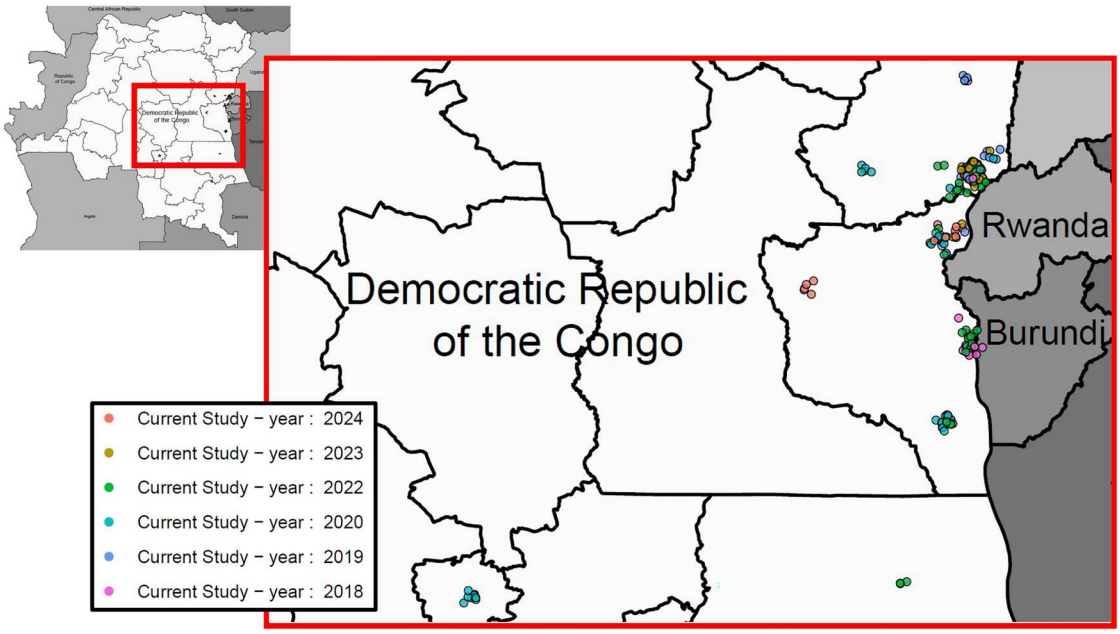

Figure 1. Map of the DR Congo showing the localization of the sample collection by time and province.

sublineages, AFR10e and AFR10d, associated with sequence types ST69 and ST515 respectively, had 65 distinct core SNPs, as shown in Figure 2 and Supplementary Figure S1.

All isolates from the Kasai Oriental ($n = 9$) and Tanganyika provinces ($n = 2$) characterized in this study belonged to the AFR10e sublineage (Figure 3). From the 79 *V. cholerae* O1 isolates characterized in South-Kivu province between 2018 and 2024, the vast majority ($n = 77$) belonged to the AFR10e lineage, whereas only two isolates (collected in 2018) belonged to the AFR10d sublineage. Similarly, in North-Kivu, the incidence of AFR10d declined dramatically by the end of 2019, with only one AFR10d isolate identified in 2020, and none reported since. During the 2018–2024 period, there were 148 isolates from the AFR10e and 25 from the AFR10d sublineage. Therefore, in contrast to our observations during the 2014–2017 period, the AFR10e sublineage significantly outpaced its AFR10d counterpart in both North-Kivu and South-Kivu provinces (Figure 3).

Table 1. Phenotypic characterization data and antimicrobial susceptibility profiles of Inaba and Ogawa isolates.

O1 Inaba ($n = 27$)				O1 Ogawa ($n = 236$)			
AMR genes (%)		Antimicrobial agents (%)		AMR genes (%)		Antimicrobial agents (%)	
*gyr*A (S83I)	100	CIP	S (100)	*gyr*A (S83I)	100	CIP	R (100)
*par*C (S85L)	0			*par*C (S85L)	100		
SXT/R391	100	SXT	R (100)	*SXT/R391*	100	SXT	R (100)
dfr1	100			*dfr1*	100		
Sul2	100			*Sul2*	100		
*cat*B9	100	CHL	S (100)	*cat*B9	100	CHL	S (100)
*flo*R	100			*flo*R	100		
varG	100	AMP	S (100)	*varG*	100	AMP	S (100)
blaSPE	0			*blaSPE*	0		
*tet*A	0	TET	S (100)	*tetA*	0	TET	S (100)
*tet*E	0	DOX	S (100)	*tetE*	0	DOX	S (100)
*alm*E	100	PXB	R (100)	*alm*E	100	PXB	R (100)
*alm*F	100			*alm*F	100		
*alm*G	100			*alm*G	100		

Notes: AMP: Ampicillin; CHL: Chloramphenicol; CIP: Ciprofloxacin; DOX: Doxycycline; PXB: Polymyxin B; SXT: Sulfamethoxazole-Trimethoprim; TET: Tetracycline. Antimicrobial susceptibility profile of isolates is expressed as the percentage of resistant isolates of *V. cholerae*. AMR genes are in in italic.

AMR genes

The analysis of antimicrobial (AMR) genes revealed that all DRC AFR10d and AFR10e isolates, consistent with previous studies, harboured the SXT/R391 integrative conjugative element ICEVchBan5, which carries genes associated with reduced susceptibility to cotrimoxazole [38] (Table 1 and Supplementary Table S2 and S3). Similar to findings in all isolates of the AFR10 sublineage [22], both AFR10d and AFR10e isolates exhibited the S83I substitution in the *gyr*A protein. Additionally, all AFR10e isolates presented the S85L substitution in the *par*C protein, a combination associated with reduced susceptibility to quinolones in *V. cholerae* [22]. Notably, no AFR10d isolate in this study was found to carry both the S83I substitution in gyrA and the S85L substitution in parC. In contrast, clinical *V. cholerae* isolates from Tanganyika ($n = 3$) and Kasai Oriental ($n = 9$), all phylogenetically belonging to the AFR10e sublineage, carried both the S83I gyrA, and S85L parC substitutions. Moreover, all isolates possessed the *cat*B9 AMR, which is associated with resistance to chloramphenicol, and the beta-lactamase *varG* gene. All but one isolate (CTMA 2005) possessed the phenicol resistance *flo*R gene, and the sulphonamide resistance *sul*II gene. Moreover, all *V. cholerae* O1 isolates characterized in this study possessed the *cat*B9 AMR,

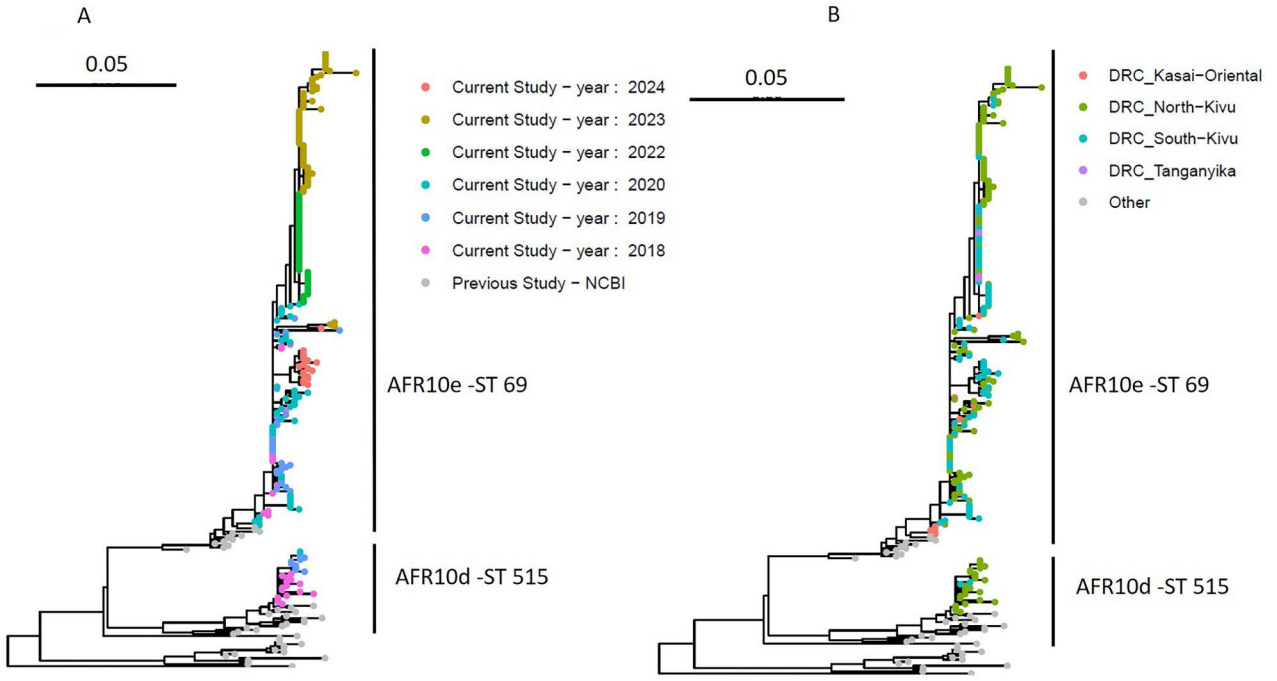

Figure 2. Phylogenomic analysis of *V. cholerae* O1 isolates (2014–2024) from four investigated DRC provinces, emphasizing the years of isolate collection (2A) and the collection province (2B). Notes: The trees were derived from 750 core genome SNPs mapped against the 7PET *V. cholerae* O1 biotype EL Tor N19691, which was used as reference genome and as an outgroup to root the tree.

which is associated with resistance to chloramphenicol, and the beta-lactamase *var*G gene. All but one isolate (CTMA 2005) possessed the phenicol resistance *flo*R gene, and the sulphonamide resistance *sul*2 gene. Surprisingly, the presence of these AMR genes did not compromise the susceptibility of AFR10 isolates to ampicillin, chloramphenicol, and tetracycline.

Virulence genes

AFR10d and AFR10e isolates analysed in DRC between 2018 and 2024 possessed the same virulence genes as those characterized in the earlier 2014–2017 study [21]. These include:

(i) Virulence genes associated with the CTX-Φ-3 prophage [39,40], as previously described in all AFR10 isolates, except for three isolates collected in 2023 (CTMA_1964, CTMA_1965, CTMA_1974) which had a deletion of all the CTX-Φ prophage genes on the large chromosome.
(ii) The complete Vibrio pathogenicity island-1 (VPI-1) [41], the Vibrio Seventh Pandemic Island-I (VSP-I), as well as the Vibrio Seventh Pandemic Island (VSP-II) [42]. Notably, VSP-II exhibits a deletion spanning from ORF *V. cholerae*_0495 to *V. cholerae*_0512.

Furthermore, no genetic changes were identified in the virulence genes compared to those in the previous study [21].

Genetic changes in chromosomal genes

Major genetic changes between AFR10d and AFR10e sublineages are summarized in Supplementary Table S4, the majority of which have already been reported [21]. In addition to these known changes, three deletions/insertions were identified in the large chromosome of AFR10d isolates:

(i) an insertion of A at position 612959 in the HA/protease gene regulator *hap*R,
(ii) a deletion of AAATCA at position 1568200 in the type VI secretion system (T6SS) that secretes self-protecting proteins TsiV1-3, and
(iii) a deletion of GTGT at position 1667905 in the RidA family protein (positions according to *V. cholerae* O1 El Tor N16961 accession number: NZ_LT906614).

The following deletions/insertions were identified in AFR10e: (i) an insertion of TTGTCGA in the IS3 family protein gene at position 532251 in the large chromosome, (ii) and a deletion of G at position 967750 in the small chromosome (*V. cholerae* O1 El Tor N19691; accession number: NZ_LT906615).

Pan-genome analysis

PGAP annotation identified an average of 3569 protein coding genes per genome. Among all identified genes, 3263 were identified as part of the core genome by BPGA while the number of accessory

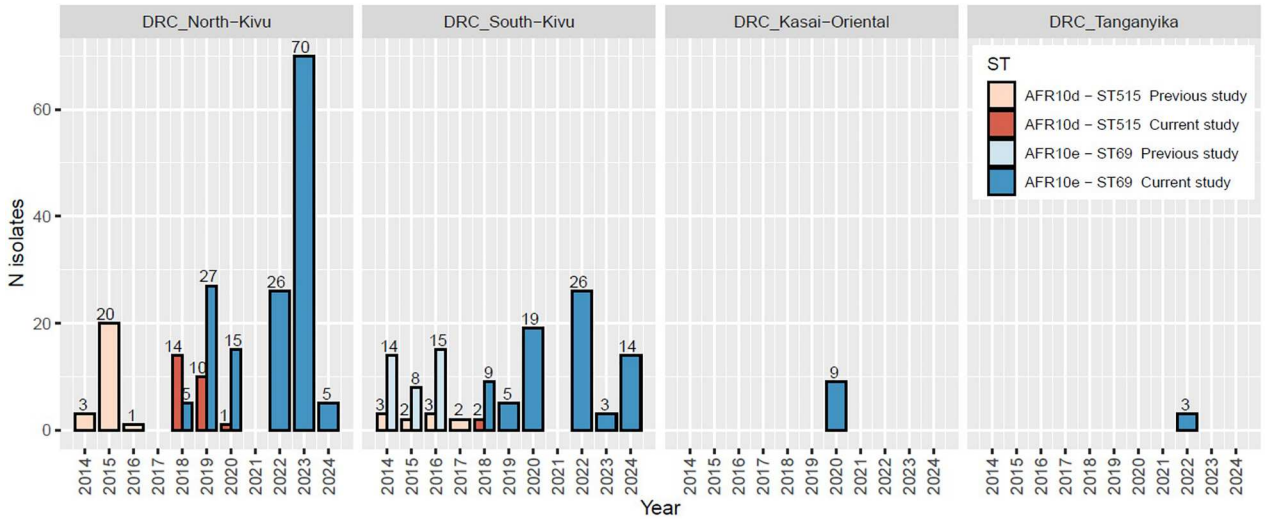

Figure 3. Temporal prevalence of AFR10d and AFR10e *V. cholerae* O1 isolates in the four DRC provinces under investigation. For the Kasai Oriental and Tanganyika provinces, strains were collected only in 2020 and 2022, respectively.

genes per genome ranged between 194 and 215. Moreover, BPGA analysis identified six marker genes that were unique to one ST (either 69 or 515), and completely absent from the other (Table 2).

CTX-Φ prophage organization in O1 isolates

The mapping of Illumina reads against CTX-Φ prophage representative sequences enabled the detection of unique variations in gene copy numbers. Between 2018 and 2020, all isolates were characterized by a single copy of the CTX core prophage genes (*psh, cep, pIIIctx, ace, zot, ctx*A, *ctx*B) (Supplementary Figure S2A). In contrast, isolates collected in 2022 exhibited multiple copies of *psh, cep, pIIIctx, ace,* and *zot* (Supplementary Figure S2B). A few isolates (N = 3) from 2023 lacked *ctx*A and ctxB (Supplementary Figure S2C), while others from this period showed coverage profiles similar to those from 2022 and earlier.

Based on Illumina results, representative isolates (n = 24) showing a distinct coverage profile in CTX-Φ prophage genes (Supplementary Table S1) were selected for de novo assembly genome analysis using GridION long reads with Canu v2.2.1. The localization of CTX-Φ prophage genes in GridION-derived assemblies using Blastn software confirmed Illumina observations and provided a detailed organization of the CTX prophage arrays.

Isolates collected between 2018 and 2020 (Figure 4 (A)) harboured a complete CTX-Φ-3 prophage, which includes the RS2 satellite phage and the core CTX-Φ region carrying a classical *ctx*B gene. This prophage was integrated into the large chromosome between the loci VC1465 and the *rtx*A gene of the RTX toxin locus. In most of these isolates, the CTX-Φ-3 prophage was flanked upstream by one canonical RS1 satellite phage ($RS1^{ca}$), aligning with observations of several CTX-Φ prophages in atypical El Tor *V. cholerae* strains [39,40].

Isolates collected from 2022 displayed significant changes in the organization of the CTX-Φ prophage array and its satellite phages on both the large and the small chromosomes. On the large chromosome, one $RS1^{ca}$ satellite phage flanked the CTX-Φ-3 prophage upstream, while two tandem copies of an environmental RS1, named $RS1^{env\text{-}DRC}$, flanked the CTX-3 Φ prophage downstream of the *ctx*B toxin gene (Figure 4 (B)). The $RS1^{env\text{-}DRC}$ variant is characterized by canonical *rst*A and *rst*B genes, an *rstC* gene with a single amino acid mismatch (L64P) compared to the canonical *rst*C gene in CTX-Φ prophage, and an *rst*R gene coding for a 86-aa long protein that matches perfectly (100%) with the *rst*R characterized in the environmental Mozambican *V. cholerae* IB1617 isolate [43].

Another notable finding was observed in DRC *V. cholerae* O1 bacteria isolated from 2022, specifically on the small chromosome with the presence of two tandemly arrayed copies of a pre-CTX-Φ prophage inserted between VCA0569 and VCA0570 loci [40] (Figure 4(B)). This arrangement resembles the configuration of the pre-CTX-Φ prophage on the small chromosome of the *V. cholerae* O139 VC1374 [44]. Despite sharing a similar pre-CTX-Φ prophage configuration and an identical *rst*R gene (named

Table 2. ST-specific genes identified by the pan-genome analysis using bacterial pan-genome analysis tool (BPGA).

Product	Symbol	ST-Specific
Acyltransferase	-	ST515_AFR10d
ExeA family protein	-	ST69_AFR10e
FkbM family methyltransferase	-	ST69_AFR10e
Restriction endonuclease subunit S	-	ST69_AFR10e
RidA family protein	-	ST69_AFR10e
BREX-1 system adenine-specific DNA-methyltransferase PglX	pglX	ST69_AFR10e

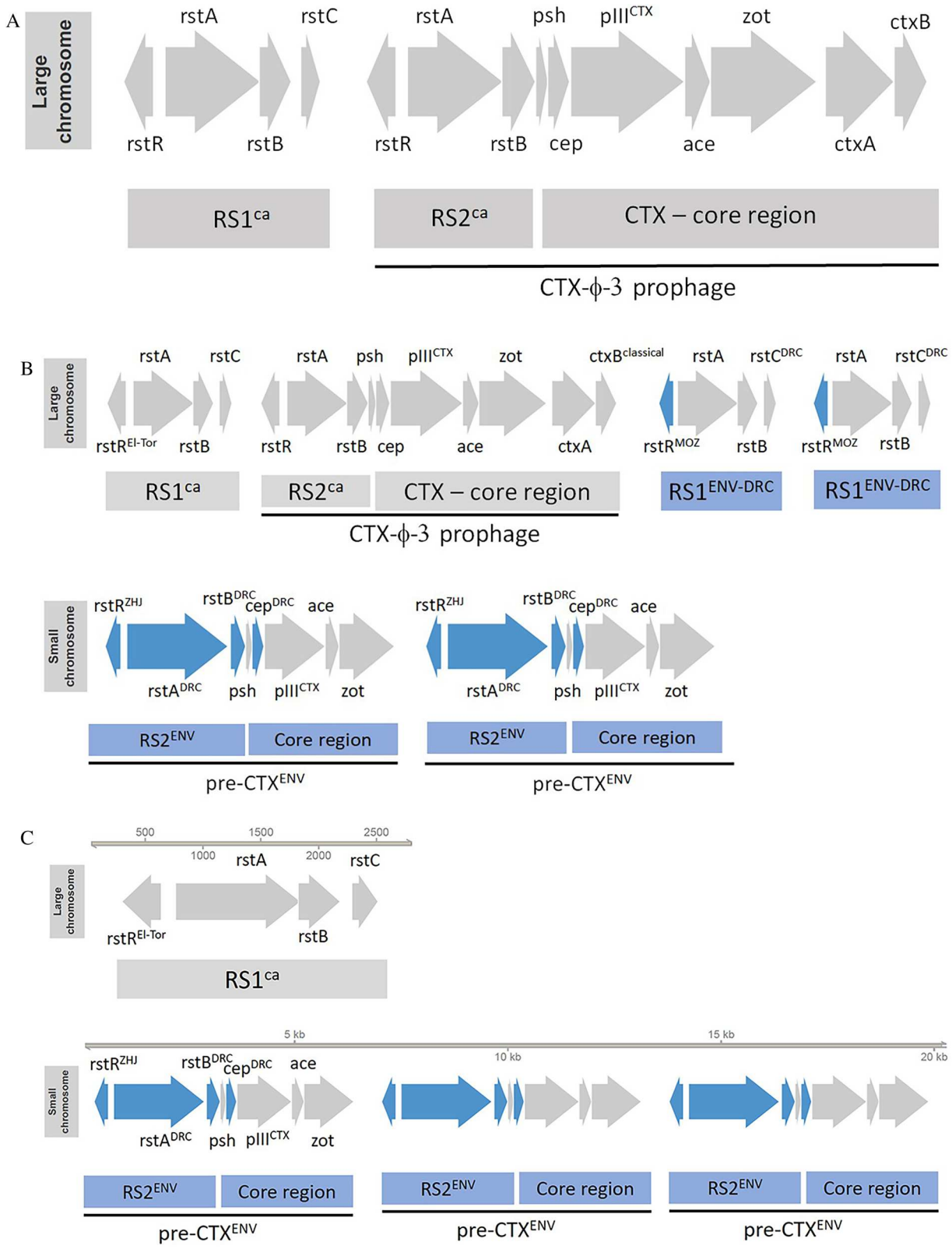

Figure 4. (A) Representative organization of the CTX-Φ prophage and its satellite in *V. cholerae* isolates collected between 2018 and 2020 in the DRC. Notes: Block arrows represent genes present in each element of the CTX-Φ-3-phage and indicate the direction of transcription. RS1^{ca} and RS2^{ca} represent RS1 and RS2 canonical satellite phages. (B) Representative organization of the CTX-Φ prophage and its satellites in both the large and small chromosomes in the majority (125/147) of *V. cholerae* isolates collected since 2022 in DRC. Notes: Canonical sequences for each gene are shown in grey, while environmental sequences are shown in blue. (C) CTX-Φ-3 prophage organization observed in three *V. cholerae* isolates collected in 2023.

*rst*R^{ZHJ}) with the *V. cholerae* O139 VC1374 isolate, the 2022-DRC *V. cholerae* O1 isolates exhibited several amino acid substitutions in the rstA, rstB and cep proteins compared to both the *V. cholerae* O139 VC1374, and to *V. cholerae* non-O1/non-O139 VCE232 [45]. These variants in the rstA, rstB, and

cep proteins have been designated as rstA^{DRC}, rstB^{DRC}, and cep^{DRC}, respectively, to distinguish them from their environmental counterparts [44,45]. In 2022, *V. cholerae* O1 isolates showed a distinct configuration of the CTX-Φ prophage array and its satellites on the large chromosome: RS1^{ca} – CTX prophage (RS2^{ca} – core region (*psh* - *cep* - *pIII*^{ctx} - *ace* - *zot* - *ctx*A - *ctx*B^{classical}) - RS1$^{env\text{-}DRC}$ - RS1$^{env\text{-}DRC}$), and on the small chromosome two to four consecutive copies of an environmental pre-CTX constituted of a RS2env (*rst*R^{ZHJ} - *rst*A^{DRC} - *rst*B^{DRC}) and a core region (*psh* - *cep*^{DRC} - *pIII*^{ctx} - *ace* - *zot*). Notably, such a hybrid configuration has not been previously observed in 7PET *V. cholerae* O1 from the third wave.

The 2022 hybrid CTX prophage configuration persisted throughout 2023 and 2024, accounting for 76.1% of all AFR10e isolates (n = 92). In 19 isolates (20.6%), the canonical CTX-Φ-3 prophage described prior to 2022 was still present. A novel prophage rearrangement was observed in three (3.3%) of the *V. cholerae* O1 AFR10e isolated from Goma and its Karisimbi neighbourhood (CTMA-1964, CTMA-1965, and CTMA-1974). This rearrangement was characterized by the complete deletion of the core CTX-Φ-3-prophage on the large chromosome and the presence of two to four copies of the environmental pre-CTX of 2022 on the small chromosome (Figure 4(C)).

Discussion

The objective of this study was to phylogenetically characterize *V. cholerae* isolates associated with cholera outbreaks in DRC from 2018 to 2022, building upon a previous study that analysed isolates from the eastern DRC between 2014 and 2017 [21]. Our results are consistent with previous reports that *V. cholerae* O1 isolates from the third wave and T10 transmission (AFR10) remain the major clade responsible for cholera outbreaks in the eastern part of the country [13,21,22]. This evidence marks the sustained prevalence of a single clade as unprecedented when compared to the establishment of other *V. cholerae* O1 clades in neighbouring African countries [46–48].

Furthermore, our findings are in line with recent studies that show an AFR10d decline and the dominance of the AFR10e sublineage [22]. In particular, we demonstrate that *V. cholerae* O1 isolates from Kasai Oriental province, the first from central DRC to be molecularly characterized, belong to the AFR10e sublineage.

These isolates are genetically linked to cholera outbreaks in the GLR provinces, highlighting the geographical spread and persistence of this clade within the DRC.

Despite the geographical isolation of the Kasai Oriental province, which lacks direct natural linkage to the GLR through rivers or lakes, the identification of AFR10e isolates phylogenetically related to AFR10e isolates from the GLR strengthens the hypothesis of a westward expansion of cholera in DRC [21], contrasting with other hypotheses on possible transmission routes in central DRC [49]. The presence of *V. cholerae* AFR10e isolates from GLR in this area suggests potential transmission routes between the GLR and central regions. Human movement, including train travel, could facilitate such transmission. It has been suggested that the primary cause of AFR10d decline was the emergence of ciprofloxacin resistance in AFR10e due the coexistence of gyrA S83I and parC S85L substitutions [22]. While this factor appeared to play a significant role in the potential AFR10d extinction, other genetic factors may have also been involved, as AFR10d isolates with both substitutions and reduced susceptibility to ciprofloxacin were identified in our previous study [21]. Genetic changes in the core genome, of AFFR10e warrant therefore further investigation (Supplementary Table S4). Furthermore, AFR10e *V. cholerae* bacteria isolated in 2022 and beyond exhibited significant changes in the CTX-Φ prophage configuration. Although rearrangements in CTX-Φ prophage have often been linked to specific waves of *V. cholerae* O1 [39], our results show that all analysed *V. cholerae* O1 isolates still belong to the 10th transmission event.

Noteworthy, an upstream RS1^{ca} satellite and two downstream tandem copies of RS1$^{env\text{-}DRC}$, a novel environmental RS1 satellite, flanked the CTX-Φ-3 prophage on the large chromosome. Additionally, two to four consecutive pre-CTX-Φ^{env} prophages containing the *rst*R^{ZHJ} gene, similar to those described in a *V. cholerae* O139 isolate [44], were discovered on the small chromosome of these DRC AFR10e *V. cholerae* O1 isolates. This rearrangement results in the co-existence of a CTX-3-Φ prophage, an environmental pre-CTX-Φ prophage, and RS1 satellites with both clinical and environmental origins within the same bacterial cell. Although epidemiological evidence is lacking, the similarity of this new CTX array to that of the *V. cholerae* non-O1 isolates (O139 and O4) previously characterized in Asia [43,44] suggests a possible recombination between AFR10e from Africa and these Asian *V. cholerae* strains. This recombination may have resulted in a hybrid CTX-3-Φ prophage with elements present on both the large and the small chromosomes. While O139 has never reported outside Asia, the recent and prolonged presence of several foreign military contingents in eastern DRC (e.g. MONUSCO with contingents from several Asian countries) may have served as a catalyst, as evidenced by the *V. cholerae* outbreak in Haiti following the presence of United Nations peacekeepers from Nepal [50].

Alternatively, RS1 and pre-CTX-Φ proteins play a crucial role in regulating the replication, morphogenesis, and the virulence of the CTX-Φ phage [51]. These observed genomic rearrangements may reflect and significantly influence the evolutionary dynamics of *V. cholerae* in DR Congo. This includes the replication, excision of CTX prophages elements [51,52] and their role in the transfer of CTX prophage-associated genes, with global warming facilitating significant genetic changes in *V. cholerae* [53].

The discovery of these genetic changes in AFR10e isolates collected in 2022 from various provinces across the DRC not only underscores their expansion into new territories but also suggests an enhanced capacity for environmental adaptation. This adaptability may provide them with a competitive advantage over other *V. cholerae* O1 isolates, potentially facilitating their rapid spread and dominance.

In conclusion, our study not only confirms the continued dominance of AFR10e sublineage in cholera outbreaks within the DRC but also highlights potential transmission routes and reveals significant genetic alterations within the CTX-Φ prophage. These findings underscore the critical need for ongoing monitoring of *V. cholerae* O1 to better predict and mitigate future outbreaks, emphasizing the importance of understanding genetic evolution in the development of public health responses.

Acknowledgements

We extend our gratitude to the technical teams of C.D.R.M/AFIA-TELE (South-Kivu Province) and AMI-LABO (North-Kivu province) for their invaluable assistance in the sampling and processing of faecal samples from patients suspected of having cholera, which facilitated the isolation of *V. cholerae*. Special thanks are due to the laboratories and authors who submitted sequences to the GISAID EpiFlu and the NCBI databases. Further acknowledgment goes to the EMBL-EBI team (Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK) for providing ENA accession numbers, thus enhancing the accessibility of our data.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was funded by the Belgian Cooperation Agency of the ARES (Académie de Recherche et d'Enseignement Supérieur) under grant COOP-CONV-20-022. The funder had no role in the study design, collection, analysis and interpretation of data, preparation of the manuscript, or the decision to submit the paper for publication.

References

[1] Clemens JD, Nair GB, Ahmed T, et al. Cholera. Lancet. 2017;390(10101):1539–1549.
[2] Lippi D, Gotuzzo E, Caini S. Cholera. Microbiol Spectr. 2016;4(4).
[3] OUNOftCoH Affairs. Latin America & The Caribbean Situation Update (12-18 December 2022) OCHA; 2022. Available from: https://reliefweb.int/report/haiti/latin-america-caribbean-weekly-situation-update-12-18-december-2022-19-december-2022.
[4] Kanungo S, Azman AS, Ramamurthy T, et al. Cholera. Lancet. 2022;399(10333):1429–1440.
[5] Ali M, Nelson AR, Lopez AL, et al. Updated global burden of cholera in endemic countries. PLoS Negl Trop Dis. 2015;9(6):e0003832.
[6] Lessler J, Moore SM, Luquero FJ, et al. Mapping the burden of cholera in sub-Saharan Africa and implications for control: an analysis of data across geographical scales. Lancet. 2018;391(10133):1908–1915.
[7] Mengel MA, Delrieu I, Heyerdahl L, et al. Cholera outbreaks in Africa. Curr Top Microbiol Immunol. 2014;379:117–144.
[8] Zheng Q, Luquero FJ, Ciglenecki I, et al. Cholera outbreaks in sub-Saharan Africa during 2010-2019: a descriptive analysis. Int J Infect Dis. 2022;122:215–221.
[9] Gaffga NH, Tauxe RV, Mintz ED. Cholera: a new homeland in Africa? Am J Trop Med Hyg. 2007;77(4):705–713.
[10] Machado A, Amorim E, Bordalo AA. Major stressors favoring Cholera trigger and dissemination in Guinea-Bissau (West Africa). Int J Environ Res Public Health. 2021;18(21):11296.
[11] Ingelbeen B, Hendrickx D, Miwanda B, et al. Recurrent Cholera outbreaks, Democratic Republic of the Congo, 2008-2017. Emerg Infect Dis. 2019;25(5):856–864.
[12] WHO. Disease outbreak news; Cholera - global situation; 2022. Available from: https://www.who.int/emergencies/disease-outbreak-news/item/2022-DON426.
[13] Alam MT, Mavian C, Paisie TK, et al. Emergence and evolutionary response of *Vibrio cholerae* to novel bacteriophage, Democratic Republic of the Congo(1). Emerg Infect Dis. 2022;28(12):2482–2490.
[14] WHO. A new resolve to eliminate cholera in DRC; 2023.
[15] Bompangue D, Giraudoux P, Piarroux M, et al. Cholera epidemics, war and disasters around Goma and Lake Kivu: an eight-year survey. PLoS Negl Trop Dis. 2009;3(5):e436.
[16] Africa W-ROf. Weekly bulletin on outbreaks and other emergencies—week 1 2018; 2018. Available from: http://apps.who.int/iris/bitstream/10665/259809/1/OEW1-2018.pdf.
[17] Division Provinciale de la Santé du Kasai Oriental RC. Nombre de cas de choléra dans la province du Kasai Oriental. Ministère de la Santé; 2020.
[18] Ramamurthy T, Mutreja A, Weill FX, et al. Corrigendum: revisiting the global epidemiology of Cholera in conjunction with the genomics of *Vibrio cholerae*. Front Public Health. 2019;7:237.
[19] Mutreja A, Kim DW, Thomson NR, et al. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature. 2011;477(7365):462–465.
[20] Weill FX, Domman D, Njamkepo E, et al. Genomic history of the seventh pandemic of cholera in Africa. Science. 2017;358(6364):785–789.
[21] Irenge LM, Ambroise J, Mitangala PN, et al. Genomic analysis of pathogenic isolates of *Vibrio cholerae* from eastern Democratic Republic of the Congo (2014-2017). PLoS Negl Trop Dis. 2020;14(4):e0007642.

[22] Hounmanou YMG, Njamkepo E, Rauzier J, et al. Genomic microevolution of *Vibrio cholerae* O1, Lake Tanganyika Basin, Africa. Emerg Infect Dis. 2023;29 (1):149–153.
[23] Ramamurthy T, Das B, Chakraborty S, et al. Diagnostic techniques for rapid detection of *Vibrio cholerae* O1/O139. Vaccine. 2020;38(Suppl 1):A73–A82.
[24] Andrews S. FastQC: a quality control tool for high throughput sequence data. Cambridge: Babraham Bioinformatics, Babraham Institute; 2010.
[25] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics. 2014;30(15):2114–2120.
[26] Bankevich A, Nurk S, Antipov D, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19 (5):455–477.
[27] Gurevich A, Saveliev V, Vyahhi N, et al. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29(8):1072–1075.
[28] Seppey M, Manni M, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness. In: Walker JM, editor. Methods in Molecular Biology. New York (NY): Humana; 2019. p. 227–245.
[29] Croucher NJ, Page AJ, Connor TR, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic Acids Res. 2015;43(3):e15.
[30] Price MN, Dehal PS, Arkin AP. Fasttree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol. 2009;26(7):1641–1650.
[31] Huson DH, Richter DC, Rausch C, et al. Dendroscope: An interactive viewer for large phylogenetic trees. BMC Bioinformatics. 2007;8(460):1–6.
[32] Yu G. Using ggtree to visualize data on tree-like structures. Currt Protoc Bioinformatics. 2020;69(1):e96.
[33] Liu B, Zheng D, Zhou S, et al. VFDB 2022: a general classification scheme for bacterial virulence factors. Nucleic Acids Res. 2022;50(D1):D912–D9D7.
[34] Chaudhari NM, Gupta VK, Dutta C. BPGA – an ultra-fast pan-genome analysis pipeline. Sci Rep. 2016;6:24373.
[35] Koren S, Walenz BP, Berlin K, et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27(5):722–736.
[36] Camacho C, Coulouris G, Avagyan V, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:1–9.
[37] Matson JS, Yoo HJ, Hakansson K, et al. Polymyxin B resistance in El Tor *Vibrio cholerae* requires lipid acylation catalyzed by MsbB. J Bacteriol. 2010;192 (8):2044–2052.
[38] Wozniak RA, Fouts DE, Spagnoletti M, et al. Comparative ICE genomics: insights into the evolution of the SXT/R391 family of ICEs. PLoS Genet. 2009;5(12):e1000786.
[39] Kim EJ, Lee D, Moon SH, et al. CTX prophages in *Vibrio cholerae* O1 strains. J Microbiol Biotechnol. 2014;24(6):725–731.
[40] Pant A, Das B, Bhadra RK. CTX phage of *Vibrio cholerae*: genomics and applications. Vaccine. 2020;38 (Suppl 1):A7–A12.
[41] Kumar A, Das B, Kumar N. Vibrio Pathogenicity Island-1: the master determinant of Cholera pathogenesis. Front Cell Infect Microbiol. 2020;10:561296.
[42] Dziejman M, Balon E, Boyd D, et al. Comparative genomic analysis of *Vibrio cholerae*: genes that correlate with cholera endemic and pandemic disease. Proc Natl Acad Sci USA. 2002;99(3):1556–1561.
[43] Choi SY, Lee JH, Kim EJ, et al. Classical RS1 and environmental RS1 elements in *Vibrio cholerae* O1 El Tor strains harbouring a tandem repeat of CTX prophage: revisiting Mozambique in 2005. J Med Microbiol. 2010;59(Pt 3):302–308.
[44] Li X, Zhao L, Gao H, et al. A novel pre-CTX prophage in the *Vibrio cholerae* serogroup O139 strain. Infect Genet Evol. 2020;81:104238.
[45] Verma J, Bag S, Saha B, et al. Genomic plasticity associated with antimicrobial resistance in *Vibrio cholerae*. Proc Natl Acad Sci U S A. 2019;116(13):6226–6231.
[46] Mwaba J, Debes AK, Murt KN, et al. Three transmission events of *Vibrio cholerae* O1 into Lusaka, Zambia. BMC Infect Dis. 2021;21(1):570.
[47] Bwire G, Sack DA, Almeida M, et al. Molecular characterization of *Vibrio cholerae* responsible for cholera epidemics in Uganda by PCR, MLVA and WGS. PLoS Negl Trop Dis. 2018;12(6):e0006492.
[48] Hounmanou YMG, Leekitcharoenphon P, Kudirkiene E, et al. Genomic insights into *Vibrio cholerae* O1 responsible for cholera epidemics in Tanzania between 1993 and 2017. PLoS Negl Trop Dis. 2019;13(12):e0007934.
[49] Kayembe HCN, Linard C, Bompangue D, et al. The spread of cholera in western Democratic Republic of the Congo is not unidirectional from East-West: a spatiotemporal analysis, 1973-2018. BMC Infect Dis. 2021;21(1):1261.
[50] Hendriksen RS, Price LB, Schupp JM, et al. Population genetics of *Vibrio cholerae* from Nepal in 2010: evidence on the origin of the Haitian outbreak. mBio. 2011;2(4):e00157–11.
[51] Davis BM, Kimsey HH, Kane AV, et al. A satellite phage-encoded antirepressor induces repressor aggregation and cholera toxin gene transfer. EMBO J. 2002;21(16):4240–4249.
[52] Ramamurthy T, Pragasam AK, Taylor-Brown A, et al. *Vibrio cholerae* O139 genomes provide a clue to why it may have failed to usher in the eighth cholera pandemic. Nat Commun. 2022;13(1):3864.
[53] Chowdhury FR, Nur Z, Hassan N, et al. Pandemics, pathogenicity and changing molecular epidemiology of cholera in the era of global warming. Ann Clin Microbiol Antimicrob. 2017;16(1):10.