

# Genome sequences of *Vibrio cholerae* strains isolated in the DRC between 2009 and 2012

Alexandre Lemopoulos,<sup>1</sup> Berthe Miwanda,<sup>2</sup> Natália C. Drebes Dörr,<sup>1</sup> Sandrine Stutzmann,<sup>1</sup> Didier Bompangue,<sup>3</sup> Jean-Jacques Muyembe-Tamfum,<sup>2,3</sup> Melanie Blokesch<sup>1</sup>

**AUTHOR AFFILIATIONS** See affiliation list on p. 2.

**ABSTRACT** *Vibrio cholerae* has caused seven cholera pandemics in the past two centuries. The seventh and ongoing pandemic has been particularly severe on the African continent. Here, we report long read-based genome sequences of six *V. cholerae* strains isolated in the Democratic Republic of the Congo between 2009 and 2012.

**KEYWORDS** *Vibrio cholerae*, genome sequences

Cholera is an acute, life-threatening diarrheal disease caused by *Vibrio cholerae* and is readily transmitted through the fecal-oral route, especially in areas with poor access to sanitation and clean drinking water (1). The ongoing seventh cholera pandemic started in 1961 in Indonesia (1) and has subsequently been introduced onto the African continent from South Asia multiple times (2, 3). Cholera is endemic in the African Great Lakes Region (4), with the Democratic Republic of the Congo (DRC) severely affected (5). Here, we announce the genome sequences of six *V. cholerae* strains originating from an unusually large epidemic that arose in the region surrounding Kisangani, later spreading westward along the Congo River during 2011–2012 (6, 7). The strains were isolated by the DRC's Institut National de Recherche Biomédicale between 2009 and 2012 as part of their cholera surveillance program. Upon shipment to Switzerland, isolates were streaked on thiosulfate citrate bile salts agar, colony-purified, grown in lysogeny broth (LB) medium, and stored as glycerol stocks at –80°C. Following inoculation from the glycerol stock and aerobic growth in LB medium at 30°C, genomic DNA was isolated as previously outlined, employing Qiagen's 100/G and 500/G columns in conjunction with the manufacturer's genomic DNA buffer set (8, 9). Sequencing and assembly were performed by the Genomic Technology Facility of the University of Lausanne and followed a reported protocol (9). Briefly, high-molecular-weight DNA was sheared to obtain 10- to 15-kb fragments using a Megaruptor (Diagenode, USA). DNA (500 ng) was used to prepare a SMRTbell library using the PacBio SMRTbell Express Template Prep Kit v.2.0 (Pacific Biosciences, USA). The resulting libraries were pooled and size-selected with Ampure PacBio beads to eliminate fragments below 3 kb. The sequencing was based on v.3.0/v.3.0 chemistry and diffusion loading on a PacBio Sequel II instrument at 900-min movie length with a pre-extension time of 120 min using a SMRT cell 8M v.3. Sequence quality control, genome assembly, circularization, and rotation were performed using the protocol Microbial Assembly in SMRT Link v.10.1 with default parameters. Assembled genomes were submitted to National Center for Biotechnology Information and annotated using the Prokaryotic Genome Annotation pipeline v.6.5 (10). Information on the assembled genomes is provided in Table 1.

Analysis of the genome sequences using Snippy v.4.6.0 (<https://github.com/tseemann/snippy>), Gubbins v.3.3.0 (11), RAxML v.8.12.12 (12), and Geneious v.11.0.14.1 indicated that two different strain types were isolated. The first strain type belongs to the seventh pandemic El Tor (7PET) lineage. Specifically, these strains sit within the

**Editor** Catherine Putonti, Loyola University Chicago, Chicago, Illinois, USA

Address correspondence to Melanie Blokesch, [melanie.blokesch@epfl.ch](mailto:melanie.blokesch@epfl.ch).

The authors declare no conflict of interest.

See the funding table on p. 3.

**Received** 4 September 2023

**Accepted** 21 January 2024

**Published** 12 February 2024

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

TABLE 1 Basic strain information and genome sequencing statistics

	DRC001 (#2506)	DRC052 (#2512)	DRC072 (#2516)	DRC186 (#2501)	DRC187 (#9154)	DRC193A (#1954)
Biosample	SAMN36815617	SAMN36815618	SAMN36815616	SAMN36815615	SAMN36815619	SAMN36815614
Year of isolation	2009	2012	2009	2011	2011	2011
Place of isolation (province)	Kalemie (Katanga)	Nord Kivu	Sud Kivu	Sud Kivu	Sud Kivu	Bandundu
GenBank accession numbers	JAZDAQ010000001 (chr 1) JAZDAQ010000003 (chr 2) JAZDAQ010000002 (contig 3)	CP132187 (chr 1) CP132188 (chr 2)	CP132182 (chr 1) CP132183 (chr 2)	JAVAXO010000001 (chr 1) JAVAXO010000002 (chr 2) JAVAXO010000003 (contig 3)	CP132190 (chr 1) CP132189 (chr 2)	CP132180 (chr 1) CP132181 (chr 2)
SRA <sup>c</sup> accession numbers	SRR26435098	SRR26435097	SRR26435099	SRR26435100	SRR26435096	SRR26435101
Number of contigs	3	2	2	3	2	2
Contig sizes	2,852,572 <sup>a</sup> 1,045,468 219,361 <sup>ab</sup>	3,063,607 1,059,659	3,060,728 1,059,647	2,963,376 <sup>a</sup> 1,107,495 36,710 <sup>ab</sup>	2,959,192 <sup>a</sup> 1,107,495	3,058,770 1,045,384
Total genome size	4,117,401	4,1232,66	4,1203,75	4,107,581	4,066,687	4,104,154
% GC content	47.6% (chr 1) 47.0% (chr 2) 48.3% (contig 3) <sup>b</sup>	47.7% (chr 1) 46.9% (chr 2)	47.7% (chr 1) 46.9% (chr 2)	47.6% (chr 1) 46.3% (chr 2) 51.4% (contig 3) <sup>b</sup>	47.5% (chr 1) 46.3% (chr 2)	47.7% (chr 1) 47.0% (chr 2)
Mean coverage	140×	248×	214×	294×	277×	271×
Long read N50	13,922	16,068	15,237	16,101	16,193	18,860
Raw read count	107,377	244,481	213,984	285,783	268,327	203,915

<sup>a</sup>Contig non-circularized.<sup>b</sup>Contig3 assembled separately by software but likely part of chr 1.<sup>c</sup>SRA, Sequence Read Archive.

two AFR10 sublineages recently described by Hounmanou et al. (3), with strains DRC052/DRC072/DRC193A and DRC001 belonging to clade AFR10d and AFR10e, respectively. Based on this similarity, our data support recent reports suggesting the 2011–2012 cholera epidemic in DRC was caused by a 7PET variant clonal complex (7). The second strain type (e.g., DRC186/DRC187) falls outside the 7PET clade. These non-O1 antigen strains are *ctx*- and *tcp*-negative. At this point, we cannot rule out the possibility that the sampled patients were coinfecting with 7PET strains, particularly considering the limitation of strain isolation based on single colonies.

## ACKNOWLEDGMENTS

The authors thank Silvan Vesenbeckh, Andrea Rinaldo, and members of the Blokesch group for fruitful discussion on cholera in DRC; Anne-Catherine Portmann and Rasmus L. Marvig for preliminary analyses of short-read sequencing data sets; and David W. Adams for proofreading the manuscript. We also acknowledge the work by the staff of the Lausanne Genomic Technologies Facility and the EPFL/UNIL Bioinformatics Competence Center for sample processing, sequencing, and genome assembly.

This work was supported by a SEED money grant by EPFL's Cooperation & Development Center and by EPFL intramural funding. M.B. is a Howard Hughes Medical Institute international research scholar (grant 55008726).

## AUTHOR AFFILIATIONS

<sup>1</sup>Laboratory of Molecular Microbiology, Global Health Institute, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

<sup>2</sup>Institut National de Recherche Biomédicale (INRB), Kinshasa, Democratic Republic of Congo

<sup>3</sup>Département de Microbiologie, Faculté de Médecine, Université de Kinshasa, Kinshasa, Democratic Republic of Congo

#### AUTHOR ORCID*s*

Alexandre Lemopoulos  <http://orcid.org/0000-0002-7997-5526>

Natália C. Drebes Dörr  <http://orcid.org/0000-0002-4088-8548>

Melanie Blokesch  <http://orcid.org/0000-0002-7024-1489>

#### FUNDING

Funder	Grant(s)	Author(s)
Howard Hughes Medical Institute (HHMI)	55008726	Melanie Blokesch

#### AUTHOR CONTRIBUTIONS

Alexandre Lemopoulos, Data curation, Writing – original draft, Formal analysis, Writing – review and editing | Berthe Miwanda, Resources | Natália C. Drebes Dörr, Writing – original draft, Formal analysis | Sandrine Stutzmann, Investigation | Didier Bompangue, Resources | Jean-Jacques Muyembe-Tamfum, Resources | Melanie Blokesch, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Supervision, Validation, Writing – original draft, Writing – review and editing

#### DATA AVAILABILITY

The genome assemblies and raw read sequences were deposited in National Center for Biotechnology Information's GenBank under Bioproject [PRJNA1001814](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1001814) and Sequence Read Archive accession numbers [SRR26435096](https://www.ncbi.nlm.nih.gov/sra/SRR26435096)–[SRR26435101](https://www.ncbi.nlm.nih.gov/sra/SRR26435101).

#### REFERENCES

- Kanungo S, Azman AS, Ramamurthy T, Deen J, Dutta S. 2022. Cholera. *Lancet* 399:1429–1440. [https://doi.org/10.1016/S0140-6736\(22\)00330-0](https://doi.org/10.1016/S0140-6736(22)00330-0)
- Weill FX, Domman D, Njamkepo E, Tarr C, Rauzier J, Fawal N, Keddy KH, Salje H, Moore S, Mukhopadhyay AK, et al. 2017. Genomic history of the seventh pandemic of cholera in Africa. *Science* 358:785–789. <https://doi.org/10.1126/science.aad5901>
- Hounmanou YMG, Njamkepo E, Rauzier J, Gallandat K, Jeandron A, Kamwiziku G, Porten K, Luquero F, Abedi AA, Rumedeka BB, Miwanda B, Michael M, Okitayemba PW, Saidi JM, Piarroux R, Weill FX, Dalsgaard A, Quilici ML. 2023. Genomic microevolution of *Vibrio cholerae* O1, Lake Tanganyika Basin, Africa. *Emerg Infect Dis* 29:149–153. <https://doi.org/10.3201/eid2901.220641>
- Bompangue D, Giraudoux P, Handschumacher P, Piarroux M, Sudre B, Ekwanzala M, Kebela I, Piarroux R. 2008. Lakes as source of cholera outbreaks, Democratic Republic of Congo. *Emerg Infect Dis* 14:798–800. <https://doi.org/10.3201/eid1405.071260>
- Rebaudet S, Sudre B, Faucher B, Piarroux R. 2013. Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. *J Infect Dis* 208 Suppl 1:S46–S54. <https://doi.org/10.1093/infdis/jit195>
- Bompangue D, Vesenbeckh SM, Giraudoux P, Castro M, Muyembe JJ, Kebela Ilunga B, Murray M. 2012. Cholera ante portas - the re-emergence of cholera in Kinshasa after a ten-year hiatus. *PLoS Curr* 4:RRN1310. <https://doi.org/10.1371/currents.RRN1310>
- Miwanda B, Moore S, Muyembe JJ, Nguéack-Tsague G, Kabangwa IK, Ndjakani DY, Mutreja A, Thomson N, Thefenne H, Garnotel E, Tshapenda G, Kakongo DK, Kalambayi G, Piarroux R. 2015. Antimicrobial drug resistance of *Vibrio cholerae*, Democratic Republic of the Congo. *Emerg Infect Dis* 21:847–851. <https://doi.org/10.3201/eid2105.141233>
- Drebes Dörr NC, Blokesch M. 2020. Interbacterial competition and anti-predatory behavior of environmental *Vibrio cholerae* strains. *Environ Microbiol* 22:4485–4504. <https://doi.org/10.1111/1462-2920.15224>
- Vesel N, Iseli C, Guex N, Lemopoulos A, Blokesch M. 2023. DNA modifications impact natural transformation of *Acinetobacter baumannii*. *Nucleic Acids Res* 51:5661–5677. <https://doi.org/10.1093/nar/gkad377>
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>
- Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 43:e15. <https://doi.org/10.1093/nar/gku1196>
- Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <https://doi.org/10.1093/bioinformatics/btu033>