Emergence of *Vibrio cholerae* O1 Sequence Type 75, South Africa, 2018–2020

Anthony M. Smith, François-Xavier Weill, Elisabeth Njamkepo, Hlengiwe M. Ngomane, Ntsieni Ramalwa, Phuti Sekwadi, Juno Thomas

We describe the molecular epidemiology of cholera in South Africa during 2018–2020. *Vibrio cholerae* O1 sequence type (ST) 75 recently emerged and became more prevalent than the *V. cholerae* O1 biotype El Tor pandemic clone. ST75 isolates were found across large spatial and temporal distances, suggesting local ST75 spread.

The seventh cholera pandemic, caused by *Vibrio cholerae* O1 biotype El Tor (7PET), arrived in Africa during 1970 and became endemic in many countries on the continent (*1*). Cholera was first reported in South Africa in 1974 (*2*). However, South Africa is not considered a cholera-endemic area; outbreaks typically are associated with importation, particularly from neighboring countries. The last cholera outbreak in South Africa was triggered by imported cases from an outbreak in Zimbabwe during 2008; South Africa reported 12,706 cases during November 2008–April 2009 (*3*).

Globally, 7PET isolates are genetically homogeneous and linked to the Bay of Bengal in South Asia (*4*,*5*). Most 7PET isolates are multidrug-resistant sequence type (ST) 69 (*6*). Rarely, 7PET has a single-locus variant, ST515, in isolates from Africa belonging to lineage T10 (*7*). As of September 2021, all cholera isolates from South Africa have been characterized as 7PET ST69 by multilocus sequence typing (MLST).

South Africa actively surveils for cholera. Since the 2008–2009 outbreak, few cases have been identified: 5 during 2010–2014, most of which were imported, and none during 2015–2017. During 2008–2009, large outbreaks occurred in 3 provinces, Mpumalanga, Limpopo, and KwaZulu-Natal (*3*), but all were caused by imported cases from neighboring Zimbabwe and Mozambique. Therefore, given their experience, healthcare workers and laboratorians in these provinces typically will test for cholera in all cases of acute watery diarrhea.

In South Africa, the National Institute for Communicable Diseases (NICD) is notified of suspected cholera cases. NICD's Centre for Enteric Diseases supports case investigations and receives all human and environmental *V. cholerae* isolates for further investigation. The case definition for confirmed cholera is isolation of *V. cholerae* O1 or O139 from a person with diarrhea. We investigated the molecular epidemiology of *V. cholerae* in South Africa during 2018–2020.

The Study

During February 2018–January 2020, NICD received 102 *V. cholerae* isolates for testing; 9 were identified as *V. cholerae* O1. We characterized the bacteria by whole-genome sequencing, comparative genomics, and phylogenetic analysis (Appendix 1, https://wwwnc.cdc.gov/EID/article/27/11/21-1144-App1.pdf). The Human Research Ethics Committee of the University of the Witwatersrand (Johannesburg, South Africa) provided ethics approval for this study (protocol no. M160667).

Of 9 *V. cholerae* O1 isolates tested, we identified 2 ST69 (7PET) and 7 ST75 isolates. The ST69 isolates were collected in October 2018 from 2 cholera patients in a family cluster. The index case-patient had traveled to Zimbabwe, where an outbreak was ongoing (*8*), within the 7-day cholera incubation period before symptom onset. We confirmed these ST69 isolates belonged to the previously described highly antimicrobial-resistant Zimbabwe outbreak strain (*8*). The 7 ST75 isolates originated from KwaZulu-Natal and Limpopo Provinces. Five isolates were collected

Author affiliations: National Institute for Communicable Diseases, Johannesburg, South Africa (A.M. Smith, H.M. Ngomane, N. Ramalwa, P. Sekwadi, J. Thomas); University of Pretoria, Pretoria, South Africa (A.M. Smith, N. Ramalwa); Institut Pasteur, Paris, France (F.-X. Weill, E. Njamkepo)

DOI: https://doi.org/10.3201/eid2711.211144

Table 1. Clinical and demographic characteristics of 5 patients hospitalized with *Vibrio cholerae* O1 ST75 diagnosed from stool cultures and risk factors for *V. cholerae* infection, South Africa, 2018–2020*

Isolate no.	Province	Sample collection date	Patient age, y/sex	Clinical manifestations	Source of drinking water	Sanitation	Linked environmental samples	Type of environmental sample, isolate no.
YA00085869	KwaZulu-Natal	2018 Feb 8	37/F	Acute watery diarrhea, dehydration	Untreated river water	NA	N	NA
YA00132994	Limpopo	2018 Nov 9	38/M	Acute watery diarrhea, vomiting, dehydration	Untreated borehole water	Pit latrine and open defecation	N	NA
YA00134463	Limpopo	2018 Nov 20	45/M	Acute watery diarrhea, dehydration	Untreated borehole water	Flush toilets	Y	Sewage, OA01603367
YA00192016	KwaZulu-Natal	2019 Dec 29	49/M	Acute watery diarrhea, abdominal cramps, dehydration	Untreated river water	Pit latrine	Y	River water, CF00214281
YA00193061	KwaZulu-Natal	2020 Jan 12	57/F	Acute watery diarrhea, dehydration	NA	NA	N	NA

*All cases were diagnosed from stool cultures. All patients survived. NA, Not available; ST, sequence type.
†Environmental samples tested positive for *V. cholerae* O1 ST75.

from patients with cholera, all adults 37–57 years of age; 2 isolates were from environmental samples collected during case investigations, 1 from sewage in Limpopo Province and 1 from river water in KwaZulu-Natal Province (Table 1). The 3 KwaZulu-Natal cases occurred ≈200–600 km apart; the first occurred in February 2018 and the last in January 2020. The 2 Limpopo cases occurred ≈70 km apart in the same district during November 2018. The Limpopo cases were ≥900 km from the KwaZulu-Natal cases. Epidemiologic investigations involved interviewing case-patients by using a standard case investigation form; visiting case-patients' residences to inspect water and sanitation services and interview other household members; collecting stool samples from household members; and collecting environmental samples when indicated. Investigators found no evidence of importation from another country, epidemiologic links between cases, or secondary transmission.

The 7 ST75 isolates showed notable features (Table 2). In particular, all carried the cholera toxin (CTX) prophage resembling CTX-2 with *ctxB1* genotype; *Vibrio* pathogenicity island 1 (VPI-1) encoding the toxin co-regulated pilus; and a variant form of *Vibrio* pathogenicity island 2 (VPI-2). However, isolates did not contain *Vibrio* seventh pandemic island I (VSP-I) and VSP-II. We noted several genomic islands (GIs), including VC-GI 119, but GI-05 was not present (Appendix 2, https://wwwnc.cdc.gov/EID/article/27/11/21-1144-App2.xlsx).

The only antimicrobial-resistance determinant found in all ST75 isolates was the *qnrVC4* gene, located in the chromosomal superintegron. Various *qnrVC* alleles previously have been reported in the *Vibrionaceae* family and sometimes are associated with fluoroquinolone resistance (*10*,*11*). However, all ST75 isolates we analyzed showed fluoroquinolone susceptibility, MIC of ciprofloxacin 0.06 μg/mL, and susceptibility to all other tested antimicrobial drugs. This pansusceptibility sharply contrasts antimicrobial resistance trends observed in 7PET isolates from Africa, which reportedly became increasingly antimicrobial resistant over time; after the 2000s, none were susceptible to antimicrobial agents (*5*).

Table 2. Features of *Vibrio cholerae* O1 ST75 isolates, South Africa, 2018–2020*

Strain no.	Serotype	Biotype	AMR phenotype	AMR gene	Plasmids	*ctxB* allele	*tcpA*	*wbeT* mutation†	Lineage‡
YA00085869	Ogawa	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	WT	L3b.1
YA00132994	Inaba	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	B08	L3b.1
YA00134463	Inaba	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	B08	L3b.1
OA01603367	Inaba	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	B08	L3b.1
YA00192016	Ogawa	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	WT	L3b.1
CF00214281	Ogawa	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	WT	L3b.1
YA00193061	Ogawa	El Tor	Pansusceptible	*qnrVC4*	None	*ctxB1*	*tcpA*^{N16961}	WT	L3b.1

*AMR, antimicrobial resistance; ST, sequence type; WT, wild-type.
†Nomenclature according to F.-X. Weill et al. (*5*).
‡Nomenclature according to H. Wang et al. (*9*).

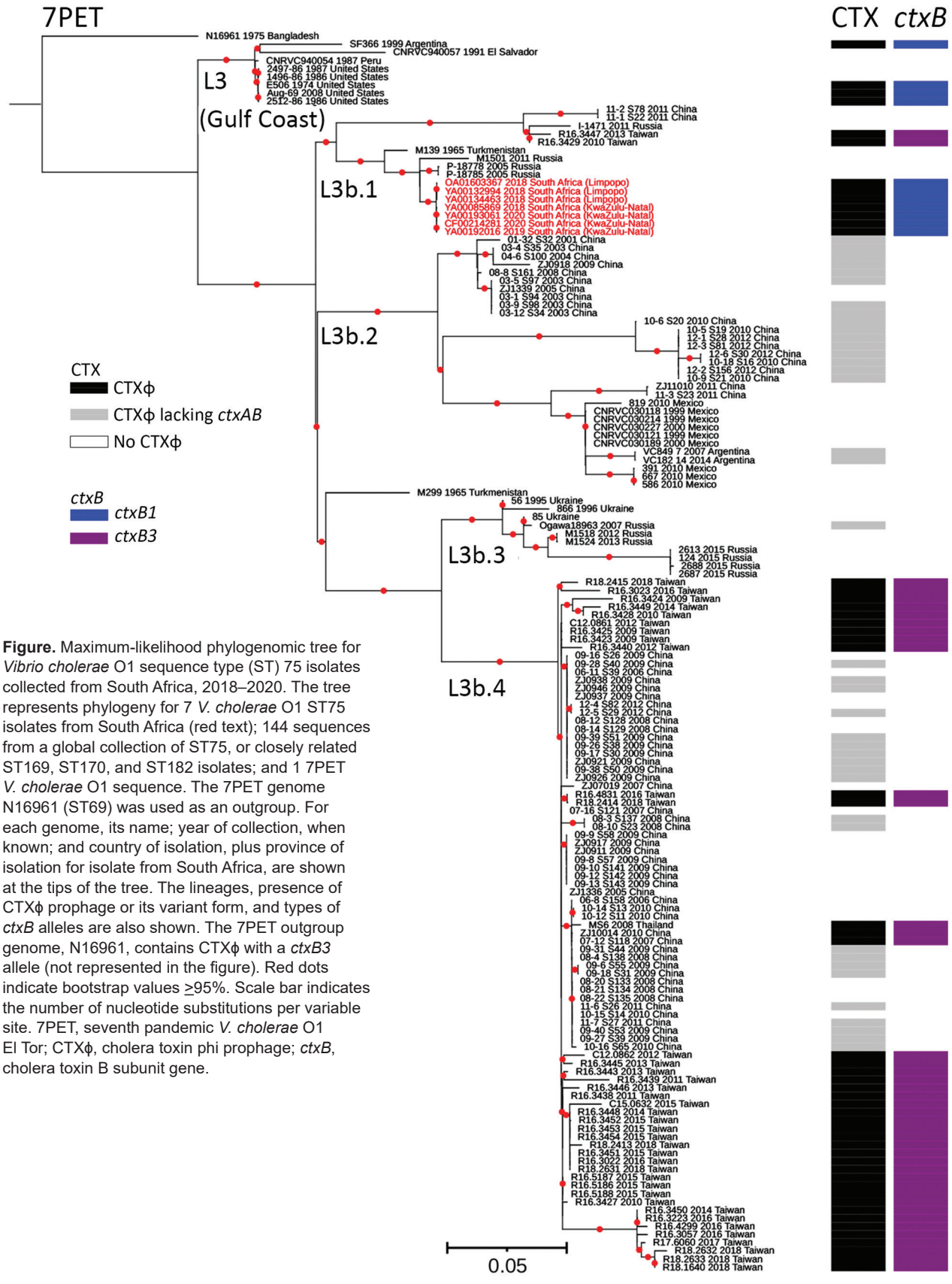


Figure. Maximum-likelihood phylogenomic tree for *Vibrio cholerae* O1 sequence type (ST) 75 isolates collected from South Africa, 2018–2020. The tree represents phylogeny for 7 *V. cholerae* O1 ST75 isolates from South Africa (red text); 144 sequences from a global collection of ST75, or closely related ST169, ST170, and ST182 isolates; and 1 7PET *V. cholerae* O1 sequence. The 7PET genome N16961 (ST69) was used as an outgroup. For each genome, its name; year of collection, when known; and country of isolation, plus province of isolation for isolate from South Africa, are shown at the tips of the tree. The lineages, presence of CTXϕ prophage or its variant form, and types of *ctxB* alleles are also shown. The 7PET outgroup genome, N16961, contains CTXϕ with a *ctxB3* allele (not represented in the figure). Red dots indicate bootstrap values ≥95%. Scale bar indicates the number of nucleotide substitutions per variable site. 7PET, seventh pandemic *V. cholerae* O1 El Tor; CTXϕ, cholera toxin phi prophage; *ctxB*, cholera toxin B subunit gene.

We further compared the ST75 isolates from South Africa with a larger global collection of 144 ST75, or closely related ST169, ST170, and ST182, genomes (Appendix 2), and constructed a maximum-likelihood phylogeny by using 49,540 SNPs (Figure). Our phylogenetic analysis showed that the 7 isolates from South Africa clustered in the L3b.1 clade, defined by H. Wang et al. (*9*), with a maximum pairwise distance of 22 SNPs. Isolates from Limpopo Province had a maximum pairwise distance of 1–6, but KwaZulu-Natal Province isolates had no SNP differences. Core-genome MLST showed Limpopo Province isolates differed from the KwaZulu-Natal Province isolates by 4–5 alleles (Appendix 1 Figure). The closest related isolates were collected in Russia from Rostov Oblast in 2005 and Republic of Kalmykia in 2011 and from Turkmenistan in Central Asia in 1965, but none of those isolates contained the CTX prophage. L3b.1 isolates from Taiwan containing the CTX prophage *ctxB3* allele were more distant.

Emergence of ST75 L3b.1 clade in South Africa is cause for concern. Recent studies on *V. cholerae* O1 isolated in Taiwan (*12*) and China (*13*) reported emerging and potential toxigenic ST75. Genomic signatures of these ST75 isolates closely resembled the US Gulf Coast *V. cholerae* O1 clone that emerged in 1973 (*14*). In particular, an investigation of *V. cholerae* O1 isolated during 2002–2018 in Taiwan showed that ST75 emerged there in 2009 and now is more prevalent than the ST69 pandemic clone (*12*). Our findings from South Africa align with the findings from Taiwan, showing that ST75 isolates outnumber ST69 isolates.

One limitation of our study is that we used reference laboratory data and a review of published *V. cholerae* O1 data to conclude that all previous cholera isolates in South Africa characterized by MLST were *V. cholerae* O1 biotype El Tor ST69. However, we cannot exclude the possibility that *V. cholerae* O1 isolates not characterized by MLST, particularly those from environmental samples, could have been non-ST69.

Epidemic 7PET lineage cholera demands an aggressive public health response to prevent outbreaks. In contrast, sporadic *V. cholerae* O1 infections mediated by other lineages, including those carrying toxin co-regulated pilus and CTX genes, typically are not epidemic-prone; most are associated with sporadic cases that rarely lead to secondary transmission (*15*). Tailoring the public health response to the degree of epidemic risk would be invaluable, especially in resource-limited settings. In countries that are not cholera-endemic but are at high risk for cholera introductions, conventional laboratory determination of *V. cholerae* O1, even complemented by identifying *ctxA* or *ctxB* genes, might be insufficient. Typing resolution of genomics, which distinguishes between 7PET and nonepidemic lineages, can elucidate the local and global epidemiology of cholera and inform public health decisions.

Conclusions

The emergence and dominance of nonepidemic, non-7PET, *V. cholerae* ST75 L3b.1 in South Africa requires close monitoring. The spatiotemporal pattern suggests local spread, possibly indicating a geographically widespread risk for sporadic disease from this strain. South Africa should strengthen its disease and environmental surveillance systems to identify nonpandemic ST75 strains, define local epidemiology, and inform an appropriate public health response.

Acknowledgments

We thank the following departments and divisions for their assistance during our study: the KwaZulu-Natal Provincial Centre for Disease Control and Prevention Directorate; health officials in King Cetshwayo, Ugu, and Umkhanyakude districts in KwaZulu-Natal Province; the Limpopo Province Department of Health; health officials in Capricorn District, Limpopo Province; National Health Laboratory Service laboratories and personnel, KwaZulu-Natal and Limpopo Provinces; Ampath Laboratories, KwaZulu-Natal Province; and Division of Public Health Surveillance and Response, National Institute for Communicable Diseases.

This study was funded by the United Kingdom Department of Health and Social Care, managed by the Fleming Fund, and performed under the auspices of the SEQAFRICA project. The Fleming Fund is a £265 million aid program supporting ≤24 low- and middle-income countries to generate, share, and use data on antimicrobial resistance. The Fleming Fund works in partnership with Mott MacDonald, the Management Agent for the Country and Regional Grants and Fellowship Programme.

About the Author

Dr. Smith is a principal medical scientist at the Centre for Enteric Diseases, National Institute for Communicable Diseases, Johannesburg, South Africa. He also holds an extraordinary professor appointment at the University of Pretoria, Pretoria, South Africa. His research interests include surveillance and epidemiology of enteric bacterial pathogens in South Africa.

References

1. Mintz ED, Tauxe RV. Cholera in Africa: a closer look and a time for action. J Infect Dis. 2013;208:S4–7. https://doi.org/10.1093/infdis/jit205
2. Küstner HG, Gibson IH, Carmichael TR, Van Zyl L, Chouler CA, Hyde JP, et al. The spread of cholera in South Africa. S Afr Med J. 1981;60:87–90.
3. Ismail H, Smith AM, Tau NP, Sooka A, Keddy KH; Group for Enteric, Respiratory and Meningeal Disease Surveillance in South Africa. Cholera outbreak in South Africa, 2008–2009: laboratory analysis of *Vibrio cholerae* O1 strains. J Infect Dis. 2013;208:S39–45. https://doi.org/10.1093/infdis/jit200
4. Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, Kariuki S, et al. Evidence for several waves of global transmission in the seventh cholera pandemic. Nature. 2011;477:462–5. https://doi.org/10.1038/nature10392
5. Weill FX, Domman D, Njamkepo E, Tarr C, Rauzier J, Fawal N, et al. Genomic history of the seventh pandemic of cholera in Africa. Science. 2017;358:785–9. https://doi.org/10.1126/science.aad5901
6. Ramamurthy T, Mutreja A, Weill FX, Das B, Ghosh A, Nair GB. Revisiting the global epidemiology of cholera in conjunction with the genomics of *Vibrio cholerae*. Front Public Health. 2019;7:203. https://doi.org/10.3389/fpubh.2019.00203
7. Irenge LM, Ambroise J, Mitangala PN, Bearzatto B, Kabangwa RKS, Durant JF, 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:e0007642. https://doi.org/10.1371/journal.pntd.0007642
8. Mashe T, Domman D, Tarupiwa A, Manangazira P, Phiri I, Masunda K, et al. Highly resistant cholera outbreak strain in Zimbabwe. N Engl J Med. 2020;383:687–9. https://doi.org/10.1056/NEJMc2004773
9. Wang H, Yang C, Sun Z, Zheng W, Zhang W, Yu H, et al. Genomic epidemiology of *Vibrio cholerae* reveals the regional and global spread of two epidemic non-toxigenic lineages. PLoS Negl Trop Dis. 2020;14:e0008046. https://doi.org/10.1371/journal.pntd.0008046
10. Fonseca EL, Dos Santos Freitas F, Vieira VV, Vicente ACP. New *qnr* gene cassettes associated with superintegron repeats in *Vibrio cholerae* O1. Emerg Infect Dis. 2008;14:1129–31. https://doi.org/10.3201/eid1407.080132
11. Fonseca EL, Vicente ACP. Epidemiology of *qnrVC* alleles and emergence out of the *Vibrionaceae* family. J Med Microbiol. 2013;62:1628–30. https://doi.org/10.1099/jmm.0.062661-0
12. Tu YH, Chen BH, Hong YP, Liao YS, Chen YS, Liu YY, et al. Emergence of *Vibrio cholerae* O1 sequence type 75 in Taiwan. Emerg Infect Dis. 2020;26:164–6. https://doi.org/10.3201/eid2601.190934
13. Luo Y, Octavia S, Jin D, Ye J, Miao Z, Jiang T, et al. US Gulf-like toxigenic O1 *Vibrio cholerae* causing sporadic cholera outbreaks in China. J Infect. 2016;72:564–72. https://doi.org/10.1016/j.jinf.2016.02.005
14. Wachsmuth IK, Bopp CA, Fields PI, Carrillo C. Difference between toxigenic *Vibrio cholerae* O1 from South America and US Gulf Coast. Lancet. 1991;337:1097–8. https://doi.org/10.1016/0140-6736(91)91744-F
15. Domman D, Quilici ML, Dorman MJ, Njamkepo E, Mutreja A, Mather AE, et al. Integrated view of *Vibrio cholerae* in the Americas. Science. 2017;358:789–93. https://doi.org/10.1126/science.aao2136

Address for correspondence: Anthony Smith, Centre for Enteric Diseases, National Institute for Communicable Diseases, Private Bag X4, Sandringham, 2131, Johannesburg, South Africa; email: anthonys@nicd.ac.za

Article DOI: https://doi.org/10.3201/eid2711.211144

Emergence of *Vibrio cholerae* O1 Sequence Type 75, South Africa, 2018–2020

Appendix 1

Materials and Methods

Phenotypic Characterization

We confirmed the identification of *Vibrio cholerae* isolates by using standard phenotypic microbiological identification and serotyping techniques. In brief, we subcultured bacteria onto 5% blood agar (National Health Laboratory Service, Diagnostic Media Products, https://www.nhls.ac.za) and thiosulfate citrate bile salts sucrose agar (Diagnostic Media Products) to check for purity of cultures. We identified bacteria by using the VITEK-2 COMPACT 15 (bioMérieux, https://www.biomerieux.com) automated microbial identification system. We determined serogrouping and serotyping by the slide agglutination method with polyvalent antisera and monospecific Inaba and Ogawa antisera (Mast Group Ltd, https://www.mast-group.com).

We performed antimicrobial susceptibility testing as follows. We determined the MIC of ampicillin, amoxicillin/clavulanate, cefepime, trimethoprim/sulfamethoxazole, chloramphenicol, ciprofloxacin, tetracycline, kanamycin, streptomycin, imipenem, and azithromycin by using the ETEST method (bioMérieux). We used the Clinical and Laboratory Standards Institute (CLSI) interpretative criteria for antimicrobial susceptibility testing of *Vibrio* species (M45) (*1*), when available. For antimicrobial drugs not listed on this *Vibrio* species in M45, we used the CLSI interpretative criteria for *Enterobacteriaceae*/*Salmonella* species (M100) (*2*). Interpretive criteria for streptomycin are not available in either CLSI document; thus, we used susceptibility ≤16 μg/mL and resistance ≥32 μg/mL.

PCR for Toxin Detection and Biotyping

We used real-time PCR to detect cholera toxin, *ctxA* gene (*3*). We used conventional PCR and analyzed PCR products using agarose gel electrophoresis to detect the presence of allelic variants of the toxin co-regulated pilus, *tcpA* gene, which determined the biotype as classical or El Tor type of *V. cholerae* O1 (*4*).

Genomic DNA Isolation from Bacteria and Whole-Genome Sequencing

We isolated genomic DNA from bacteria by using an Invitrogen PureLink Microbiome DNA Purification Kit (ThermoFisher Scientific, https://www.thermofisher.com). We performed whole-genome sequencing (WGS) by using NextSeq (Illumina, https://www.illumina.com) next-generation sequencing technology, and prepared DNA libraries by using a Nextera DNA Flex Library Preparation Kit (Illumina), before running 2 × 150 bp paired-end sequencing runs with ≈80 times coverage.

Additional Genomic Data

We downloaded and included raw sequence files from 71 *V. cholerae* genomes described by D. Domman et al. (*5*), A.R. Reimer et al. (*6*), Y. Luo et al. (*7*), S.S. Watve et al. (*8*), and Y.H. Tu et al. (*9*) from the European Nucleotide Archive (https://www.ebi.ac.uk/ena/browser/home) in the study (Appendix 2, https://wwwnc.cdc.gov/eid/article/27/11/21-1144-App2.xlxs). We downloaded and included 74 unpublished, complete, or assembled genomes described by J.F. Heidelberg et al. (*10*), K. Okada et al. (*11*), D. Hu et al. (*12*), K.V. Kuleshov et al. (*13*), H. Wang et al. (*14*) from GenBank in this study (Appendix 2).

Genomic Sequence Analyses

We used FqCleanER version 3.0 (https://gitlab.pasteur.fr/GIPhy/fqCleanER) to eliminate adaptor sequences (*15*), correct sequencing errors (*16*), and discard low-quality short-reads. We generated assemblies by using SPAdes version 3.15.0 (*17*).

We performed in silico multilocus sequence typing (MLST) on assembled sequences for the entire *V. cholerae* dataset by using MLST version 2.0 (https://cge.cbs.dtu.dk/services/MLST/) (*18*). We analyzed various genetic markers by using BLAST version 2.2.26 (https://blast.ncbi.nlm.nih.gov) against reference sequences of the CTX prophage (CTX-1, GenBank accession no. AE003852, coordinates 1566967–1573281; CTX-2, GenBank accession no. CP001486, coordinates 852233–858550), the *ctxB* gene (GenBank accession no. AE003852, coordinates 1566967–1567341), the toxin co-regulated pilus (TCP) genes (GenBank accession no. AE003852, coordinates 890449–891123), *Vibrio* pathogenicity island (VPI) 1 (VPI-1, GenBank accession no. AE003852, coordinates 873242–914124), VPI-2 (GenBank accession no. AE003852, coordinates 1896092- 1952861), *Vibrio* seventh pandemic island (VSP) I (VSP-I, GenBank accession no. AE003852, coordinates 175343–189380), and VSP-II (GenBank accession no. AE003852, coordinates 523156–550021). We used VCGIDB (http://leb.snu.ac.kr/vcgidb/index) to determine the presence and type of genomic islands and ResFinder v4.0.1 (https://cge.cbs.dtu.dk/services/ResFinder) to determine acquired antimicrobial resistance genes.

Phylogenetic Analysis

We performed phylogenetic analysis on the 7 *V. cholerae* O1 ST75 isolates from this study and 145 additional genomes. We mapped the paired-end reads and genome assemblies onto the reference genome of *V. cholerae* O1 El Tor N16961, also known as A19 (GenBank accession nos. LT907989 and LT907990) by using Snippy version 4.1.0 and BWA version 0.7.17 (https://github.com/tseemann/snippy). We called SNPs by using Snippy version 4.6.0 and Freebayes version 1.3.2 (https://github.com/tseemann/snippy) under the following constraints: mapping quality of 60, minimum base quality of 13, minimum read coverage of 4, and a 75% read concordance at a locus for a variant to be reported. We aligned core genome SNPs in Snippy version 4.1.0 for phylogeny inference. We masked repetitive regions, such as insertion sequences and the TLC-RS1-CTX region, and recombinogenic VSP-II regions in the alignment (*19*). We built a maximum-likelihood phylogenetic tree from 49,540 aligned chromosomal SNPs, by using RAxML version 8.2.12, under the general time-reversible model with 200 bootstraps (*20*). The final tree was rooted on the N16961 genome, which we visualized by using Interactive Tree of Life (iTOL) version 6 (https://itol.embl.de) (*21*).

Core-Genome Multilocus Sequence Typing

We uploaded and investigated raw genomic sequencing data from FastQ files of paired-end reads to the EnteroBase web-based platform (http://enterobase.warwick.ac.uk/species/index/vibrio). We used the EnteroBase core-genome multilocus sequence typing (cgMLST) tool to perform a genomic comparison of isolates. The cgMLST scheme (Vibrio cgMLST + HierCC) incorporates 1,128 core genes. We displayed phylogenetic cluster analysis of cgMLST data by using a GrapeTree-generated minimum spanning tree with the MSTree V2 algorithm (https://bitbucket.org/enterobase/enterobase-web/wiki/GrapeTree) (*22*). We deposited short-read sequence data in the NCBI (https://www.ncbi.nlm.nih.gov) Sequence Read Archive under BioProject identification no. PRJEB39740.

All isolates showed $\leq$4 allele differences when compared against each other, indicative of highly genetically related isolates. We used the cgMLST hierarchical cluster tool of EnteroBase to further interrogate the isolates. At hierarchical cluster level 5 (HC5), where isolates are clustered at 5 allele differences, all isolates grouped into a single HC5:1000. At HC2, the isolates are split into 2 clusters: HC2:2146, which included isolates recovered in KwaZulu-Natal Province; and HC2:1003, which included isolates recovered in Limpopo Province. This finding suggests that the isolates from KwaZulu-Natal Province are slightly genetically different to those from Limpopo Province.

References

1. Clinical and Laboratory Standards Institute. Methods for antimicrobial dilution and disk susceptibility testing of infrequently isolated or fastidious bacteria; approved guideline, third edition M45). The Institute: Wayne (PA); 2015.

2. Clinical and Laboratory Standards Institute. Performance standards for antimicrobial susceptibility testing; twenty-seventh informational supplement (M100–S27). The Institute: Wayne (PA): 2017.

3. Koskela KA, Matero P, Blatny JM, Fykse EM, Olsen JS, Nuotio LO, et al. A multiplatform real-time polymerase chain reaction detection assay for *Vibrio cholerae*. Diagn Microbiol Infect Dis. 2009;65:339–44. PubMed https://doi.org/10.1016/j.diagmicrobio.2009.07.009

4. Keasler SP, Hall RH. Detecting and biotyping *Vibrio cholerae* O1 with multiplex polymerase chain reaction. Lancet. 1993;341:1661. PubMed https://doi.org/10.1016/0140-6736(93)90792-F

5. Domman D, Quilici ML, Dorman MJ, Njamkepo E, Mutreja A, Mather AE, et al. Integrated view of *Vibrio cholerae* in the Americas. Science. 2017;358:789–93. PubMed https://doi.org/10.1126/science.aao2136

6. Reimer AR, Van Domselaar G, Stroika S, Walker M, Kent H, Tarr C, et al.; V. cholerae Outbreak Genomics Task Force. Comparative genomics of *Vibrio cholerae* from Haiti, Asia, and Africa. Emerg Infect Dis. 2011;17:2113–21. PubMed https://doi.org/10.3201/eid1711.110794

7. Luo Y, Octavia S, Jin D, Ye J, Miao Z, Jiang T, et al. US Gulf–like toxigenic O1 *Vibrio cholerae* causing sporadic cholera outbreaks in China. J Infect. 2016;72:564–72. PubMed https://doi.org/10.1016/j.jinf.2016.02.005

8. Watve SS, Chande AT, Rishishwar L, Mariño-Ramírez L, Jordan IK, Hammer BK. Whole-genome sequences of 26 *Vibrio cholerae* isolates. Genome Announc. 2016;4:e01396. PubMed https://doi.org/10.1128/genomeA.01396-16

9. Tu YH, Chen BH, Hong YP, Liao YS, Chen YS, Liu YY, et al. Emergence of *Vibrio cholerae* O1 sequence type 75 in Taiwan. Emerg Infect Dis. 2020;26:164–6. PubMed https://doi.org/10.3201/eid2601.190934

10. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, et al. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. Nature. 2000;406:477–83. PubMed https://doi.org/10.1038/35020000

11. Okada K, Na-Ubol M, Natakuathung W, Roobthaisong A, Maruyama F, Nakagawa I, et al. Comparative genomic characterization of a Thailand-Myanmar isolate, MS6, of *Vibrio cholerae* O1 El Tor, which is phylogenetically related to a "US Gulf Coast" clone. PLoS One. 2014;9:e98120. PubMed https://doi.org/10.1371/journal.pone.0098120

12. Hu D, Liu B, Feng L, Ding P, Guo X, Wang M, et al. Origins of the current seventh cholera pandemic. Proc Natl Acad Sci U S A. 2016;113:E7730–9. PubMed https://doi.org/10.1073/pnas.1608732113

13. Kuleshov KV, Vodop'ianov SO, Markelov ML, Dedkov VG, Kermanov AV, Kruglikov VD, et al. Draft genome sequences of *Vibrio cholerae* O1 El Tor strains 2011EL-301 and P-18785, isolated in Russia. Genome Announc. 2013;1:e00659. PubMed https://doi.org/10.1128/genomeA.00659-13

14. Wang H, Yang C, Sun Z, Zheng W, Zhang W, Yu H, et al. Genomic epidemiology of *Vibrio cholerae* reveals the regional and global spread of two epidemic non-toxigenic lineages. PLoS Negl Trop Dis. 2020;14:e0008046. PubMed https://doi.org/10.1371/journal.pntd.0008046

15. Criscuolo A, Brisse S. AlienTrimmer: a tool to quickly and accurately trim off multiple short contaminant sequences from high-throughput sequencing reads. Genomics. 2013;102:500–6. PubMed https://doi.org/10.1016/j.ygeno.2013.07.011

16. Liu Y, Schröder J, Schmidt B. Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. Bioinformatics. 2013;29:308–15. PubMed https://doi.org/10.1093/bioinformatics/bts690

17. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77. PubMed https://doi.org/10.1089/cmb.2012.0021

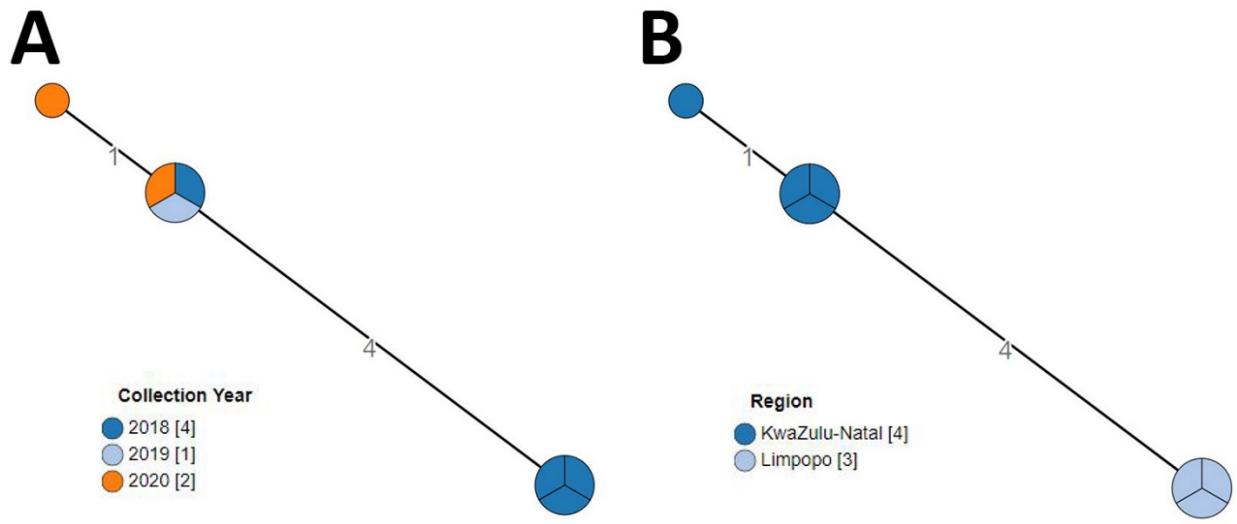
18. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, et al. Multilocus sequence typing of total-genome-sequenced bacteria. J Clin Microbiol. 2012;50:1355–61. PubMed https://doi.org/10.1128/JCM.06094-11

19. Weill FX, Domman D, Njamkepo E, Tarr C, Rauzier J, Fawal N, et al. Genomic history of the seventh pandemic of cholera in Africa. Science. 2017;358:785–9. PubMed https://doi.org/10.1126/science.aad5901

20. Stamatakis A. RAxML-VI-HPC: maximum likelihood–based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22:2688–90. PubMed https://doi.org/10.1093/bioinformatics/btl446

21. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49(W1):W293–6. PubMed https://doi.org/10.1093/nar/gkab301

22. Zhou Z, Alikhan NF, Sergeant MJ, Luhmann N, Vaz C, Francisco AP, et al. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. Genome Res. 2018;28:1395–404. PubMed https://doi.org/10.1101/gr.232397.117



Appendix 1 Figure. Minimum spanning tree drawn by using cgMLST data from 7 *Vibrio cholerae* O1 sequence type 75 isolates collected in South Africa, 2018–2020. A) Isolate collection year; B) region of isolate collection. Circular nodes represent isolates having identical cgMLST profiles; the larger the node, the more isolates included. Number values between adjacent nodes indicate the number of allele differences between nodes. cgMLST, core-genome multilocus sequence typing.